Industry Partners make a difference:

Assessing the Impact of

Problem-Based Learning on

Critical and Creative Thinking

in Primary School

by

Susan Gaardboe

BAppSci, PGCE (UK)

A thesis submitted for the degree of

Doctor of Philosophy

UniSA STEM



University of
South Australia

March 2024

Contents

Contents ..... **iii**

List of Figures ..... **xii**

List of Tables ..... **xiii**

Table of Abbreviations ..... **xvi**

Summary ..... **xvii**

Declaration ..... **xix**

Acknowledgements ..... **xx**

Dedication ..... **xxi**

Chapter 1. Can You Prove It? The Role of Problem-Based Learning and Industry Partners, in the Application of Critical and Creative Thinking Skills **1**

1.1. Research Background and Motivation *1*

1.2. Research Questions *3*

1.3. Methodology *7*

1.4. Structure of the Thesis *7*

1.5. Contributions of the Thesis *10*

Curriculum Writers and Developers 10

Teachers and the Primary School Sector 11

Researchers and the University Sector 12

Part A. Making Sense of Skill Development and Problem-Based Learning **14**

A.1. Methodology of Conceptual Studies *14*

Chapter 2. Investigating Critical and Creative Thinking in the Australian Curriculum **17**

2.1. Background to the National Curriculum *17*

2.2. The Structure of the Australian Curriculum *19*

2.3. The Theories Behind the Critical and Creative Thinking Skills *24*

2.4. Theoretical Basis to the 'Levels' in the Continuum *27*

2.5. The Accuracy of the Learning Continuum *28*

2.6. Conclusions Regarding the Continuum *32*

Chapter 3. Theoretical Foundations and Measurement of Critical and Creative Thinking..... 34

3.1. Introduction..... *34*

3.2. General Concepts of Critical Thinking *34*

3.3. Comparing Critical Thinking Theories with the Critical and Creative Thinking Continuum *38*

3.4. Measuring Critical Thinking *44*

3.5. General Concepts of Creativity and Creative Thinking..... *48*

3.6. Models of Creativity *50*

3.7. The 4Ps and Phase Models Combined..... *52*

3.8. Creative Thinking in Education..... *55*

3.9. The Relationship Between Creative Thinking and Critical Thinking..... *58*

3.10. The Domain Specific Nature of Critical and Creative Thinking..... *59*

3.11. Measuring Creative Thinking *60*

3.12. Problem-Solving Assessment Tasks..... *64*

3.13. Choice of Assessment Task to Measure Critical and Creative Thinking *66*

3.14. Current Gaps in Knowledge..... *67*

Chapter 4. Theoretical Foundations and Measurement of the Effect of Problem-Based Learning 69

4.1. Introduction..... *69*

4.2. What is Problem-Based Learning?..... *69*

4.3. PBL and Critical and Creative Thinking *76*

4.4. PBL and Motivation and Engagement *78*

4.5. The Problem in Problem-Based Learning *80*

4.6. The Impact of Industry Partners in PBL..... *82*

4.7. Current Gaps in Knowledge *83*

Chapter 5. Theoretical Foundations and Measurement of Motivation and Engagement **84**

5.1. Introduction *84*

5.2. What is Motivation? *85*

5.3. What is Engagement? *88*

5.4. Measures of Motivation and Engagement *91*

5.5. Choice of Task to Measure Motivation and Engagement *94*

5.6. Current Gaps in Knowledge *94*

Chapter 6. Conclusion to Part A **96**

Part B. Investigating Problem-Based Learning and Skill Application in the Classroom **100**

B.1. Methodology of Empirical Studies *100*

B.2. Research Design and Methods *101*

Chapter 7. Study One: Initial Development and Pilot Study of a Measure of Critical and Creative Thinking **105**

7.1. Introduction *105*

7.2. Research Design and Methods *107*

Scope of Research 107

Research Design 107

7.3. Step One: Development of Statements *108*

7.4. Step Two: Measurement Validity Testing *113*

7.5. Step Three: Presentation and Response Format *115*

Student Assessment Task 1 117

Student Assessment Task 2 117

The Teacher Task 117

Presentation of the Measure 118

7.6. Step Four: Data Collection *119*

Participants 119

Measures 120

7.7. Data Collection Approach *120*
Method for Statistical Analysis 122
7.8. Results of Data Collection and Analysis *123*
Statistical Analysis of Measures 123
Response by Students to Characteristics of the Measures 129
Response by Students to Presentation of the Measure 129
Response by Students to the Response Formats 130
7.9. Study One Discussion *130*
Response Type 131
Teacher Task 131
Student Assessment Task 1 132
Student Assessment Task 2 133
7.10. Limitations of Study One *141*
7.11. Conclusions from Study One *141*
Chapter 8. Study Two: Two Alternatives for a Novel Assessment of Critical and Creative Thinking 144
8.1. Introduction *144*
8.2. Impact of Presentation Formats *145*
8.3. Theories of Multimedia Learning *145*
8.4. Study Two Research Design and Methods *148*
Validity and Reliability 148
Measures 148
Procedure 149
Step One: Review Assessment Task 150
Step Two: Data Collection 155
Step Three: Data Analysis 160
Data Analysis Approach 161
8.5. Results *164*
Preliminary Analysis 164
Identification of Outliers 165
8.6. Investigation of Reasons for Outliers *169*

8.7. Comparison of Presentation Formats *171*

8.8. Discussion *172*

Initial Test analysis 173

Investigating the Possible Causes of Outliers 173

Causes for Low Rate of Agreement 174

Preferred Assessment Task 180

8.9. Limitations, Constraints and Recommendations *181*

8.10. Conclusions from Study Two *182*

Chapter 9. Study Three: The Impact of Problem-Based Learning and Industry Partners on Critical and Creative Thinking 183

9.1. Introduction *183*

9.2. Method *187*

Research Design 187

Participants 189

Measures 191

Procedure 192

Method for Statistical Analysis 203

9.3. Results of Statistical Analysis *205*

Critical and Creative Thinking Statistical Results 206

Motivation and Engagement Statistical Results 211

9.4. Discussion *214*

New Insights into Critical and Creative Thinking Skills: H 3a(i) and H 3b(i) 214

The Long-Term Effect of PBL on Critical and Creative Thinking: H 3a(ii) and H 3b(ii). 219

New Insights into Motivation and Engagement: H 4a(i), and H 4b(i) 219

The Long-Term Effect of PBL on Motivation and Engagement: H 4a(ii) and H 4b(ii). 220

The Benefits of PBL and Industry Partners 225

9.5. Limitations, Constraints and Recommendations. *225*

9.6. Conclusions from Study Three *227*

Chapter 10. Conclusions from Part B: Investigating Problem-Based Learning and Skill Development in the Classroom 229

10.1. Introduction *229*

10.2. Contributions *230*

10.3. Addressing Gaps in the Literature *230*

10.4. Limitations of the Studies and Recommendations *233*

Chapter 11. Can you Prove it? **234**

References **237**

Appendix A STEM Learning Evaluation Grant Report **274**

Appendix B Measures of Critical Thinking **297**

Appendix C Measures of Creativity, Creative Potential, Creative Thinking and Creative Problem Solving **301**

Appendix D Measures of Motivation and Engagement **311**

Appendix E Search Outlines **316**

Appendix F Study One Supporting Documents **318**

Appendix F-1. Study One. Assessment Task One Script *318*

Appendix F-2. Study One. Student Assessment Task One *324*

Appendix F-3. Study One. Student Assessment Task Two *326*

Appendix F-4. Study One. Teacher Task *330*

Appendix F-5. Study One. Teacher Survey *332*

Appendix F-6. Study One. Teacher Information *334*

Appendix F-7. Study One. Participant Information *338*

Appendix G Study Two Supporting Documents **343**

Appendix G-1. Study Two. Critical and Creative Thinking measure Script *343*

Appendix G-2. Study Two. Student Assessment. Presentation Format One *350*

Appendix G-3. Study Two. Student Assessment. Presentation Format Two *352*

Appendix G-4. Study Two. Teacher Task *356*

Appendix H Study Two Tests of Normality 359

Table H-1. Study Two. Tests of normality for critical and creative thinking presentation format - student responses 359

Table H-2. Study Two. Tests of normality for critical and creative thinking presentation format - teacher responses 361

Appendix I Study Two Fleiss' Kappa Analysis 363

Table I-1. Study Two. Results of preliminary analysis with Fleiss' Kappa comparing teacher and student responses 363

Table I-2. Study Two. Fleiss' Kappa results for individual teachers 365

Table I-3. Study Two. Teacher variables: number of years teaching 371

Table I-4. Study Two. Teacher variables: number of years teaching the current year level 373

Table I-5. Study Two. Teacher variables: Has previously used the AC C&CrT continuum to assess students 375

Table I-6. Study Two. Teacher variables: Confidence to use the AC C&CrT continuum to assess students 377

Table I-7. Study Two. Student variables: Year level of students 379

Table I-8. Study Two. Student variables: Sex of students 381

Table I-9. Study Two. Student variables: Learning complexities – Presentation Format 1 383

Table I-10. Study Two. Student variables: Learning complexities - Presentation Format 2 385

Table I-11. Study Two. Results of analysis with Fleiss' Kappa following removal of outliers 387

Appendix J Study Three Supporting Documents 389

Appendix J-1. Study Three. Teacher Survey 389

Appendix J-2. Study Three. Critical and Creative Thinking Self-Report Measure 390

Table J-3. Study Three. Motivation and engagement statements 396

Appendix J-4. Study Three. Teacher post-study questionnaire 400

Appendix J-5. Study Three. Participation information and consent form 401

Appendix J-6. Study Three. STEM Problem-Based Learning teacher training programme and resources 412

Appendix J-7. Study Three. PBL problems provided for the intervention *419*

Appendix K Study Three Data Analysis 423

Table K-1. Study Three. Tests of normality for critical and creative thinking assessment, control group: Asking Questions and Finding and Organising Information *423*

Table K-2. Study Three. Tests of normality for critical and creative thinking assessment, control group: Connecting Information and Creating New Ideas *425*

Table K-3. Study Three. Tests of normality for critical and creative thinking assessment, control group: Evaluating a Plan or Idea *427*

Table K-4. Study Three. Tests of normality for critical and creative thinking assessment, theory group: Asking Questions and Finding and Organising Information *428*

Table K-5. Study Three. Tests of normality for critical and creative thinking assessment, theory group: Connecting Information and Creating New Ideas *430*

Table K-6. Study Three. Tests of normality for critical and creative thinking assessment, theory group: Evaluating a Plan or Idea *432*

Table K-7. Study Three. Tests of normality for critical and creative thinking assessment, industry partner group: Asking Questions and Finding and Organising Information *433*

Table K-8. Study Three. Tests of normality for critical and creative thinking assessment, industry partner group: Connecting Information and Creating New Ideas *435*

Table K-9. Study Three. Tests of normality for critical and creative thinking assessment, industry partner group: Evaluating a Plan or Idea *437*

Table K-10. Study Three. Tests of normality for motivation and engagement scale: Total positive motivation and total positive engagement *438*

Table K-11. Study Three. Tests of normality for motivation and engagement scale: Total negative motivation and total negative engagement *439*

Table K-12. Study Three. Motivation and engagement scale analysis for internal consistency – time points compared, all schools combined *440*

Table K-13. Study Three. Motivation and engagement scale analysis for internal consistency – schools compared, Time Point 1 ..... *441*

Table K-14. Study Three. Motivation and engagement scale analysis for internal consistency – schools compared, Time Point 2 ..... *442*

Table K-15. Study Three. Motivation and engagement scale analysis for internal consistency – schools compared, Time Point 3 ..... *443*

Table K-16. Study Three. Motivation and engagement scale analysis for internal consistency – EALD/ Non-EALD students compared – Time Point 1..... *444*

Table K-17. Study Three. Motivation and engagement scale analysis for internal consistency – EALD/ Non-EALD students compared – Time Point 2..... *445*

Table K-18. Study Three. Motivation and engagement scale analysis for internal consistency – EALD/ Non-EALD students compared – Time Point 3..... *446*

Table K-19. Study Three. Cochran's Q analysis of critical and creative thinking – control group... *447*

Table K-20. Study Three. Cochran's Q analysis of critical and creative thinking – PBL group *449*

Table K-21. Study Three. Cochran's Q analysis of critical and creative thinking – theory group.... *451*

Table K-22. Study Three. Cochran's Q analysis of critical and creative thinking – industry partner group..... *452*

Table K-23. Study Three. Friedman test analysis of motivation and engagement survey – control group and PBL group..... *453*

Table K-24. Study Three. Friedman test analysis of motivation and engagement survey – theory group and industry partner group *454*

List of Figures

Figure 1 *Positioning the research focus* 4
Figure 2 *The structure of the Australian Curriculum* 19
Figure 3 *Terminology of the structure of the Critical and Creative Thinking learning continuum (V8.4)* 21
Figure 4 (over) *The Australian Curriculum, Critical and Creative Thinking learning continuum* 21
Figure 5 *Comparison of Blooms taxonomy levels with the C&CrT continuum* 29
Figure 6 *Continuity of skills descriptions in sub-element 'Imagine possibilities and connect ideas'* 31
Figure 7 *A systems model of creativity* 49
Figure 8 *Relationships between the terms critical and creative thinking, divergent and convergent thinking* 59
Figure 9 *Relationship between three inquiry-based approaches: Inquiry Learning, Problem-Based Learning and Project-Based Learning* 70
Figure 10 *Research design for Study One* 108
Figure 11 *Terminology of the structure of the Critical and Creative Thinking learning continuum (V8.4)* 109
Figure 12 *Data collection activities: Study One* 121
Figure 13 *Illustrating the effect of introducing new topics at different year levels* 136
Figure 14 *Illustrating the effect of introducing new topics building on previous knowledge* 138
Figure 15 *Research design: Study Two* 149
Figure 16 *Process of data analysis* 159
Figure 17 *Individual teachers identified as outliers* 165
Figure 18 *Research design: Study Three* 188
Figure 19 *The Engineering Design Process* 195

List of Tables

Table 1 *Research Questions* 5
Table 2 *The origins of the skills in the continuum* 24
Table 3 (over) *Comparison of terminology for Australian Curriculum V8.4 C&CrT continuum sub-elements with Ritchhart et al.'s 'Making Thinking Visible'* 25
Table 4 *Critical thinking skills and dispositions identified by the Delphi Report* 36
Table 5 *Comparison of terminology for critical thinking skills between a selection of theories at the level of 'elements' in the Australian Curriculum* 39
Table 6 *Comparison of terminology for critical thinking skills between the C&CrT continuum, Delphi Report, and OECD critical thinking rubric* 41
Table 7 *Criteria for suitability for measures of Critical Thinking, used in the primary school classroom* 45
Table 8 *Measures of critical thinking skills* 46
Table 9 *Comparison of three creative process models* 51
Table 10 *The Innovative Phase Model structure* 53
Table 11 *Aligning the Extended Phase Model with convergent and divergent thinking* 53
Table 12 *Comparison of education definitions of creative thinking (CrT) with the psychology understanding of creative thinking* 57
Table 13 *Revised Criteria for suitability for measures of critical and creative thinking skills, used in the primary school classroom* 61
Table 14 *A creativity assessment comparison grid with examples of measures* 63
Table 15 *Classification of creativity measures* 65
Table 16 *Research questions (1st revision)* 68
Table 17 *Characteristics defining Inquiry Learning, Problem-Based Learning and Project-Based Learning* 72
Table 18 (over) *Comparison between features of inquiry-based approaches* 73
Table 19 *Criteria for suitability for measures on motivation and engagement* 92
Table 20 *Motivation and engagement measures validated for application in primary school* 93
Table 21 *Findings of the conceptual study — Research Questions 1 and 2* 97
Table 22 *Findings of the conceptual study — Research Questions 3 and 4* 98
Table 23 *Revised research questions 2a, 3a, 3b, 4a and 4b, and associated hypotheses* 99
Table 24 (over) *Critical and creative thinking measure statements and their categories* 109
Table 25 *Main features of the components of the critical and creative thinking measure* 116

Table 26 *Participants in Study One, Pilot Study* 121
Table 27 *Student responses to Assessment Tasks 1 and 2* 124
Table 28 *Study One. Results of exact sign test, Assessment Task 2* 127
Table 29 *SOLO's developmental modes* 135
Table 30 *SOLO's levels of understanding* 135
Table 31 *Examples of SOLO questioning applied to year level topics* 137
Table 32 *Examples of SOLO questioning applied to a subject* 139
Table 33 *Recommendations for further investigation* 143
Table 34 *Recommendations from Study One* 151
Table 35 *Reworded statements in assessment task* 152
Table 36 *Teacher participants in Study Two* 156
Table 37 *Student participants in Study Two* 157
Table 38 *Description of Kappa levels* 163
Table 39 *Summary of Fleiss' Kappa results showing number of results with positive Kappa for individual teachers (Part 1)* 166
Table 40 *Summary of results of analysis with all teachers and students, comparing a variety of variables* 170
Table 41 *Summary of results of analysis comparing students with and without learning issues* 171
Table 42 *Summary showing statistically significant results of analysis comparing Presentation Format 1 and Presentation Format 2 following removal of outliers* 172
Table 43 *Participants in Study Three* 191
Table 44 *Sample lesson outline for 10 week STEM PBL unit* 196
Table 45 *Observations of variations in interventions in different schools* 201
Table 46 *Analysis of critical and creative thinking comparing control and PBL groups* 209
Table 47 *Analysis of critical and creative thinking comparing theory and industry partner groups* 210
Table 48 *Analysis of motivation and engagement higher order factors comparing control and PBL groups* 212
Table 49 *Analysis of motivation and engagement higher order factors comparing theory and IP PBL groups* 213
Table 50 *Research Question 3 and sub-questions 3a and 3b with associated hypotheses and results* 215
Table 51 *Research Question 4 and sub-question 4a and 4b, with associated hypotheses and results* 216

Table 52 *Analysis of motivation and engagement first order factors comparing control and PBL groups* 221
Table 53 *Analysis of motivation and engagement first order factors comparing control and PBL groups* 222

Table of Abbreviations

AC	Australian Curriculum
ACARA	Australian Curriculum Assessment and Reporting Authority
ACER	Australian Council for Educational Research
C&CrT	Critical and Creative Thinking
C&CrT continuum	Critical and Creative Thinking learning continuum
DECD	Department for Education and Child Development (South Australia)
DfE	Department for Education (South Australia)
IP	Industry Partner
OECD	Organisation for Economic Co-operation and Development
PBL	Problem-Based Learning
PISA	Programme for International Student Assessment
PjBL	Project-Based Learning
SACSA	South Australian Curriculum, Standards and Accountability framework
SOLO	Structure of the Observed Learning Outcome
STEM	Integrated learning based on Science, Technologies, Engineering and Mathematics

Summary

The overarching question, 'What is the impact of an industry partner in STEM Problem-Based Learning on the development of critical and creative thinking skills in primary school students?' provided the direction for the research described in this thesis. The research is presented in two parts: part one describes the conceptual research required to formulate the questions which were answered by the empirical research, described in part two.

The conceptual research covers four main areas. First, the Critical and Creative Thinking learning continuum in the Australian Curriculum is investigated to identify its background, structure and relevance to the use of skills by students, in the classroom. Second, critical thinking and creative thinking are examined. The two terms are clarified and defined in the light of theories of critical thinking, creativity, psychology and education. Criteria are developed to enable suitable measures of critical thinking and creative thinking to be identified for use in the classroom. Although no suitable measures are identified from the literature, this chapter provides a foundation for the development of a measure later in the study. The third area covered by the conceptual research is Problem-Based Learning (PBL). The characteristics of PBL are identified and compared with other forms of Inquiry learning. Research on the 'problem' in Problem-Based Learning, the evidence of the effect on student skills and the use of industry partners providing a STEM problem is sparce in relation to primary school students. Based on the research literature, further research questions and hypotheses are developed. Motivation and engagement are identified through the PBL research as possibly holding a key to explain the claimed success of PBL. The fourth area covered by the conceptual research is therefore motivation and engagement. Through an analysis of theories of motivation and engagement, and examination of the current measures used with students, a suitable measure is identified to assess students' levels of motivation and engagement during learning.

The second half of the thesis uses empirical research to investigate the remaining questions emerging from part one. In Study One, two measures of critical and creative thinking skills are designed and tested through a pilot study. The way in which students respond to the measures informs recommendations for an alternative structure for the Critical and Creative Thinking learning continuum. The results of Study One inform the development and testing of a Critical and Creative Thinking Self-Report Measure in Study Two. Two presentation formats of the measure are developed to test the application of multimedia learning theory for its effect on student responses to the measure. The level of success of the two formats is gauged by comparing the teachers' assessment of each student, with each student's own self-report of their use of skills during

everyday learning. The results of this analysis bring to light the lack of agreement between the teachers and students, revealing a variation in the percentage agreement rate between teachers and students of 3% to 47%, with a surprisingly low mode of agreement rate of 10%. These results point to a worrying absence of shared understanding between teachers and students about the skills and how to recognise their use by students.

In Study Three the novel assessment task developed through empirical studies One and Two is applied in primary school classrooms pre-, post-, and in follow-up, to a STEM Problem-Based Learning intervention. The effects of the use in PBL learning of an industry partner and their complex, real-world STEM problem, are compared with a control group, and with students undergoing the STEM PBL intervention with a hypothetical problem. The positive effects of an industry partner STEM problem on the range and level of skills used by primary school students is demonstrated, over and above the skills utilised by students engaged in a hypothetical STEM problem. The lasting value of this learning, i.e., beyond the conclusion of the intervention, is also demonstrated. This study provides the first empirical data on the effectiveness of PBL in assisting primary school students to apply critical and creative thinking skills in their learning. Student motivation and engagement are measured, and improvements are demonstrated as expected. However, the details of the improvements (reductions in negative motivation and engagement) are unexpected. This leads to recommendations for further work in this area to understand what appears to be a complex relationship.

This thesis contributes knowledge to three sectors of the community: teachers and the schooling sector, particularly at a primary or elementary level of schooling; researchers and the university sector; and curriculum writers and developers.

Declaration

I declare that:

this thesis presents work carried out by myself and does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university; to the best of my knowledge it does not contain any materials previously published or written by another person except where due reference is made in the text; and all substantive contributions by others to the work presented, including jointly authored publications, are clearly acknowledged.



Susan Gaardboe

1st March, 2024

Acknowledgements

This research was supported by an Australian Government Research Training Program Scholarship, and additionally, recognised by an Australian Post-Graduate Award (APA).

I acknowledge the generous support and advice of the following people:
My supervisors – Prof David Cropley for his unstinting generosity of time and patience in the face of certain frustration; Associate Prof Simon Leonard, Dr Martin Westwell and Dr Tim Patston, for excellent lunchtime conversation and critical feedback.
Professor Denise Meredyth and Dr Diana Newport-Peace, who planted the first thought.
Dr Rebecca Marrone, Dr Kim van Broekhoven, Dr Michelle Oppert, Dr Chanvi Singh and Maria Vieira, my co-conspirators in the PhD journey.
Christie Evans and Deb Lasscock, who enabled the thought to find its voice, and provided the environment to nourish and support those fragile, early stages of developing ideas.

Mads Gaardboe without whose unfailing support and encouragement this adventure would never have begun and would certainly not have ended as it has.

Thanks also to the Department for Education SA, and Catholic Education SA for enabling me to test my ideas in the classroom, and the peer review team: Deb Lasscock, Christie Evans, Vanessa Jones, Sharon Cosh, Deb Tune, and teachers who welcomed me into their classrooms.

I also owe a huge debt of thanks to Dr Diana Newport-Peace for editorial advice.

You begin with a thought, that blossoms and develops into an idea, which becomes a concept, and from there transitions into a theory. At each point along the winding path you feel the exhilaration of growth as you see further, understand more, and reveal the next layer of the puzzle. Exhilaration is tempered, or even replaced by despondency as you reach out your hand to touch what seems so close, but now appears to recede out of reach. The winding path leads you onwards. My companions on this journey have supported and inspired me. They have provided physical, emotional and intellectual nourishment. Their contributions cannot be unravelled from the whole, nor can their influence be understated. My gratitude to them all is unbounded.

Dedication

This work is dedicated to my children, who said they wished they'd been taught like this,

and to my grandchildren, who I hope benefit from this research.

Chapter 1.

Can You Prove It? The Role of Problem-Based Learning and Industry Partners, in the Application of Critical and Creative Thinking Skills

Give a man a fish and you feed him for a day.
Teach him how to fish and you feed him for his lifetime.

Proverb

The room was quiet.
The 30 teachers attending my presentation were hanging on my words.
"What we've noticed," I said, "is that using Problem-Based Learning with an industry partner seems to help the students develop their critical and creative thinking skills."
"But does it?" said a little voice in the back of my head. "Can you prove it?"

(Department for Education STEM PBL Professional Development Day, Brighton Primary School, 19 September 2019)

1.1. Research Background and Motivation

Problem-Based Learning (PBL) came to my attention in 2015, as a pedagogical approach to teaching and learning, suitable for building engagement in STEM subjects in primary schools in Australia. This was both supported and promoted at a national and local level (Department for Education (DfE) South Australia, n.d.-b; Department for Education and Child Development (DECD), 2016, p. 6; Education Council, 2015). Schools were encouraged to "Build on students' curiosity and connect STEM learning to solving real world problems, including through collaborative and individual learning experiences that are hands-on and inquiry-based and support the achievement of deep knowledge." (Education Council, 2015). As a teacher employed in a local primary school in South Australia, it was in this environment of experimenting with STEM teaching that I trialled using Problem-Based Learning with an industry partner, to engage primary school students in STEM learning. Predictably, questions were asked about the effectiveness of the teaching methods being employed.

In 2017, the Department for Education in South Australia initiated an evaluation project to investigate the effectiveness of STEM teaching in schools — both the development of students' critical and creative thinking skills and the measures that were being used to evaluate their use. As a school-based participant in this STEM Learning Evaluation study (Department for Education (DfE) South Australia, 2017; Gaardboe, 2019 included in Appendix A), I began to investigate the efficacy of the teaching methods we were employing. This formed the impetus which drove the research described in this thesis.

The major finding in my team's component of the STEM Evaluation Study was that there was an overall positive change in the students' reported self-perception of their abilities in critical and creative thinking, following their involvement in STEM Problem-Based Learning (PBL) with an industry partner (IP)[1]. The study was accomplished without a control group, and without a comparison of problem types, which led to further questions about the factors which might be activating these changes, and how they could be accurately measured. After the project had concluded these issues remained to be explored, leading to the overarching question, 'What is the impact of an industry partner in STEM Problem-Based Learning on the development of critical and creative thinking skills in primary school students?' This question provided the motivation for the research described in this thesis.

The research in this thesis is predicated on the dual assumptions that all individuals possess the inherent potential for rationality and creativity, and that the effective use of critical and creative thinking skills is a route to acquiring knowledge and achieving success in society. This stance aligns with the proposition of the former Australian Chief Scientist and others (e.g., Office of the Chief Scientist, 2014; World Economic Forum, 2016), that to increase critical and creative thinking skills in school students will improve employability and employment success in new graduates, particularly in this era of Artificial Intelligence and the changing face of the workplace.

The overarching question, 'What is the impact of an industry partner in STEM Problem-Based Learning on the development of critical and creative thinking skills in primary school students?' was underpinned by an epistemological assumption that the cause of the development of critical and creative thinking skills could be explored through empirical observation and study. However, over the course of this research this epistemological view was refined to focus specifically on the *application* of the skills in question, as the development of skills could not in fact be

[1] The use of the term 'industry partner' refers to any partner who is external to the school. This usage aligns with Australian government and research terminology, e.g., Torii (2018); Watters et al. (2016).

empirically demonstrated at the desired level. The detailed graduations of development of the skills described in the Critical and Creative Thinking continuum in the Australian Curriculum have not been empirically tested and cannot yet be empirically measured. The unique characteristics of the research environment included the practical constraints of working in primary school classrooms through the COVID-19 pandemic. In the context of Problem-Based Learning, within the pre-defined framework of the Critical and Creative Thinking learning continuum, this necessitated a more nuanced and multifaceted approach, extending beyond the confines of empirical observation alone.

Within the overarching question there is also an assumption that critical and creative thinking skills exist as identifiable entities that can be enhanced through the specific context of STEM Problem-Based learning with industry partners. The question further assumes that the critical and creative thinking skills in the Australian Curriculum Critical and Creative Thinking learning continuum (C&CrT continuum) provide an accurate description of the development of critical and creative thinking skills, and that there is a way to measure and understand the development of these skills in the context of STEM Problem-Based learning with industry partners. These assumptions that the C&CrT continuum is accurate and that the skills described can be accurately measured, sit at the heart of the overarching question. It was imperative therefore that these assumptions were investigated, and if necessary challenged, prior to empirical studies being carried out.

Therefore, although the overarching question was initially conceived as one to which an empirical research design would be applicable, the research design incorporated a combination of conceptual research followed by empirical research.

1.2. Research Questions

The overarching question, 'What is the impact of an industry partner in STEM Problem-Based Learning on the development of critical and creative thinking skills in primary school students?' focusses on the intersection of areas relating to what our students should be able to do, how we can effectively teach it, and how we can measure what they do. It is therefore in the intersection of the following areas that my research sits: critical and creative thinking skills (what students do); PBL being used to teach primary school students, including the use of industry partners (how we teach); and the assessment of skills (how we measure what students do). The positioning of the research focus is illustrated in Figure 1. The research questions reflect the relationship between these areas of focus and are shown in Table 1. Research Questions 1 and 2 address critical and creative thinking and their intersection with assessment methods through conceptual research. Similarly, Research Question 3 is addressed initially through conceptual research and is positioned in the area of STEM

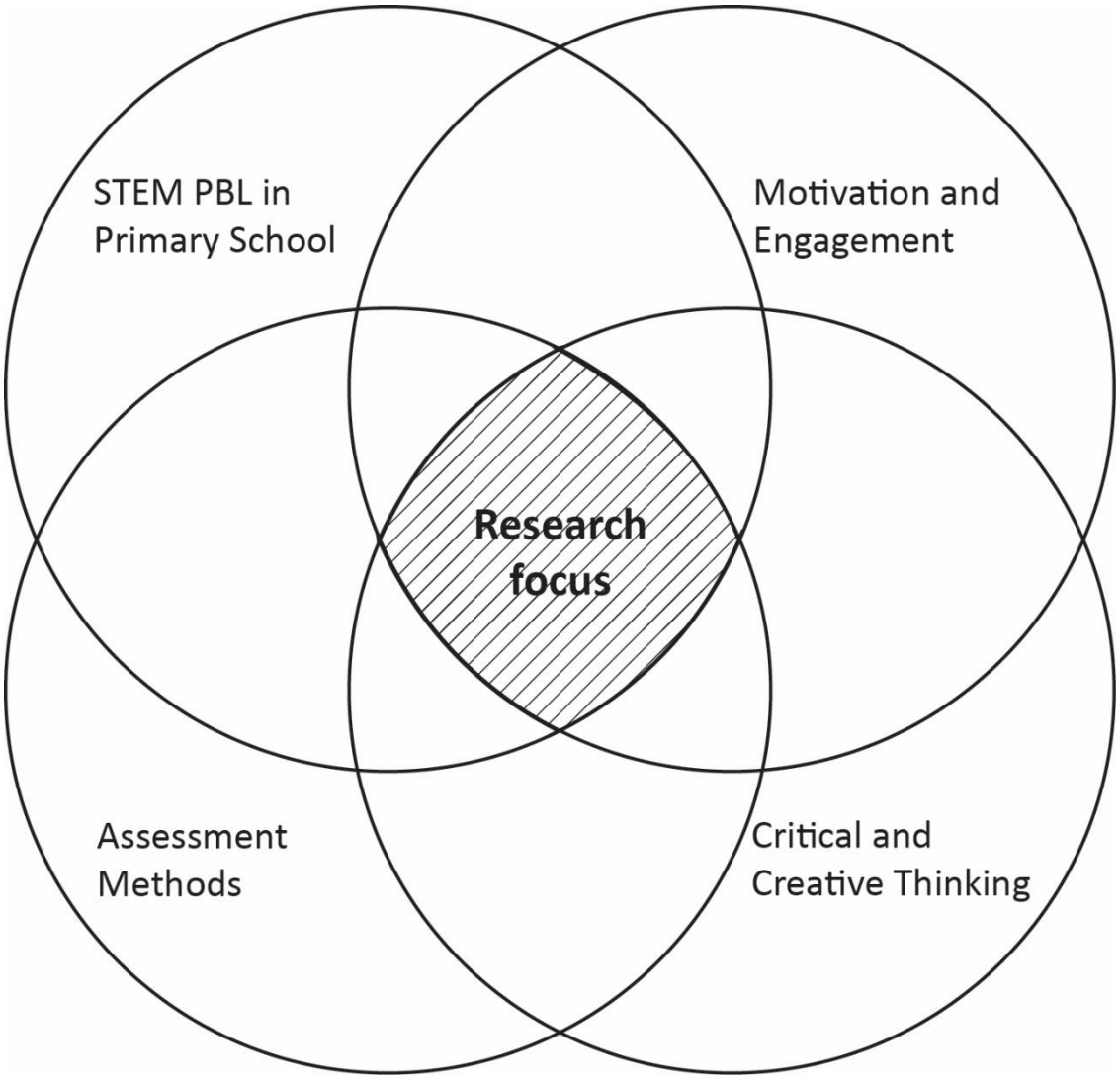


Figure 1

Positioning the research focus

Table 1

Research Questions

Overarching Question: What is the impact of an industry partner in STEM PBL on the development of critical and creative thinking skills in primary school students?	Addressed in Chapter
Research Question 1: What is the theoretical basis for the Australian Curriculum Critical and Creative Thinking learning continuum?	2
Research Question 2: How are critical and creative thinking skills measured during learning?	3, 7, 8
Research Question 2a: What should be measured to assess critical and creative thinking in the classroom during learning activities?	7
Research Question 2b: How does the method of communicating a self-report measure to students affect the accuracy of their responses?	8

Table 1 (continued)

	Addressed in Chapter
Research Question 3: How do the characteristics of STEM PBL affect the use of skills during learning?	4, 9
Research Question 3a: What is the impact of PBL on student critical and creative thinking skill application?	9
Research Question 3b: What is the impact of using an industry partner problem in STEM PBL on student critical and creative thinking skill application?	9
Research Question 4: How are student motivation and engagement measured in primary school classrooms during learning activities?	5, 9
Research Question 4a: What is the impact of PBL on student motivation and engagement?	9
Research Question 4b: What is the impact of using an industry partner problem in STEM PBL on student motivation and engagement?	9

PBL in primary school, overlapping with Critical and Creative Thinking and assessment. The insight gained from the conceptual research through Research Questions 2 and 3 also brought to light the potential significance of motivation and engagement on the effectiveness of the industry partner I-9 and PBL, enabling the development of Research Question 4. This question is addressed initially through conceptual research and is positioned in the intersection of motivation and engagement, and assessment methods. The insights gained from the conceptual research enabled the development of refined questions to be investigated through empirical research. The research questions investigated by the research described in this thesis are shown in Table 1.

1.3. Methodology

From the four initial research questions, ('What is the theoretical basis for the Australian Curriculum Critical and Creative Thinking learning continuum?', 'How are critical and creative thinking skills measured during learning?' 'How do the characteristics of STEM PBL affect the use of skills during learning?' and 'How are student motivation and engagement measured in primary school classrooms during learning activities?'), it was evident that a more extensive phase of conceptual analysis would be required before any empirical work could be undertaken. The results of the conceptual research would be required to provide the level of underlying knowledge on which the empirical research could then be structured. Conceptual research enables existing theories to be bridged, work across disciplines to be linked, and the scope of thinking about topics to be broadened (see Gilson & Goldberg, 2015 for a general discussion on conceptual research). Part A of this thesis, 'Making Sense of Problem-Based Learning and Skill Development', describes the conceptual research which was undertaken. The foci were thoroughly investigated, after which refined research questions and testable hypotheses were developed, enabling the empirical studies to be formulated. These empirical studies are described in Part B of this thesis, 'Investigating Problem-Based Learning and Skill Development in the Classroom'.

1.4. Structure of the Thesis

This research thesis is presented in two main parts with the addition of an introductory chapter (the current chapter) and a concluding chapter (Chapter 11). Part A, 'Making Sense of Problem-Based Learning and Skill Development', covers conceptual research over five chapters. Part B, 'Investigating Problem-Based Learning and Skill Development in the Classroom', includes three empirical studies and the description of the combined research findings. Chapter 11 concludes the thesis with a synthesis of the findings, and recommendations for future research.

Part A begins with Chapter 2 in which the Critical and Creative Thinking learning continuum from the Australian Curriculum is introduced and critically interrogated. As the research is situated in South Australia, where the Australian Curriculum (Version 8.4)[2] provides the structure for teaching and assessing school students, this background provides important conceptual understanding. A contribution to knowledge is made through theory adaptation, providing deeper insight into the influence of scholars credited with the development of critical and creative thinking in the Australian Curriculum. This insight enables a clearer understanding and critique of the structure of the continuum and enabled suggested considerations to be shared with the developers of Version 9.0 of the Critical and Creative Thinking continuum. Research Question 1 is addressed, and recommendations are made for further research.

In Chapter 3 the literature on critical thinking and creative thinking is examined. Research on critical thinking is analysed and models of critical thinking are compared for their relevance to the Australian Curriculum. The term 'creative thinking' is clarified and defined in the light of theories of creativity, psychology and education. The relationships between critical and creative thinking, convergent and divergent thinking are analysed, and a novel model is proposed, providing an original contribution to knowledge. Research literature relating to the measurement of critical and creative thinking is examined for relevance in assessing the critical and creative thinking skills in the Australian Curriculum. A list of criteria is developed to enable suitable measures to be identified for use in the classroom. Recommendations are made about where empirical research can make meaningful and novel contributions in skills assessments. As Research Question 2 is not successfully answered a revised question, Research Question 2a, is formulated.

In Chapter 4 Problem-Based Learning and related inquiry styles of teaching are identified and compared through the development of a new comparison chart. Research literature is analysed on the 'problem' in Problem-Based Learning, the evidence of the effect on student skills and the use of industry partners providing a STEM problem. Research Question 3 is investigated, and Revised Research Questions 3a and 3b and Research Question 4 are developed as a result.

The potential roles of motivation and engagement relating to Problem-Based Learning and the use of industry partners are investigated in Chapter 5 in response to the conceptual research in Chapter 4. The assessment of motivation and engagement, their reported connections with industry partners, and their roles in Problem-Based Learning are analysed, leading to the choice of a suitable form of measurement of motivation and engagement for Australian primary school students.

[2] Although Version 9.0 of the Australian Curriculum has been released and is currently being introduced into South Australian classrooms, at the time of these studies Version 8.4 was in use in the classroom.

Research Question 4 is answered, and as a result of the conceptual investigation, revised Research Questions 4a and 4b are developed.

Chapter 6 concludes Part A, drawing together the results of the conceptual research and presenting the revised Research Questions and their associated hypotheses.

Having investigated the factors which are unique and specific to this research through conceptual studies, the following chapters proceed with applying empirical research to answer the remaining research questions. Part B consists of four chapters describing the empirical research, in which a novel assessment task is developed, enabling the remaining research questions to be investigated.

Part B commences with Chapter 7, in which Study One is described. This study identifies *what* should be measured when examining the application of the critical and creative thinking skills in an Australian primary school classroom, followed by testing two measures in the classroom through a pilot study. Through this original theoretical contribution further recommendations are provided about the descriptions of the skills in the Australian Curriculum Critical and Creative Thinking learning continuum, and about the desirable features of an assessment task to measure the application of the skills. As a result of this study, Research Question 2a is answered and Research Question 2b is developed.

In Chapter 8, Study Two focusses on the application of multimedia learning theory to the method of communicating a novel critical and creative thinking self-report measure. Two presentation formats of the measure are compared, and the subsequent effect on students' results is addressed. This innovative application of multimedia learning theory adds a meaningful contribution to understanding the role of multimedia presentations in assessments. The level of success of the two formats is gauged by comparing the teachers' assessment of each student, with each student's own self-report of their use of skills during everyday learning. The results of this analysis bring to light the lack of agreement between the teachers and students, revealing a variation in the percentage agreement rate between teachers and students of 3% to 47%, with a surprisingly low mode of agreement rate of 10%. These results point to a worrying absence of shared understanding about the skills and how to recognise their use by students. This contribution to knowledge has implications for how teachers teach and assess the critical and creative thinking skills, and how effectively teachers are able to report on student critical and creative thinking skills. Implications for teacher education courses are discussed. Research Question 2b is answered.

In Chapter 9, Study Three is described. The novel assessment task developed in the earlier empirical studies is applied in primary school classrooms pre-, post-, and in follow-up, to a STEM Problem-Based Learning intervention. The effects of the use in PBL learning of an industry partner

and their complex, real-world STEM problem, are compared with a control group, and with students undergoing the PBL intervention with a theoretical problem. This original study provides a number of significant contributions to knowledge. Both application of critical and creative thinking skills, and motivation and engagement are assessed, with some unexpected results. The positive effect of an industry partner on the critical and creative thinking skills is clearly demonstrated, with PBL also demonstrating a measurably positive effect. The relationship between student motivation and engagement with PBL and industry partners is less clear, though the results support the claims of PBL proponents. This points to the need for further research to unravel the complex relationship indicated by this research. Research Questions 3a, 3b, 4a and 4b are all addressed in this study.

The final chapter in Part B, Chapter 10, provides a discussion of the findings from the three studies, and the ways in which the research described helps fill the gaps identified in the research literature.

In Chapter 11, the concluding chapter, the unique contributions of all the studies are described, including theoretical advances, the impact of the research on current teaching and teacher education. Practical implications and future research avenues are suggested.

Finally, the Appendices present further detail on the contents of the chapters, including an unpublished report presented to the Department for Education in South Australia, the complete scripts and assessment tasks presented to students, comprehensive lists of creative thinking and critical thinking measures, and the full results of the data analysis on which the conclusions in the research chapters are based.

1.5. Contributions of the Thesis

This thesis contributes knowledge to three sectors of the community: teachers and the schooling sector, particularly at a primary or elementary level of schooling; researchers and the university sector; and curriculum writers and developers. The research findings and potential impact on each sector are presented below.

Curriculum Writers and Developers

As the value of encouraging the development of critical and creative thinking in students is being recognised worldwide, writers of school curricula are including them in their documents (Heilmann & Korte, 2010; Patston et al., 2021) . My findings in respect of the gaps and inconsistencies in the continuum in the Australian Curriculum will be useful to curriculum developers who are identifying and describing the skills to be taught and assessed in schools. Specifically, I recommend that curriculum writers resist the temptation to detail the developmental stages of any specific skills until sufficient empirical research has been undertaken to understand the use and development of critical

and creative thinking skills more fully. I also recommend that consideration is given as to whether their presentation in a continuum is suitable, particularly given the findings regarding the way some skills are being used by students. An alternative way of presenting these skills may more accurately reflect how they develop and are applied during learning.

Teachers and the Primary School Sector

One of the first findings in the research carried out in this thesis was that there is very little research reported on primary school students in the intersection of the fields of PBL, critical and creative thinking skills use, and motivation and engagement. Given the findings in this thesis, this is significant, as it points to a lack of information about this age group on important components of their learning. Carrying out research in primary schools is complicated. Australian classroom teachers are unused to researchers in their classrooms, and commonly use the excuse of 'a crowded curriculum' as the reason they don't have time to allow a researcher access to their students. One alternative would be to encourage teachers to become teacher-researchers. Teachers are often exhorted to use 'research-based evidence' in their teaching practice, and yet they have little access to the results of research other than through professional development – which I have observed through personal experience, may or may not be evidence-based. By encouraging teachers to engage in classroom-based research this will: provide greater access to data on primary school students; help develop more diverse skills in teachers; potentially provide teachers with increased feelings of self-worth which, according to current media reports, is currently at a low ebb; and produce evidence-based classroom-focussed research. To develop teacher-researchers would require input from both the school sector and the university sector. The school sector would need to recognise teacher involvement by developing a pay structure which renumerated teachers for their commitment. Collaborations between teachers and university researchers would facilitate the production of academic articles which disseminate the research to an audience beyond the limited circles of the teacher alone. Potential input from the university sector is detailed below.

The finding that incorporating STEM PBL as a pedagogical approach to teaching and learning in the primary school classroom enables students to use more critical and creative thinking skills, at greater levels of complexity than more passive styles of teaching, should equip schools with the confidence to expand their use of PBL. Furthermore, STEM PBL should be introduced at all levels of primary (and potentially secondary) school, as it has been proven to be effective in increasing critical and creative thinking among primary school children. It cannot be assumed however that all methods of PBL are equal in their effectiveness, and it is therefore essential that effective teacher professional development is provided which identifies the characteristics of effective PBL.

The use of an industry partner to provide the problem which students investigate is a characteristic which has been shown to make a significant difference to the student use of skills. However, teachers need assistance to access industry partners with suitable problems to share with students, and training to understand how to support the industry partners' work in the classroom. The Industry Engagement division in the Department for Education in South Australia provides a strong starting point from which to build a comprehensive 'library' of suitable industry partners.

Researchers and the University Sector

The third sector of the community to benefit from the research described in this thesis is the research and university sector, and there are four findings of significance of relevance to this sector. The first, as previously described, is that there is relatively little research carried out in primary school classrooms relating to the intersection of the fields of PBL, critical and creative thinking skills use, and motivation and engagement. Two barriers were identified through the research carried out in this thesis: gaining access to primary school classrooms is difficult without connections in the individual school; and the requirement for ethics applications by each educational sector (university, public school, Catholic school, private schools) is cumbersome. If the requirement for ethics applications was simplified to enable access to all sectors through one single application, this would streamline the research process. The first point, regarding the difficulty of gaining access to primary school classrooms has already been discussed in the previous section. Developing the concept of teacher-researchers would potentially enable greater access into classrooms for researchers. In addition to encouraging a relationship between classroom teachers and researchers however, pre-service teacher providers (teacher training facilities) would need to consider including a research component in the pre-service teacher education. Pre-service teachers would benefit from being introduced to research methodology which would build the capacity and interest of pre-service teachers. The Swedish and Finnish models may be of interest here (e.g., Koski et al., 2023; Malinen et al., 2012; Olin & Pörn, 2023). As well as increasing the volume of research on primary school students, research from the perspective of an experienced teacher adds a dimension to the research which I believe is not always present in, for example, research on creativity. The value of practice-based evidence should not be overlooked (Australian Psychological Society, 2022). Encouraging greater collaboration between university researchers and practicing teachers – from the conceptual stage of research and not just when it comes to implementing studies in the classroom – would be a beneficial outcome.

The second finding of significance is that there is much we don't know about how critical and creative thinking skills develop and are used by students in the classroom. My research has

shown that a relatively modest investigation can advance knowledge significantly. Further information including more accurate descriptions of the skills and how they develop and are used, based on empirical research, would aid teachers in identifying the needs of their students.

In addition to further research on the critical and creative thinking skills, building an awareness of the relevance of the skills to student learning is advised. Pre-service teachers would benefit from an explicit focus on the skills and the benefits of providing opportunities for students to use and develop their critical and creative thinking skills. My limited observations of pre-service teachers graduating from universities in South Australia[3] is that they have at best a cursory familiarity with the critical and creative thinking skills in the Australian Curriculum, and most graduates (and many teachers) have little or no experience with the Critical and Creative Thinking learning continuum. The role of PBL in enabling the use of critical and creative thinking during learning is certainly an area which should be emphasised during pre-service training.

The third significant finding of relevance to the university and research sector is that the relationship between motivation and engagement, PBL and industry partners is complex. Further research is required to understand these relationships and the role they may play in the successful development of critical and creative thinking skills in students – particularly, but not exclusively, of primary school students. The effects of different forms of 'problem' in Problem-Based Learning (i.e., hypothetical problems versus genuine industry partner problems), on the use of critical and creative thinking skills, requires investigation to enable a more thorough understanding. Additionally, identifying whether the apparent benefits gained by using an industry partner are due to the person (the 'stranger in the room'), or if they are due to the nature of the problem provided by the industry partner, requires further investigation.

Finally, of interest to researchers is that the use of critical and creative thinking skills can be measured using the self-report measure developed in this thesis. The self-report measure can be used to identify the use of critical and creative thinking skills during learning, rather than on the students' potential to use skills, or their ability to use skills in artificial situations.

[3] This experience is gained from tutoring pre-service teachers at two institutions and supervising pre-service practicums over approximately 5 years.

Part A.
Making Sense of Skill Development and Problem-Based Learning

In Chapter One the impetus and motivation for the research described in this thesis are explained, and an overarching question and four research questions are identified. The questions cover four broad areas: the Australian Curriculum Critical and Creative Thinking learning continuum (the C&CrT continuum); critical and creative thinking (C&CrT) and its assessment; the effect of STEM PBL and industry partners on students' use of skills; and the potential role of motivation and engagement. The initial investigation of each of these areas reveals a lack of consensus in some areas, and a lack of relevant research in most of the areas of interest. It was important therefore to understand and clarify the current state of knowledge in each of these four areas prior to the application of empirical research on students in the classroom. Conceptual research which exposes the theoretical frameworks and engages in conceptual clarification and synthesis within these four areas is required. Part A of this thesis therefore consists of four conceptual research chapters.

The conceptual research reported here is approached through an analytical framework. It draws on literature reviews, theoretical analyses, and the synthesis of existing frameworks. Current theories from psychology and education are evaluated to understand the structure of the existing Critical and Creative Thinking learning continuum more deeply (theory adaptation), to clarify definitions of critical and creative thinking for relevance in education (theory synthesis), and to provide a refined description of the relationship between Problem-Based learning and other forms of inquiry-based approaches (Typology: Jaakkola, 2020). By critically evaluating the theoretical landscape, the study identifies gaps and potential areas for further theoretical development, aiming not only to describe the existing theories but also to propose conceptual models that can guide future research relating to the areas under examination. The application of these conceptual models enables the empirical studies to be formulated.

A.1. Methodology of Conceptual Studies

The first of the conceptual studies investigates the accuracy of the descriptions of the skills in the Australian Curriculum Critical and Creative Thinking learning continuum through Research Question 1, 'What is the theoretical basis for the Australian Curriculum Critical and Creative Thinking learning continuum?' This comprehensive conceptual research study sought to understand the conceptual

frameworks that underlie the delineation of the Critical and Creative Thinking skills and developmental progressions within the Critical and Creative Thinking learning continuum, providing a nuanced understanding of the theoretical landscape. The study aimed to go beyond surface-level descriptions by engaging in a theoretical exploration to uncover the foundational principles that guide the integration of these skills into the educational framework.

The second conceptual study investigates the theoretical foundations of critical and creative thinking and the most suitable measure to assess the use of critical and creative thinking skills by students during learning activities by answering Research Question 2, 'How are critical and creative thinking skills measured during learning?'. The study involved an extensive literature review scrutinizing existing assessment tools to evaluate their applicability. This led to the development of a further research question (RQ2a) which required further investigation. This conceptual research phase played a pivotal role in shaping the subsequent empirical design – addressed in Part B – as it provided a nuanced understanding of the complexities associated with assessing critical and creative thinking in the dynamic context of learning activities.

The third conceptual study investigates the extent to which using a PBL approach affects student skill use through a comprehensive exploration of the theoretical underpinnings of Problem-Based Learning (PBL) and its potential impact on student skill development. Further, it includes an in-depth review of educational literature, examining the core principles of PBL, and the mechanisms through which PBL fosters skill acquisition and application, considering both cognitive and affective dimensions. This study addresses Research Question 3, 'How do the characteristics of STEM PBL affect the use of skills during learning?', with the result that two additional research questions are developed – RQ 3a and RQ 3b. These questions are addressed in Part B through empirical research. Research Question 4, 'How are student motivation and engagement measured in primary school classrooms during learning activities?', is also developed as a result of this study, and addressed in the final study in Part A. Two additional research questions connecting RQ 4 to PBL, (RQ 4a 'What is the impact of PBL on student motivation and engagement?' and RQ4b 'What is the impact of using an industry partner problem in STEM PBL on student motivation and engagement?') are developed as a result and are also addressed in Part B.

The fourth and final conceptual study in this section investigates how motivation and engagement can be effectively measured in primary school classrooms during learning activities with students: undertaking STEM learning; having a range of literacy levels; and with limited access to technology. This involved a thorough exploration of diverse literature and theoretical frameworks related to the measurement of motivation and engagement in primary school settings. This

conceptual research aimed to identify suitable indicators and dimensions for measuring motivation and engagement that are sensitive to the diverse learning needs of Australian primary school students engaged in STEM activities. Building on this conceptual groundwork, existing measurement tools for assessing motivation and engagement were critically examined, enabling a suitable measure to be identified. This study concluded the conceptual research component and by application of the conceptual models, this enabled the empirical studies described in Part B to be formulated.

The final chapter in Part A presents the findings and conclusions from the conceptual studies.

Chapter 2.

Investigating Critical and Creative Thinking in the Australian Curriculum

In order to understand and assess critical and creative thinking skills, it is necessary to understand their place in the Australian Curriculum, the reasons for their inclusion and the intention of curriculum designers in respect of how the skills would be taught and learned in schools.

2.1. Background to the National Curriculum

In 2008 The Melbourne Declaration (Ministerial Council on Education, 2008) noted the changing face of the workplace and the need for graduates to be, among other things, creative problem solvers. A national curriculum was promised that would "support young people to develop a range of generic and employability skills that have particular application to the world of work"(p. 13) and that would enable young people to "develop the capacity to think creatively, innovate, solve problems" (p. 13).

The mandated curriculum in South Australia's public schools at this time was SACSA – the South Australian Curriculum, Standards and Accountability framework. Student skills were not a strong focus of SACSA, with more emphasis being placed on learning areas or subject knowledge (Cockburn, 2005; Yates & Collins, 2010). The Melbourne Declaration (Ministerial Council on Education, 2008) provided the impetus and direction for the development of the new Australian Curriculum which was gradually implemented across Australia (Donnelly & Wiltshire, 2014; for a more comprehensive description of the factors which influenced the development of the curriculum as a whole see e.g., Reid, 2018; Scarino, 2019). In South Australia implementation in primary school classrooms commenced in 2012 with Maths and Science, additional subjects being added each year, culminating with Languages in 2017 (Department for Education and Child Development (DECD), 2013).

Concurrent with the development of the new curriculum were concerns over the falling numbers of STEM graduates. In 2014 the Office of the Australian Chief Scientist (2014, p. 11) reported the lowest level of participation in science subjects in schools for 20 years. Downward

trends in the numbers of students participating in STEM subjects[4] had also been flagged in other Asian and Western countries (Gamse et al., 2017; Kelley & Knowles, 2016). This decline in the rates of participation in Australia was reflected in the relative performance of school students in science and mathematics in PISA[5] from 2006 to 2012 (Office of the Chief Scientist, 2014; Thomson et al., 2017), and again in 2018 (Thomson, 2021). The slump in numbers of Australia's STEM graduates[6] threatened Australia's current economic prosperity and future economic success (Education Council, 2015; Office of the Chief Scientist, 2013).

Adding to concerns around the decline in numbers of STEM graduates was the anticipated changes in the workplace brought about by Industry 4.0 and the resultant impact on the skills required by students entering the workforce (Torii, 2018; World Economic Forum, 2016, 2018, 2020). The Fourth Industrial Revolution[7] , we were told, would "reduce the number of workers required for certain work tasks. …[while] increased demand for new roles…[would result in] difficult transitions for millions of workers and the need for…a new surge of agile learners and skilled talent" (World Economic Forum, 2018, p. v). Australia's workforce would require increased abilities in what are labelled 21st Century, STEM, or soft skills (Ananiadou & Claro, 2009; Lucas & Smith, 2018; Vincent-Lancrin et al., 2019). These skills include critical thinking and analysis, creativity and problem solving (Education Council, 2015; World Economic Forum, 2016). The combination of low rates of students in STEM subjects, falling success rates of Australian students in international tests, and a growing recognition of the importance of skills was an ideal environment in which to introduce a new national curriculum. Covering teaching from Foundation to Year 10, the Australian Curriculum was developed with an increased focus not just on what students should know, but also on the skills they should be able to use (Gilbert, 2019; Gonski, 2018; Yates & Collins, 2010). As described by the developers of the Australian Curriculum — the Australian Curriculum Assessment and Reporting Authority (ACARA):

> …21st century learning does not fit neatly into a curriculum solely organised by learning areas or subjects that reflect the disciplines. Increasingly, in a world where

[4] 'STEM subjects' refers to any specialist subjects which fall under the categories of Science, Technologies, Engineering or Mathematics. Although the terms STEMM (including Medicine) and STEAM (including the arts) are frequently used, these references were focussed specifically on the core STEM subjects.
[5] PISA is the Programme for International Student Assessment administered by the OECD to 15-year-old students. The testing is carried out every three years. In 2022 (the 2021 test was delayed due to COVID-19) 85 countries world-wide took part. The main focus of the survey varies, e.g., in 2018 the focus was on Reading, in 2015 Science and Collaborative Problem Solving, in 2012 Mathematics. https://www.oecd.org/pisa/
[6] 'STEM graduates' refers to students with qualifications which fall under the categories of Science, Technologies, Engineering or Mathematics.
[7] The terminology of Industry 4.0 and the Fourth Industrial Revolution were specified in these references, although the Digital Age and now, Industry 5.0 are also terms used in current literature to describe the growth in use and complexity of technologies which are becoming embedded in our daily lives.

knowledge itself is constantly growing and evolving, students need to develop a set of knowledge, skills, behaviours and dispositions, or general capabilities that apply across subject-based content and equip them to be lifelong learners able to operate with confidence in a complex, information-rich, globalised world. (Australian Curriculum Assessment and Reporting Authority (ACARA), 2012, p. 15).

2.2. The Structure of the Australian Curriculum

The curriculum was designed in what are described as three distinct dimensions — Learning Areas (subject-knowledge curriculum); General Capabilities; and Cross-curriculum Priorities (Australian Curriculum Assessment and Reporting Authority (ACARA), 2017). The structure of the curriculum, illustrated in Figure 2, presents the three dimensions as separate components which are designed to be woven together through the use of effective teaching strategies (Australian Curriculum Assessment and Reporting Authority (ACARA), 2013b). The 21st century skills, located in the



Figure 2

The structure of the Australian Curriculum[8]

[8] *Note*. Australian Curriculum Assessment and Reporting Authority (ACARA). (n.d.-c). The Three Dimensions of the Australian Curriculum. https://www.australiancurriculum.edu.au/media/1423/the-three-dimensions-of-the-australian-curriculum.png?width=500&height=343.9575033200531 In the public domain.

dimension of the General Capabilities, are described as “an integrated and interconnected set of knowledge, skills, behaviours and dispositions” (Australian Curriculum Assessment and Reporting Authority (ACARA), 2017, 'The whole curriculum' section). The term Critical and Creative Thinking was coined to replace the term 21st Century Skills, as the skills in question are not unique to the 21st Century, indeed their use has been recognised for generations (Cropley, 2019b, 2020; McGaw, 2013). Critical thinking and creative thinking, although initially treated separately in the early stages of development of the curriculum, became difficult to distinguish as individual entities, resulting in their amalgamation in the final iteration of the Critical and Creative Thinking learning continuum (McGaw, 2013, p. 8).

During the period in which the research described in this thesis was being carried out, version 8.4 of the Australian Curriculum was in use in schools. In 2020 a review of the Australian Curriculum commenced, and version 9.0 was endorsed in mid-2022 (Australian Government Department of Education, 2022). Version 8.4 remains in use in schools currently, with version 9.0 being implemented from 2023 “according to [schools] own timelines.” (Australian Government Department of Education, 2022, Review of the Australian Curriculum section). The results of the revision of the Critical and Creative Thinking skills are discussed below.

Critical and Creative Thinking in version 8.4 of the Australian Curriculum incorporates four categories of skills. Described as ‘elements’, these consist of the categories of: Inquiring - identifying, exploring and organising information and ideas element; Generating ideas, possibilities and actions element; Reflecting on thinking and processes element; and Analysing, synthesising and evaluating reasoning and procedures element. These four elements are then broken down further, each into three sub-elements (i.e., 12 sub-elements in total), providing the opportunity to supply further clarity to the skills included within the curriculum. The terminology of the continuum — elements, sub-elements and levels — is demonstrated in Figure 3. As shown in Figure 4, each sub-element is further articulated by a series of 6 statements, describing the stages in the development of each sub-element from Foundation to Year 10 (Australian Curriculum Assessment and Reporting Authority (ACARA), 2013b). These 72 statements, the ‘levels’, form the Critical and Creative Thinking learning continuum (C&CrT continuum).

The background literature on the Australian Curriculum describes the C&CrT continuum as a progressive continuum where “What is learned in the early years supports all subsequent learning.” (Australian Curriculum Assessment and Reporting Authority (ACARA), 2013b, p. 7). The continuum had been originally designed for use by the curriculum writers, and then re-purposed for teacher use

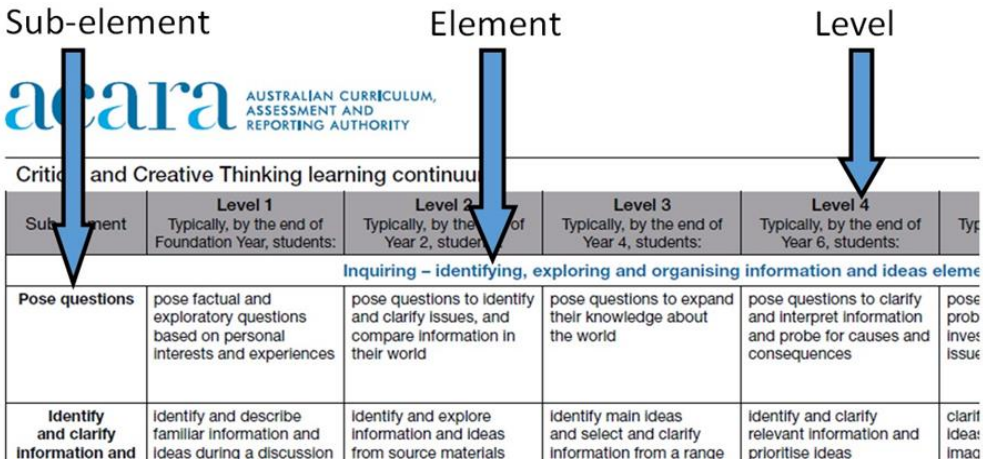


Figure 3

Terminology of the structure of the Critical and Creative Thinking learning continuum (V8.4)

Figure 4 (over)

The Australian Curriculum, Critical and Creative Thinking learning continuum[9]

[9] Australian Curriculum Assessment and Reporting Authority (ACARA). (n.d.-b). General Capabilities Creative and Critical Thinking learning continuum. In Australian Curriculum General Capabilities Version 8.4: ACARA. In the public domain.

Critical and Creative Thinking learning continuum

Sub-element	Level 1 Typically, by the end of Foundation Year, students:	Level 2 Typically, by the end of Year 2, students:	Level 3 Typically, by the end of Year 4, students:	Level 4 Typically, by the end of Year 6, students:	Level 5 Typically, by the end of Year 8, students:	Level 6 Typically, by the end of Year 10, students:
Inquiring – identifying, exploring and organising information and ideas element						
Pose questions	pose factual and exploratory questions based on personal interests and experiences	pose questions to identify and clarify issues, and compare information in their world	pose questions to expand their knowledge about the world	pose questions to clarify and interpret information and probe for causes and consequences	pose questions to probe assumptions and investigate complex issues	pose questions to critically analyse complex issues and abstract ideas
Identify and clarify information and ideas	identify and describe familiar information and ideas during a discussion or investigation	identify and explore information and ideas from source materials	identify main ideas and select and clarify information from a range of sources	identify and clarify relevant information and prioritise ideas	clarify information and ideas from texts or images when exploring challenging issues	clarify complex information and ideas drawn from a range of sources
Organise and process information	gather similar information or depictions from given sources	organise information based on similar or relevant ideas from several sources	collect, compare and categorise facts and opinions found in a widening range of sources	analyse, condense and combine relevant information from multiple sources	critically analyse information and evidence according to criteria such as validity and relevance	critically analyse independently sourced information to determine bias and reliability
Generating ideas, possibilities and actions element						
Imagine possibilities and connect ideas	use imagination to view or create things in new ways and connect two things that seem different	build on what they know to create ideas and possibilities in ways that are new to them	expand on known ideas to create new and imaginative combinations	combine ideas in a variety of ways and from a range of sources to create new possibilities	draw parallels between known and new ideas to create new ways of achieving goals	create and connect complex ideas using imagery, analogies and symbolism
Consider alternatives	suggest alternative and creative ways to approach a given situation or task	identify and compare creative ideas to think broadly about a given situation or problem	explore situations using creative thinking strategies to propose a range of alternatives	identify situations where current approaches do not work, challenge existing ideas and generate alternative solutions	generate alternatives and innovative solutions, and adapt ideas, including when information is limited or conflicting	speculate on creative options to modify ideas when circumstances change
Seek solutions and put ideas into action	predict what might happen in a given situation and when putting ideas into action	investigate options and predict possible outcomes when putting ideas into action	experiment with a range of options when seeking solutions and putting ideas into action	assess and test options to identify the most effective solution and to put ideas into action	predict possibilities, and identify and test consequences when seeking solutions and putting ideas into action	assess risks and explain contingencies, taking account of a range of perspectives, when seeking solutions and putting complex ideas into action

Critical and Creative Thinking learning continuum

Sub-element	Level 1 Typically, by the end of Foundation Year, students:	Level 2 Typically, by the end of Year 2, students:	Level 3 Typically, by the end of Year 4, students:	Level 4 Typically, by the end of Year 6, students:	Level 5 Typically, by the end of Year 8, students:	Level 6 Typically, by the end of Year 10, students:
Reflecting on thinking and processes element						
Think about thinking (metacognition)	describe what they are thinking and give reasons why	describe the thinking strategies used in given situations and tasks	reflect on, explain and check the processes used to come to conclusions	reflect on assumptions made, consider reasonable criticism and adjust their thinking if necessary	assess assumptions in their thinking and invite alternative opinions	give reasons to support their thinking, and address opposing viewpoints and possible weaknesses in their own positions
Reflect on processes	identify the main elements of the steps in a thinking process	outline the details and sequence in a whole task and separate it into workable parts	identify pertinent information in an investigation and separate into smaller parts or ideas	identify and justify the thinking behind choices they have made	evaluate and justify the reasons behind choosing a particular problem-solving strategy	balance rational and irrational components of a complex or ambiguous problem to evaluate evidence
Transfer knowledge into new contexts	connect information from one setting to another	use information from a previous experience to inform a new idea	transfer and apply information in one setting to enrich another	apply knowledge gained from one context to another unrelated context and identify new meaning	justify reasons for decisions when transferring information to similar and different contexts	identify, plan and justify transference of knowledge to new contexts
Analysing, synthesising and evaluating reasoning and procedures element						
Apply logic and reasoning	identify the thinking used to solve problems in given situations	identify reasoning used in choices or actions in specific situations	identify and apply appropriate reasoning and thinking strategies for particular outcomes	assess whether there is adequate reasoning and evidence to justify a claim, conclusion or outcome	identify gaps in reasoning and missing elements in information	analyse reasoning used in finding and applying solutions, and in choice of resources
Draw conclusions and design a course of action	share their thinking about possible courses of action	identify alternative courses of action or possible conclusions when presented with new information	draw on prior knowledge and use evidence when choosing a course of action or drawing a conclusion	scrutinise ideas or concepts, test conclusions and modify actions when designing a course of action	differentiate the components of a designed course of action and tolerate ambiguities when drawing conclusions	use logical and abstract thinking to analyse and synthesise complex information to inform a course of action
Evaluate procedures and outcomes	check whether they are satisfied with the outcome of tasks or actions	evaluate whether they have accomplished what they set out to achieve	explain and justify ideas and outcomes	evaluate the effectiveness of ideas, products, performances, methods and courses of action against given criteria	explain intentions and justify ideas, methods and courses of action, and account for expected and unexpected outcomes against criteria they have identified	evaluate the effectiveness of ideas, products and performances and implement courses of action to achieve desired outcomes against criteria they have identified

(Australian Curriculum Assessment and Reporting Authority (ACARA), 2013a), and this was reflected in the complex language describing the levels.

2.3. The Theories Behind the Critical and Creative Thinking Skills

The language used in the continuum, and indeed the ideas behind structure of the continuum have been identified in the public realm in general terms, through an acknowledgment of the influence of relevant general educational theories and their creators as shown in Table 2. In addition, the help of expert advisors (education and curriculum), and the contributions of national and international research (not specified) are acknowledged (Australian Curriculum Assessment and Reporting Authority (ACARA), 2013b, p. 69).

Although the contributions of some scholars are identified more specifically in the literature describing the development of the continuum, the greater part of their influence on the structure of this influential document is left to speculation. Gardner (2009) for example, is credited with recommending the fostering of skills through intellectual flexibility and open-mindedness, and Perkins (1994) is credited with promoting the development of critical and creative thinking through broad and adventurous thinking and metacognition. Theorists whose contribution is less clearly defined include Marzano, Richhart and Bloom.

Table 2

The origins of the skills in the continuum

Categorisation of skill	Theorists
dispositions	Perkins (1994); Ritchhart et al. (2011); Tishman et al. (1995)
philosophical inquiry	Sharp et al. (1980)
taxonomies of skills	Anderson and Krathwohl (2001); Bloom et al. (1956)
thinking strategies	Marzano et al. (2001)
habits and frames of mind	Costa and Kallick (2004); de Bono (2009); Gardner (1983)

Marzano's thinking on the subject specific/subject general nature of skills (Marzano, 2010), although not specified, may well have influenced the Australian Curriculum's position on this matter[10]. The description of critical thinking provided in the Australian Curriculum ("…at the core of most intellectual activity that involves students learning to recognise or develop an argument, use evidence in support of that argument, draw reasoned conclusions, and use information to solve problems"; Australian Curriculum Assessment and Reporting Authority (ACARA), n.d.-a) clearly reflects the six general reasoning and thinking skills identified by Marzano through the McREL study (Marzano, 2010, p. 270).

The relationship to Ritchhart, Church and Morrison's work on 'Making Thinking Visible'(2011) is seen in a strong resemblance to the phrasing in the sub-elements of the continuum. This is demonstrated in Table 3, where the sub-elements of the C&CrT continuum are compared with Ritchhart et al.'s work. Although somewhat casually presented by Richart et al. (2011) as "… types of thinking … that seem useful …"(p. 14), 12 of the 15 statements which appear in the lists of "8 thinking moves that lead to understanding" (pp. 11 & 13), "other kinds of thinking" (p. 14) and "students' thinking about thinking" (p. 15), show a clear link to the choice of sub-elements which have been incorporated into the continuum.

The third set of theorists whose influence can be recognised in the continuum, are those who describe ways of comparing increased levels of complexity within each skill. Their influence is seen in the statements relating to the levels, six of which are provided for each sub-element, accounting for the development of the skill from approximately 5 years to 16 years of age.

Table 3 (over)

Comparison of terminology for Australian Curriculum V8.4 C&CrT continuum sub-elements with Ritchhart et al.'s 'Making Thinking Visible'

[10] The skills are described as being 'general capabilities' but to be taught within subject specializations. Care, E., Kim, H., Vista, A., & Anderson, K. (2018). Education System Alignment for 21st Century Skills: Focus on Assessment. *Center for Universal Education at The Brookings Institution*. , Gilbert, R. (2018). General capabilities in the Australian Curriculum: Promise, problems and prospects. In A. Reid & D. Price (Eds.), *The Australian Curriculum: Promises, problems and possibilities*. Australian Curriculum Studies Association.

Australian Curriculum[1]		Richhart[2]
Element	Sub-element	'kinds of thinking'
Inquiring	Pose questions	Wondering and asking questions
	Identify and clarify information and ideas	Identify patterns and make generalizations
	Organise and processinformation	No equivalent
Analysing	Apply logic and reasoning,	Identifying claims, assumptions and bias, Building explanations and interpretations, Reasoning with evidence
	Draw conclusions and design a course of action	Formulating plans and monitoring actions, Capturing the heart and drawing conclusions
	Evaluate procedures and outcomes	Evaluating evidence, arguments and actions
Generating	Imagine possibilities and connect ideas	Making connections
	Consider alternatives	Considering different viewpoints and perspectives
	Seek solutions and put ideas into action	No equivalent
Reflecting	Think about thinking (metacognition)	Students thinking about thinking
	Reflect on processes	No equivalent
	Transfer knowledge into new contexts	No equivalent
No equivalent	No equivalent	Observing closely and describing what's there
	No equivalent	Uncovering complexity and going below the surface of things

Note.[1] Australian Curriculum Assessment and Reporting Authority (ACARA) (n.d.-a V 8.4). [2]Ritchhart et al. (2011, pp. 11-15)

2.4. Theoretical Basis to the 'Levels' in the Continuum

Bloom and colleagues (e.g., Bloom et al., 1956) are one of the groups of scholars who recognise the value of describing levels of knowing, i.e., not just *what* you know, but also *how well* you know it. The horizontal axis (the levels) of the C&CrT continuum shows some evidence of the influence of their work, however the theory has not been applied consistently to the continuum. The cognitive process dimension described by Krathwohl (2002, p. 215) in his revision of Bloom's Taxonomy appears to have provided inspiration for some of the descriptions of students' learning at increasing stages of complexity, demonstrated by the descriptions of the skills (i.e., the 'level statements') which at first glance appear to progress in incremental steps from Foundation to Year 10. The cognitive process dimension is described by Krathwohl as containing 6 stages: Remember, Understand, Apply, Analyse, Evaluate and Create (p. 215). As an example of how these stages have been overlaid onto the continuum, the skill of 'Identify and clarify information and ideas' can be examined, and this is illustrated in Figure 5.

At an entry level (level 1, typically achieved by the end of Foundation year) the skill is described as requiring the student to be able to "identify and describe familiar information and ideas etc." (Australian Curriculum Assessment and Reporting Authority (ACARA), n.d.-b). This type of activity requires only the first level of cognitive process — 'remembering'. By level 3 (typically by the end of year 4), the skill is described as "Identify main ideas and select and clarify information from a range of sources" (Australian Curriculum Assessment and Reporting Authority (ACARA), n.d.-b). This type of activity would require the application of the three more complex stages of: understanding; application of a skill; and analysis of the information. The expectation that each consecutive level requires greater skill mastery, demonstrated by the application of more complex cognitive processes to each of the skills, is reinforced by statements from both Krathwohl and ACARA: Krathwohl, with his acknowledgement of the hierarchical nature of the Taxonomy, i.e., that "mastery of each simpler category was prerequisite to mastery of the next more complex one", (Krathwohl, 2002, p. 212); and ACARA's description of the skills being presented in a developmental sequence[11], "where learning at the higher levels was dependent on having obtained the prerequisite knowledge and skills at lower levels" (Australian Curriculum Assessment and Reporting Authority (ACARA), 2013b, p. 69).

[11] This is supported by more recent documentation, describing changes to the General Capabilities as having 'Improved the developmental sequence across the levels of each continuum' Australian Curriculum Assessment and Reporting Authority (ACARA). (2023b). *What's changed in the new Australian Curriculum?* ACARA. Retrieved 14 June 2023 from https://v9.australiancurriculum.edu.au/resources/stories/curriculum-changes#accordion-41087ee733-item-594b90cc58. It is not clear from the documentation whether the term 'developmental' is referring to a sequence which is developing or if it is referring to the sequence of stages being predicated on the age-related development of the child.

However, as shown in Figure 5, the sequence is not successfully continued at the higher levels of the continuum (levels 5 and 6) representing descriptions of learning expected to be exhibited by students at years 8 and 10. This lack of consistency in the descriptions of skill development in consecutive levels is also observed in other sub-elements in the C&CrT continuum.

The value of the continuum should not be overlooked, however. Providing a document for teachers that demonstrated a trajectory for the development of critical and creative thinking skills – even if not fully verified – encouraged teachers to observe students and consider how the skills were being taught. As an early conceptual model from which to develop more comprehensive theories and understandings, its value is significant. The Australian Curriculum is reviewed every six years (Australian Curriculum Assessment and Reporting Authority (ACARA), 2023b), the latest update resulting in Version 9.0. During the review process opinions were sought from a range of experts, including creativity and critical thinking researchers, resulting in the sub-elements being reframed and the descriptions being re-worded for clarity and simplicity. These regular reviews are valuable as they enable new knowledge and understandings about the critical and creative thinking skills to be incorporated as discoveries are made.

2.5. The Accuracy of the Learning Continuum

The description of the C&CrT continuum as a learning continuum created the expectation that statements at each consecutive 'level' within a sub-element would describe progressive steps in each skill (Victorian Curriculum and Assessment Authority (VCAA), 2019, p. 4). For the statements to be seen as a continuum, i.e., a progression of elements varying by degrees (Merriam-Webster, n.d.), the description of each skill should increase in complexity, from one level to the next (Biggs, 1999; DeLuca et al., 2019), with each earlier stage required to enable the consecutive stages to occur. An example could be that of a child learning to crawl. A usual progression might be: a) they wriggle on their tummy, b) they pull their knees up underneath them, and c) worm forward, then d) they crawl on hands and knees. The progression would not be: a) they wriggle on their tummy, b) they use furniture to support themselves standing. The statements in this second example may both be observed as stages of development, but they do not show a progression of the development of the skill, or an effective learning continuum. In this second example, step 'a' and step 'b' describe stages of different progressions, with 'a' possibly but not necessarily being part of the progression which ends with step 'b'. Likewise, the statements in the C&CrT learning continuum do not consistently show the progressive development of each of the skills.

acara AUSTRALIAN CURRICULUM, ASSESSMENT AND REPORTING AUTHORITY

Australian CURRICULUM

Critical and Creative Thinking learning continuum

Sub-element	Level 1 Typically, by the end of Foundation Year, students:	Level 2 Typically, by the end of Year 2, students:	Level 3 Typically, by the end of Year 4, students:	Level 4 Typically, by the end of Year 6, students:	Level 5 Typically, by the end of Year 8, students:	Level 6 Typically, by the end of Year 10, students:
Identify and clarify information and ideas	identify and describe familiar information and ideas during a discussion or investigation	identify and explore information and ideas from source materials	identify main ideas and select and clarify information from a range of sources	identify and clarify relevant information and prioritise ideas	clarify information and ideas from texts or images when exploring challenging issues	clarify complex information and ideas drawn from a range of sources
Blooms levels:	remember	understand	understand apply analyse	understand analyse evaluate	understand analyse	understand analyse

Figure 5

Comparison of Blooms taxonomy levels with the C&CrT continuum[12]

[12] Based on Australian Curriculum Assessment and Reporting Authority (ACARA). (n.d.-b). General Capabilities Critical and Creative Thinking learning continuum. In *Australian Curriculum General Capabilities Version 8.4*: ACARA. In the public domain. And Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory Into Practice*, *41*(4), 212-218.

This lack of consistency in the progressive development of the skills is also seen for example in some sub-elements in the continuum where multiple skills are described. Sometimes the descriptions are provided consistently across all levels, at other times a description is missed at a year level or is replaced by a similar but different skill description. This is illustrated in Figure 6, for the sub-element 'Imagine possibilities and connect ideas' where three skills are described: being able to create things; being able to connect or combine things; and being able to compare things. As can also be seen in Figure 6 the skill of 'creating things' is mentioned at every level. The differences in the level of skill are not obvious between: creating things in new ways (Foundation level); creating things in ways that are new (Year 2 level); creating new combinations (Year 4 level); creating new possibilities (Year 6 level); and creating new ways of achieving goals (Year 8 level). Describing the stages of the skill as it develops has not been clearly articulated here. If Foundation level teachers were presented with the question, 'Do you observe Foundation level students creating things using new and imaginative combinations?', they would answer 'yes' even though the expectation is that this skill will not be achievable until around Year 4 according to the continuum. The statement that students 'create new and imaginative combinations' (the level 3 statement) then, does not adequately describe a more advanced level of skill which would demonstrate a greater level of ability in a student. It can also be seen that the very distinct and important skill of comparing has been inserted only at the Year 8 level. This is an example of where a skill is described at one or two levels but is not consistently described across all levels. This omission may have been due to a requirement for brevity from the writers of the continuum, or because earlier and later stages of the skill could not be easily identified.

An example where the concept of progression has been effectively described is in the SOLO taxonomy by Biggs and Collis (1982) and Biggs and Tang (2011). The SOLO (Structure of the Observed Learning Outcome) taxonomy describes linear additive progressions by describing an increasing complexity of levels of understanding labelled unistructural, relational, multistructural and extended abstract. Terms such as 'define' and 'identify' characterise unistructural understanding, while 'describe', 'list' and 'combine' are demonstrations of multistructural understanding. Relational and extended abstract levels are illustrated by terms such as 'compare', 'classify' and 'analyze', with 'evaluate', 'generalise' and 'hypothesise' reflecting the highest level of understanding. Reflecting back on the Blooms terminology applied to the C&CrT continuum in Figure 5 illustrates again the lack of consistency in the descriptions of the skills if they are to be understood as a linear progression. A more detailed description of the application of the SOLO taxonomy to the use of critical and creative thinking skills during learning is provided in Chapter 7.

Generating ideas, possibilities and actions element						
Imagine possibilities and connect ideas	use imagination to view or create things in new ways and connect two things that seem different	build on what they know to create ideas and possibilities in ways that are new to them	expand on known ideas to create new and imaginative combinations	combine ideas in a variety of ways and from a range of sources to create new possibilities	draw parallels between known and new ideas to create new ways of achieving goals	create and connect complex ideas using imagery, analogies and symbolism
Consider alternatives	suggest alternative and creative ways to approach a given	identify and compare creative ideas to think broadly about a given	explore situations using creative thinking strategies to propose a	identify situations where current approaches do not work, challenge	generate alternatives and innovative solutions, and adapt ideas, including	speculate on creative options to modify ideas when circumstances

Imagine possibilities: creating things

Connect ideas: combining things

Compare multiple ideas

Note. Adapted from Australian Curriculum Assessment and Reporting Authority (ACARA) (n.d.-b). In the public domain.

Figure 6

Continuity of skills descriptions in sub-element 'Imagine possibilities and connect ideas'

2.6. Conclusions Regarding the Continuum

My observations of students in the classroom, and discussions with them during the STEM Evaluation Study (Appendix A) confirmed that students did not necessarily use, and certainly hadn't mastered, all the earlier levels of each skill before moving to the next level of skill described. Furthermore, students reported that they were working at a Year 10 level in some skills, and at a Year 2 level in other skills. These observations added fuel to the concerns that the skills as presented in the C&CrT continuum do not accurately represent the way the skills develop, or how they are used by students.

The theory synthesis applied to the literature referenced for the Critical and Creative Thinking skills in the Australian Curriculum has enabled Research Question 1, 'What is the theoretical basis for the Australian Curriculum Critical and Creative Thinking learning continuum?', to be effectively answered. Further, this investigation has shown the inaccuracies in the representations of the development of the skills described within the Critical and Creative Thinking learning continuum. Although there has been an attempt to capture some conceptual nuances, this has not been completely successful. As a result the continuum provides disjointed descriptions of the skills in most cases, and incomplete descriptions of the developmental stages of the skills.

As the Australian Curriculum, and therefore also the C&CrT continuum, is mandated for teaching in Australian schools, the question of accuracy of the descriptions of the skills is significant. This finding has implications for both the teaching and assessment of the skills being used by students. Teaching and assessment are complementary activities in classrooms. As students demonstrate their learning (formative assessment), teachers identify the teaching needed to support the student to reach the next goal. Teachers are required to teach what is in the curriculum and are expecting to see the skills displayed by students, as they are described in the continuum for specific year levels. If a teacher bases their expectations on a continuum which is inaccurate, they are unlikely to recognise opportunities to stretch their students. There is also the risk that teachers only familiarise themselves with the skill descriptions up to the year level at which they are teaching, and do not familiarise themselves with the skills at higher levels. Teachers would therefore find it difficult to accurately identify the levels of skills students are capable of working at i.e., to accurately assess students. As the Critical and Creative Thinking learning continuum has been based on current educational research, this would indicate the need for further research on the skills and both how they develop, and how they are applied by students in genuine learning situations.

The research into the C&CrT continuum in this chapter has identified its origins and demonstrated its connections to leading educational theories about critical thinking. Further, this has led to the identification of weaknesses in the application of the theories to the descriptions of developmental sequences of the critical and creative thinking skills. These weaknesses in the sequencing of the skills have serious implications for their practical application in the classroom, and to how we measure these skills during learning. Rather than a linear continuum, it is possible that at least some of the skills require an alternative way of describing how they are used and develop. Additionally, if the skills are not a progression, with earlier descriptions being a precursor to later stages, this may affect how measures are structured. Following this conceptual research, existing theories of critical thinking will be compared to the structure of the continuum to determine whether suitable measures exist to assess student use of the skills during learning.

Chapter 3.

Theoretical Foundations and Measurement of Critical and Creative Thinking

This chapter looks at the theoretical underpinnings of critical and creative thinking, proposes new definitions and reviews existing assessment tools.

3.1. Introduction

The background and structure of the Australian Curriculum's Critical and Creative Thinking learning continuum (C&CrT continuum) was examined in Chapter 2. Although the continuum was informed by relevant educational theories, my analysis found that these theories were not consistently or rigorously applied. Despite these limitations, as the Australian Curriculum is the mandated tool for teachers in Australia for use in planning, teaching and assessing, it is important that any assessment of student learning is aligned with the skills described in the continuum. In posing Research Question 2, 'How are critical and creative thinking skills measured during learning?', I had initially assumed that the assessment tools or measures being used for critical and creative thinking research would be fit for purpose. However, as this chapter demonstrates, this was not the case. To be fit for purpose for use in the classroom, a measure must both assess the skills being taught in the classroom (i.e., the skills described in the C&CrT continuum) and be suitable to identify the skills used by students during their learning.

Although the terms 'critical thinking' and 'creative thinking' appear to be inextricably linked to each other in the Australian Curriculum (Ab Kadir, 2018), academic researchers generally treat them as separate concepts (e.g., Baker et al., 2001; Thornhill-Miller et al., 2023). Ritchhart (2002) for example describes creative thinking as 'looking out, up, around, and about' and critical thinking as "looking at, through, and in between" (p27). Critical thinking — its definition, representation in the curriculum, and assessment — will be examined first, followed by creative thinking.

3.2. General Concepts of Critical Thinking

Any search for information on critical thinking will inevitably lead to the Delphi Report, 'Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction' (Facione, 1990). Led by Peter Facione on behalf of the American Philosophical Association, this

formative work drew together much (but by no means all) of the critical thinking research of the time from critical thinking theory and assessment research, in the fields of philosophy, education, social and physical sciences. The report helped shift the view of critical thinking as relevant primarily in such areas as philosophy or logic, to recognition of its role as an essential skill which was relevant for academic success and lifelong learning. Further, the Delphi Report provided a conceptual framework and guidelines for teaching and assessing critical thinking (Facione, 1990). Building on the background thinking of Dewey's 'reflective thinking' (Dewey, 1910, p. 3) and the philosophical argumentation structure of logical thinking (described as the 'philosophical tradition': (Davies, 2019; Lai, 2011; Moseley et al., 2005; Sternberg, 1986), the Delphi Report identified critical thinking as:

> purposeful, self-regulatory judgement which results in interpretation, analysis, evaluation and inference as well as explanation of the evidential conceptual, methodological, criteriological or contextual considerations upon which that judgement was based. (Facione, 1990, p. 3)

This definition was accompanied by descriptions of cognitive skills, dispositions, critical thinking instructions and assessments. Together, this provided a consensus view of critical thinking to aid the educational sector at all levels to focus on the elements of critical thinking that were identified as the most essential for effective critical thinkers.

Table 4 shows the six cognitive skills and 19 dispositions identified in the Delphi report as sitting at the core of critical thinking. The six skills are all based on the premise of critical thinking as the act of being critical, i.e., a means to develop and analyse an effective argument. This is described as the argumentation-and-judgement view (Davies & Barnett, 2015). This view does not identify critical thinking as a single skill, but as a collection of skills that, when used appropriately, enable the act of thinking critically to be carried out.

The report also signalled an awareness that it was not sufficient for a person to have the capability to use critical thinking *skills*, but that they needed particular "approaches to life and living in general…[and] approaches to specific issues, questions or problems" (Facione, 1990, p. 25). These 'approaches' are the 19 'dispositions' listed in Table 4, such as inquisitiveness, open-mindedness and the persistence necessary to apply them. The significance of this will be explored further when motivation and engagement are discussed in Chapter 5.

The Delphi Report, however, reflected the view of only one branch of scholarship among many, about critical thinking in the early 1990s. Moseley et al. (2005) and Sternberg (1986) comment on the great variety of views of critical thinking (Moseley collected nearly 40 definitions of

Table 4

Critical thinking skills and dispositions identified by the Delphi Report

Cognitive Skills	Dispositions
interpretation	inquisitiveness with regard to a wide range of issues
analysis	concern to become and remain generally well-informed
evaluation	alertness to opportunities to use critical thinking
inference	trust in the processes of reasoned inquiry
explanation	self-confidence in one's own ability to reason
self-regulation	open-mindedness regarding divergent world views
	flexibility in considering alternatives and opinions
	understanding of the opinions of other people
	fair-mindedness in appraising reasoning
	honesty in facing one's own biases, prejudices, stereotypes, egocentric or sociocentric tendencies
	prudence in suspending, making or altering judgments
	willingness to reconsider and revise views where honest reflection suggests that change is warranted
	clarity in stating the question or concern
	orderliness in working with complexity
	diligence in seeking relevant information
	reasonableness in selecting and applying criteria
	care in focusing attention on the concern at hand
	persistence though difficulties are encountered
	precision to the degree permitted by subject and circumstances

Note. Based on Facione (1990)

critical thinking in 2005; p. 20). Most of the theories fall within the three main perspectives of philosophical, psychological and sociological theories and their sub-sets (Moseley et al., 2005).

The view promoted by the Delphi Report falls within what can be described as the 'cognitive-philosophical' camp (Davies & Barnett, 2015, p. 11). Sternberg warns that theories of critical thinking "so often say similar things in different ways, or even occasionally different things in similar ways" (p. 8), and Davies and Barnett describe the concept of critical thinking as remaining "more elusive than ever" (Davies & Barnett, 2015). As an example of the wide variety of theories, an alternative interpretation of the term 'critical thinking' can be described from the 'critical pedagogy' view (Davies & Barnett, 2015). This branch of thinking came from German critical theory, and interpreted 'critical' as 'critique', i.e., drawing out dimensions of meaning in an argument that helps explain it in a broader context (Kaplan, 1991, p. 362). Kaplan (1991) describes this alternative view of critical thinking as incorporating concepts of phenomenology ("observing and articulating the structures of lived experience"; *p* 362) and psychoanalysis, which goes well beyond a list of what students should know-and-do (the 'argumentation-and-judgement view' of the Delphi Report), to a broader concept of enabling "a critique of lived social and political realities with the aim of changing those realities to allow greater freedom of thought and action" (1991, p. 362).

This discussion of the breadth of understanding of the term critical thinking provides a context for the complexity of the skills described in the Critical and Creative Thinking learning continuum. The descriptions of critical thinking so far provided represent concepts relating certainly to what adult critical thinkers are expected to be capable of, and also, to some extent, what can be observed in adolescent critical thinkers. The context of the research in this thesis however concerns primary school aged students, and cognition research indicates that some of these concepts may only apply to a lesser extent in critical thinking in younger children (Kuhn, 1999). This is of significance as it brings into question the relevance of research based on tertiary and secondary students, when working with primary aged students. The lack of research relating specifically to the development of critical thinking skills in primary aged students is a significant gap in the literature.

The differences in the ability to use critical thinking, brought about by cognitive developmental stages, is demonstrated by Kuhn's research on the developmental sequences of second-order cognition (Kuhn, 1999) and has implications for how we view critical thinking in children. In her paper 'A Developmental Model of Critical Thinking' (1999), Kuhn describes the development of three forms of second-order cognition (metacognitive, metastrategic and epistemological) that are essential for critical thinking, and links them to developmental stages.

Although the critical pedagogy view may have a valid place in critical thinking, Kuhn's research suggests that the level of epistemological understanding required to engage in this level of critique is not generally reached until adolescence (1999, p. 22), i.e., past the years of primary school education, and in fact according to Kuhn may never be achieved by some adolescents or even some adults. This research by Kuhn (1999) provides additional depth to our understanding of how particular types of critical thinking are linked to the stage of development of the person. It also alerts us to the need for teachers to modify their expectations when working with students of different ages. Additionally, it demonstrates a weakness in the critical pedagogy view, which fails to take these factors into account. The implications of this for working with the Continuum are discussed below.

There is no single definition of the construct of critical thinking (Davies, 2019; Shaw, 2014) undoubtedly in part because of the multiple, overlapping concepts (e.g., the argumentation-and-judgement view, critical pedagogy, developmental stages of meta-cognition) which have previously defied agreement. There is more recently, however, an acknowledgment of the relevance of the structures of argumentation, without neglecting the elements of critical pedagogy, in what Paul (2011) describes as the 'third wave' of the critical thinking movement (see also Santos Meneses, 2020). This third wave applies what can be described as a moderating effect on critical thinking theories, enabling a variety of views to be represented. This is aptly illustrated by Ennis's widely used definition: "Critical thinking is reasonable reflective thinking focussed on deciding what to believe or do" (Ennis, 1993, p. 180). This definition, which builds on the Delphi definition by including both skills (reflective thinking) and a requirement for some judgement (deciding what to believe or do, i.e., a disposition) is described by Davies and Barnett as the 'skills-and-judgements' view (Davies & Barnett, 2015, p. 11).

Ennis's definition, and the 'skills-and-judgement' view represent the popular view of critical thinking currently in use (Davies & Barnett, 2015; Hatcher, 2015). Recognition that the very disparate views of critical thinking discussed so far co-exist within the sphere of education and are representative of only a few of the different ways of conceptualizing critical thinking, serves as a reminder of the importance of clarifying what is meant when the term 'critical thinking' is used, and critical thinking measures are being assessed (Ab Kadir, 2018).

3.3. Comparing Critical Thinking Theories with the Critical and Creative Thinking Continuum

As will be seen below, one effect of a lack of consensus in the definition of critical thinking, is the lack of consistency in the terminology used to describe the various skills included in the measures

Table 5

Comparison of terminology for critical thinking skills between a selection of theories at the level of 'elements' in the Australian Curriculum

Australian Curriculum[1]	The Delphi Consensus[2]	OECD[3]	The Critical Thinking Matrix[4]	Watson-Glaser Critical Thinking Appraisal[5]	Cornell Critical Thinking Tests[6]	California Critical Thinking Skills Test[7]
Inquiring	Analysis	Inquiring	Analysis	Deduction	Meaning	Analysis
Generating	Interpretation	Imagining	Interpretation	Interpretation	Deduction	Deduction
Reflecting	Evaluation	Doing	Evaluation	Evaluation of arguments	Value judgement	Evaluation
Analysing	Inference	Reflecting	Inference	Inference	Induction	Inference
	Explanation		Explanation	Recognition of assumptions	Assumptions	Induction
	Self-regulation		Self-regulation		Credibility	
					Observation	

Note.[1] Australian Curriculum Assessment and Reporting Authority (ACARA) (n.d.-b). [2]Facione (1990). [3] Vincent-Lancrin et al. (2019). [4]Ellerton (n.d.). [5]Watson (1980). [6] Ennis et al. (1985). [7] Facione (1991).

used. A further effect is the lack of consistency in the emphasis on the choice of skills to incorporate in any measure - or indeed in any curriculum. To inform the selection of suitable measures which I intended to use in my empirical studies, I compared the sets of skills identified in the most widely reported theories on critical thinking against the Australian Curriculum Critical and Creative Thinking learning continuum. The sets of skills described in the theories are referred to as scales and sub-scales (consistent with the Delphi Report), and the sets of skills within the curriculum are described as elements and sub-elements (consistent with the curriculum). My initial comparison was conducted at the level of the elements and scales. At this level of detail however, I found that there is insufficient information on which to base a comparison. As seen in Table 5, the term, 'Analysis' for example, is indicated across most of the lists of skills, and terms denoting inquiry or interpretation are present, but details of what is meant by the terms are not provided. While the continuum's presentation of sub-elements does define the skills in greater detail, only a few of the scales and sub-scales found in the literature include this quantity of detail.

Only three lists of skills were found which provided a sufficiently fine amount of detail to enable a comparison with the C&CrT continuum at the value of the sub-elements. Two of these are shown in Table 6 — the Delphi Report, and the OECD domain general rubric on critical thinking. The third skills list — Ellerton's Critical Thinking Matrix (Ellerton, n.d.) — although providing a suitable level of detail, is almost identical to the Delphi Report structure. As there is no measurement tool associated with Ellerton's matrix, it was not included at this level of examination.

Some of the skills described in the C&CrT learning continuum and the OECD Rubric have no equivalent in the Delphi skills list, as they refer to creative thinking skills, and the Delphi report is focused solely on critical thinking. It can also be seen that there are two skills identified in the Delphi Report ('Stating results' and 'Presenting arguments') that have no corresponding skill description in the C&CrT continuum.

My comparison of the continuum with the OECD Rubric demonstrates the importance of looking more deeply into the meaning behind the terms used to denote skills when making comparisons. Although both the C&CrT learning continuum and the OECD skills list identify 'Reflecting' as a skill at the element level, when the subscale/sub-element value statements are compared, there are significant differences between the details of the skills being described by the OECD and the C&CrT skills. By comparing the statements at the sub-element level in this way, the similarities and differences between the descriptions of skills in the various rubrics become more pronounced: e.g., the C&CrT continuum uses the terminology of 'Posing questions', and 'Identify and

Table 6

Comparison of terminology for critical thinking skills between the C&CrT continuum, Delphi Report, and OECD critical thinking rubric

C&CrT continuum[1]		OECD domain-general rubric[2]	The Delphi Report[3]
Element	Sub-element	Critical Thinking Subscale	Subscale
Inquiring	Pose questions	Identify and question assumptions, … (partial statement) (Inquiring)	Clarifying meaning (Interpretation) Examining ideas (Analysis) Identifying arguments (Analysis)
	Identify and clarify information and ideas	Understand context/frame and boundaries of the problem. (Inquiring) … check accuracy of facts and interpretations … (partial statement) (Inquiring)	
	Organise and process information	… analyse gaps in knowledge (partial statement) (Inquiring)	Categorization (Interpretation) Decoding significance (Interpretation)
Generating	Imagine possibilities and connect ideas	No equivalent	No equivalent
	Consider alternatives	Identify and review alternative theories and opinions and compare … (partial statement) (Imagining)	Conjecturing alternatives (Inference)
	Seek solutions and put ideas into action	… imagine different perspectives on the problem. (partial statement) (Imagining)	No equivalent

Note. [1]Australian Curriculum Assessment and Reporting Authority (ACARA) (n.d.-b). [2]Vincent-Lancrin et al. (2019, p. 27). [3] Facione (1990).

Table 6 (continued)

C&CrT continuum[1]		OECD domain-general rubric[2]	The Delphi Report[3]
Element	Sub-element	Subscale	Subscale
Reflecting	Think about thinking (metacognition)	Reflect on the possible bias of one's own perspectives compared to other perspectives. (Reflecting)	Self-examination (Self-regulation)
	Reflect on processes	No equivalent	Self-correction (Self-regulation)
	Transfer knowledge into new contexts	No equivalent	No equivalent
Analysing	Apply logic and reasoning,	No equivalent	Assessing claims (Evaluation) Assessing arguments (Evaluation) Querying evidence (Inference)
	Draw conclusions and design a course of action	Evaluate and acknowledge the uncertainty or limits of the endorsed solution or position (Reflecting)	Conjecturing alternatives (Inference) Drawing Conclusions (Inference)
	Evaluate procedures and outcomes	Justify a solution or reasoning on logical, ethical or aesthetic criteria/reasoning (Doing) Identify Strengths and weaknesses of evidence, arguments, claims and beliefs. (Imagining)	Justifying procedures (Explanation)
	No equivalent	No equivalent	Stating results (Explanation) Presenting arguments (Explanation)

Note. [1] Australian Curriculum Assessment and Reporting Authority (ACARA) (n.d.-b). [2]Vincent-Lancrin et al. (2019, p. 27). [3]Facione (1990).

clarify information and ideas' within the element of Inquiring. The Delphi Consensus talks about 'Clarifying meaning', 'Examining ideas' and 'Identifying arguments' straddling their skills of Interpretation and Analysis. Although given different names and grouped with different skills, the choice of terminology in the names points to similarities in their intent, indicating that these the skills are most likely comparable.

Two groups of skills described at the element level in the Delphi report skills list, 'explanation' and 'evaluation', receive considerably less attention in the C&CrT learning continuum than in the Delphi list. The *explanation* skills relate to stating results and presenting arguments. The *evaluation* skills are specifically those relating to assessing an argument or claim. Despite argumentation being the major focus of critical thinking in its original conception (the argumentation of Aristotle: Ellerton, 2017; Hung et al., 2008; Moseley et al., 2005), it receives only brief attention as part of the skill of using logic and reasoning within the C&CrT continuum. This seems to be a curious oversight in the light of the statement on the Australian Curriculum website, that critical thinking is "at the core of most intellectual activity that involves students learning to recognise or develop an argument, use evidence in support of that argument, draw reasoned conclusions" Australian Curriculum Assessment and Reporting Authority (ACARA) (n.d.-a).

Although Ennis's 'skills-and-judgement' definition ("Critical thinking is reasonable reflective thinking focussed on deciding what to believe or do": Ennis, 1993, p. 180) is described as representing the popular view of critical thinking currently in use (Davies & Barnett, 2015; Hatcher, 2015), it fails to acknowledge the vital significance of the dispositions required for effective critical thinking. For the purposes of this thesis, critical thinking will therefore be defined as a collection of skills which enable focussed, purposeful, reflective thinking, and rely on the presence of various dispositions to be successfully enacted. Where Ennis's definition restricts critical thinking to reflective thinking, this new definition allows for a broader set of skills, e.g. allowing for skills such as the ability to recognise or develop an argument, use evidence in support of that argument and draw reasoned conclusions, as well as others not specified here. Where Ennis limits the outcome of the use of critical thinking to 'deciding what to believe or do', the new definition describes the outcomes that are enabled (focussed, purposeful, reflective thinking). In addition, the new definition also emphasises that for critical thinking to occur it is not sufficient to have mastered a skill, if the disposition to use it effectively is not present also.

The Delphi report skills list and the C&CrT continuum skills show strong similarities. The OECD report, while containing sufficient detail to enable comparison, was not well aligned to the Australian Curriculum. At the conclusion of the analysis it was therefore anticipated that measures based on the Delphi Consensus interpretation of Critical Thinking were more likely to be suitable for assessing the skills described in the C&CrT continuum.

3.4. Measuring Critical Thinking

Although the Delphi Consensus has been identified as the theoretical approach most closely aligned with the C&CrT continuum, alignment to the continuum structure is only one of the requirements for a suitable measure. For any measurement task to be fit-for-purpose (e.g., see Cropley, 2023; Race, 2014), it must be suitable for use in South Australian Primary School classrooms, and for use by students with a range of language abilities. The criteria shown in Table 7 was developed based on classroom observation and experience, to assess the suitability of measures in the context of these requirements.

A review of the literature was carried out to identify measures which were described as critical thinking measures, or which measured components of critical thinking (a full list can be found in Appendix B). Twenty-one tasks were identified through the literature review. The characteristics of the measures were first identified (skills measured and format/focus) as shown in Table 8. The measures were then assessed for suitability for use in the classroom by comparing each measure against the criteria in Table 7. The most used and reported assessments are summarised in Table 8. The final column, 'Meets criteria', lists the criteria from Table 7 which apply to each of the tests.

Based on the descriptions of critical thinking measures provided by the developers of the measures, four classifications of critical thinking measures were identified: those that assess critical thinking skills; those that assess critical thinking skills and dispositions; those that assess just critical thinking dispositions; and those that assess problem-solving skills. Measures which only assess dispositions, or that measure only one aspect of critical thinking, were considered unsuitable for this research as they did not provide sufficient coverage of the C&CrT continuum skills. Problem-solving skills will be discussed later in this chapter, following the analysis of creative thinking measures.

Table 7

Criteria for suitability for measures of Critical Thinking, used in the primary school classroom

Number	Criteria
1	Validated for use with primary school students
2	Measures the skills identified in the C&CrT continuum/ Delphi report
3	Identifies the level at which the skill is being used, as described in the C&CrT continuum
4	Has sufficient granularity[1] to identify change over a period of 10 weeks
5	Does not rely on specialised equipment or computers
6	Is inclusive of students with low literacy skills[2]
7	Is easy to administer in the classroom

Note. [1]Granularity is identified as the level of detail in the information (Victorian Curriculum and Assessment Authority (VCAA), 2019)which can be identified by the measurement, (Rovers et al., 2019) e.g., a coarse-grained level would measure the presence or absence of a skill, while a fine-grained measure would be capable of measuring different levels of application of the skill.[2] This stipulation was as a response to the low reading ability of some of the students (mainly boys) who appeared to show significant development in their skills during the STEM Learning Evaluation study but were unable to complete the assessment task successfully due to their low literacy levels.

As the C&CrT continuum was developed relatively recently, and is only based loosely on research, it is not surprising (as noted by Scoular & Care, 2018), that none of the measures were found to assess the skills identified in the C&CrT continuum, nor to identify the level of skill achieved as described in the C&CrT continuum. Four measures were found that had been validated for use with primary school students, but they all either relied on the use of technology for testing (and were only available for on-line testing) or included a significant load of reading. As not all primary school students in South Australia have access to a digital device, inconsistent levels of skills may affect the results of testing using this method. A heavy reliance on text in the measure may also skew the results for students with less developed skills in reading.

Following this investigation of the literature it was determined that an initial, partial response to Research Question 2: 'How are critical and creative thinking skills measured during learning?', was that there were no existing tasks that would effectively measure the *critical* thinking skills as described in the C&CrT continuum, during student learning. A new measure would therefore

Table 8

Measures of critical thinking skills

Name	Skills measured	Reference	Age group	Format/ Focus	Meets criteria[1]
California Critical Thinking Skills Test: (CCTST) [2,3,4]	Analysis Evaluation Inference Inductive reasoning Deductive reasoning	P. Facione & N.C. Facione 1994 (Insight Assessment, 2023a)	Tertiary + upper secondary	Multiple choice Self-report On-line	2 (partial) 7
Educate Insight Grades 3-5 (Form of CCTST)	Habits of mind + Inductive Reasoning, Deductive Reasoning, Evaluation, Analysis, Inference, Numeracy	Insight Assessment (2023c)	Grades 3-5	On-line Self-report + scenario based 75 mins	1 2 (partial) 7
Cornell Critical Thinking Tests Level X [2,5] (CCTT-X)	Induction, deduction, Identification of assumptions, credibility	Ennis et al. (1985)	Grade 5 and up	Multiple choice On-line or paper Requires reading 50 mins	1 5 7
Cornell Critical Thinking Tests Level Z [2,5] (CCTT-Z)	Induction, deduction, Identification of assumptions, credibility, semantics, definition, prediction in planning	Ennis et al. (1985)	Grade 5 and up	Multiple choice On-line or paper Requires reading 50 mins	1 5 7

Table 8 (continued)

Name	Skills measured	Reference	Age group	Format/ Focus	Meets criteria
Ennis-Weir Critical Thinking Essay Test[2]	Getting the point, seeing the reasons and assumptions, stating one's point, offering good reasons, seeing other possibilities, responding to and avoiding equivocation, irrelevance, circularity, reversal of conditional relationships, overgeneralization, credibility	Ennis and Weir (1985)	Grade 7 and Upwards + Tertiary	Paper Requires reading and writing context outdated—letter to newspaper, topic not relevant to children—parking issues, driving, etc.	5 7
Watson-Glaser Critical Thinking Appraisal[2] (W-GCTA)	Inference, assumptions, deduction, interpretation, evaluation of arguments	Pearson (2023); Watson and Glaser (1980)	16 years+	On-line Paper version 30 mins Random generation of questions from item-bank	2 (partial) 7

Note. [1]Criteria listed in Table 7. [2]Ennis (1993). [3]One of the most used tests of CT according to Martin Davies (2019). [4]Hatcher (2015). [5]Hasinger (2023).

need to be developed with a focus on the skills described in the C&CrT continuum, in order to measure critical thinking skills during learning activities.

3.5. General Concepts of Creativity and Creative Thinking

One of the challenges for creativity is that it is addressed by a number of parallel disciplines, sometimes in different ways (Kaufman, 2009, p4). Creativity has its origins in psychology research (Cropley, 2015a, p. 7), with a focus on understanding the cognitive, biological, and individual factors that contribute to creativity. According to Baas et al., (2015), the stereotypical layperson's view of creativity is that it happens unexpectedly, and is the result of unfocussed, "out-of-the-box" thinking. Creativity researchers however recognise that being creative, and creative thinking, involve much more than just original, unstructured thought (Beghetto & Karwowski, 2017; Feist, 2010; Lubart et al., 2013). For the purposes of recognising and measuring creative thinking, it is therefore essential to explore the origins of current thinking about creativity (Lassig, 2013), and its relationship to creative thinking. This section commences therefore with an examination of creativity, followed by clarification of the meaning of the term 'creative thinking', and a reconciliation of the psychology view of creative thinking with its application to education.

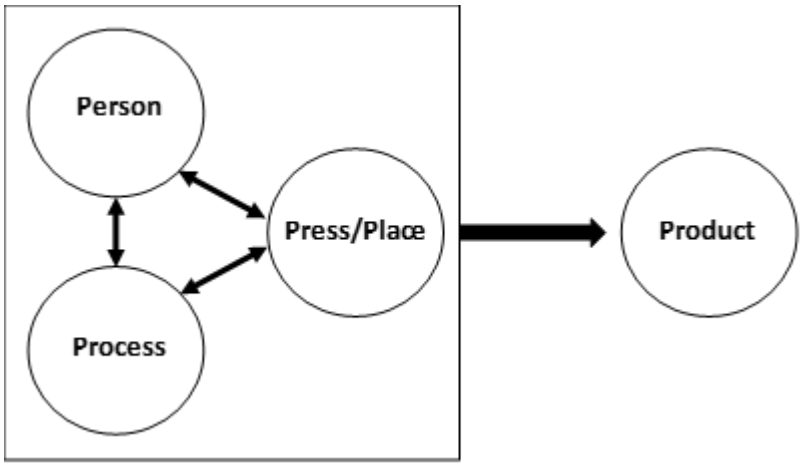
A widely used definition, developed by Plucker, Beghetto and Dow (2004) describes creativity as:

> …the interaction between aptitude, process, and environment by which an individual or group produces a perceptible product that is both novel and useful as defined within a social context (p90).

Plucker, Beghetto and Dow's 2004 definition drew from an analysis of research literature to identify the explicit terms and implicit assumptions being used by creativity researchers to define creativity. It builds on the 4Ps framework, developed in the 1960s to characterise creativity (Barron, 1969; Rhodes, 1961). While each of the 4Ps - Person, Process, Press and Product - has unique properties, effective creativity requires the four strands to operate in unity (Rhodes, 1961, p. 307). This is effectively illustrated in Cropley's illustration in Figure *7*.

The definition begins by using the terms 'aptitude', 'process', and 'environment' to identify three factors which impact on a person's capacity for creativity. First, the 'aptitudes' of the individual, are understood to include the personality, intellect, temperament, behaviours, etc., which are characteristics of the individual. This is described as **person**. These personal properties (e.g., extroversion, agreeableness, openness and conscientiousness), combined with motivation and personal feelings, are known to influence how creative a person can be (Cropley, 2015b). Second,

the **process** of being creative, refers to the *cognitive* processes (as distinct from an operational process): how the creative person thinks—their perception, learning and thinking abilities. It is in Process that *creative thinking* sits. Third is the 'environment' in which the person is operating. This includes both the influence of the social and physical environment upon the person and their mental



Note. From (Cropley, in press). Printed with permission.

Figure 7

A systems model of creativity

processes, to either support or inhibit creativity. This is labelled **press**. It also includes the response of the environment to the creative act, e.g., whether the creative act is responded to positively or negatively by those in the environment. The relationship between these three factors (person, process and press) is defined as an 'interaction', as it is recognised that while they form a system together, each of these factors also plays a part individually. Describing the relationship in this way also hints at the possibility of enhancing creativity by the beneficial improvements of any one or all of these factors (Plucker et al., 2004), which inevitably has implications for education.

A fourth factor in the definition, which impacts on whether an act is deemed to be creative, is the 'product', which is the result of the act of creativity. The product is the embodiment of an idea into a tangible form (this includes both physical and conceptual products), but to be considered creative the product must be both novel and useful (Rhodes, 1961; Runco & Jaeger, 2012; or novel and elegant as described by Cropley & Cropley 2008). Although there is continuing discussion over what constitutes novel and useful (e.g., Puryear & Lamb, 2020) it is recognised that novelty and

usefulness are defined by the social context (including geographical, chronological and cultural) in which the product is being judged.

Before focussing on the process strand and *creative thinking*, as well as understanding the factors that affect creativity, and how creative thinking relates to creativity as a whole, it also helps to understand *how* creativity happens. The part played by creative thinking can then be more easily isolated and examined.

3.6. Models of Creativity

There have been many models developed to describe the steps which are required for the making of a creative product — also described as creative problem solving. Known generally as process models, the term 'phase models' will be used here to avoid confusion between process as a cognitive function (cognitive processes), and process as an operation (i.e., the steps required to complete an activity). Phase models describe the operational phases of developing a creative solution to a problem. One of the earliest and most recognised phase models is a four-step model developed by Wallas in 1926 (Cropley, 2015b; Mumford et al., 1991; Torrance, 1993). Wallas described the creative act ("the making of a new generalization or invention, or the poetical expression of a new idea": 1926, p. 79) as consisting of the four phases: Preparation, Incubation, Illumination and Verification (Wallas, 1926, p. 80). Illumination was described as an un(sub)conscious phenomenon which may include "sudden flashes of emotional experience too momentary for description, feelings of queerness, or surprise, or recognition, or amusement" (Wallas, 1926, p. 79) perpetuating the myth of a mystical component to creativity. Osborne's (1953) subsequent model was influential in dispelling the mystical component myth by removing the term Illumination, and inserting three additional steps in the Preparation phase, as shown in Table 9. Osborne's three additional steps of Analysis, Hypothesis Generation and Synthesis effectively signalled the application of cognitive rather than ethereal processes at work (Mumford et al., 1991, p. 97). More recent versions of the creative process, e.g., the Creative Process Model, (Mumford et al., 1991) and the Extended Phase Model (Cropley, 2015b) include two additional phases (communication and validation) occurring after Wallas's Verification phase. The addition of the two further phases extends models of creativity to include the necessary acts of enabling the creative product to be proven both novel and useful. This is illustrated in Table 9 by the steps of Communication and Validation in Cropley's Extended Phase Model (Cropley, 2015b).

Having discussed both the factors which affect creativity (the 4Ps), and the steps required to achieve a creative outcome (phase models), it is important to acknowledge the relationship between

Table 9

Comparison of three creative process models

4-Step Model[1]	Osborne[2]	Extended Phase Model[3]
Preparation	Orientation Preparation Analysis Hypothesis generation	Preparation Activation
Incubation	incubation	Generation
Illumination	synthesis	Illumination
Verification	verification	Verification
		Communication Validation

Note. [1]Wallas (1926). [2] Osborn (1953). [3]Cropley (2015b).

these two frameworks in order to understand the role of creative thinking in the creative act. This relationship is effectively demonstrated by Cropley's (2015b) Innovation Phase Model.

3.7. The 4Ps and Phase Models Combined

The Innovation Phase Model plots the phases of the Extended Phase Model (Cropley, 2015b) against the factors of the 4Ps framework, as shown in Table 10. Creative thinking, as indicated by the 'Process' factor in the 4Ps model, is found in the shaded cells. Each cell in the table indicates an opportunity to influence the creative act, either by encouraging creativity or discouraging creativity. For example, the Preparation phase (where knowledge acquisition and problem recognition occur), can encourage the development of creative responses firstly through a stimulating environment (Press – cell A); secondly through suitable motivation and personal properties (Person – cell B); and lastly through engaging in the right sorts of thinking (Process – cell C). Understanding these 'sorts of thinking', described by J. P. Guilford in the 1950s, enables us to begin to understand 'creative thinking'.

In 1959 the psychologist J. P. Guilford called for an increase in research on the thinking processes involved in creativity (Guilford, 1959). His theory on types of thinking is recognised as a watershed moment in creativity research.

In his work studying human intelligence, Guilford identified and described two complimentary ways of thinking – convergent and divergent thinking (Guilford, 1959). Guilford's work was prompted by his perceived need to identify the characteristics of creativity and creative people. Guilford cited many of the same claims in the 1950s for the importance of creativity, as are cited today: the need for creativity in science and engineering (STEM), the economic value of creative ideas, and the threat of technology making all but the creative classes obsolete (1959, p. 446). Along with Wallas, Guilford recognised that if creativity was to be intentionally developed, the process of being creative, and what made a person creative, needed to be understood. Understanding the components of creative thinking enables choices to be made about appropriate assessments to effectively measure those components.

The two different types of thinking described by Guilford, convergent and divergent, are both now recognised as playing an important role in creativity and therefore in creative thinking (Cropley, 2006). Although Guilford's initial descriptions of divergent thinking resulted in the interpretation of it being synonymous with creativity (Runco & Acar, 2019), divergent thinking is now believed to occur sequentially with convergent thinking during creative thinking (Cropley, 2006; Runco & Acar, 2019). Convergent thinking skills enable us to make choices and decisions (Guilford,

Table 10

The Innovative Phase Model structure

The Factors of the 4Ps	The phases of the Extended Phase Model						
	Preparation	Activation	Generation	Illumination	Verification	Communication	Validation
Process	C						
Person	B						
Product							
Press	A						

Note. Based on Cropley (2015b, p. 223), and Cropley and Cropley (2015)

Table 11

Aligning the Extended Phase Model with convergent and divergent thinking

The Factors of the 4Ps	The phases of the Extended Phase Model						
	Preparation	Activation	Generation	Illumination	Verification	Communication	Validation
Process	Convergent thinking	Divergent thinking	Divergent thinking	Convergent thinking	Convergent thinking	Divergent / Convergent thinking	Convergent thinking

1959). They enable us, for example, to identify the one best solution by evaluating options, or determining which is the most relevant information to use (Cropley et al., 2023). Some of Guilford's convergent thinking skills (e.g., analysis and evaluation), have already been discussed as critical thinking skills. Divergent thinking on the other hand is the thinking skill which enables the generation of original and varied ideas (Acar & Runco, 2012). It has been modelled as comprising four elements: fluency – the creation of many ideas; flexibility – the production of a variety of ideas or use of a variety of approaches; originality – the creation of new ideas; and elaboration – the ability to fill in the details (Patston et al., 2022; Torrance, 1977). Divergent Thinking is 'cognition that leads in various directions' (Runco, 1999b), or 'lateral thinking' (Dingli, 2008) and is most often recognised as a skill which creative people possess (Hocevar, 1981).

The relationship between the cognitive thinking processes and the Extended Phase Model (i.e., what is happening in the shaded top row of Table 10), is shown in Table 11**.** Both divergent and convergent thinking have their own role to play during the development of a creative product. If creative thinking is the cognitive processes which are required to enable the creative product to be realised, then there is no question that both convergent thinking and divergent thinking play a role in creative thinking (Gralewski & Karwowski, 2019).

One further characteristic of creative thinking which remains to be described is the dynamic relationship between the two types of thinking (convergent and divergent). Although illustrated in Table 11 as linear, with divergent thinking neatly following convergent thinking, with each phase following the one before in an orderly fashion, a non-linear characteristic of the creative thinking process is indicated by the literature. Guilford asserts that evaluation is occurring at every step in "a perpetual system of checks and balances" (Guilford, 1967, p. 329); Cropley (2015b) describes the process as possibly incorporating "interactions, false-starts, restarts, early break-offs, and the like" (p. 49). Haner (2005) describes the "complex, partly iterative and partly simultaneous" process (p. 289) and Amabile depicts a nonlinear, iterative and non-sequential creative process (1996, p. 114)

The theories described so far, and the research supporting them, indicate:

- that creative thinking occurs in the cognitive process strand of the 4Ps framework,
- that creativity/creative problem solving is a complex activity including various steps or phases, and that creative thinking occurs in each of those phases,
- that creative thinking requires the use of two distinct types of cognitive process (convergent and divergent thinking),
- that creative thinking is a nonlinear and iterative activity.

Based on these understandings from psychology research, creative thinking can therefore be understood as a competency used to engage in the development of creative products, which incorporates divergent and convergent thinking, and is influenced by external factors such as the person and place associated with the creative act.

It remains to clarify how the view of creative thinking in psychology compares with the view of creative thinking in education. Although the investigation of the literature has included sources from both education and psychology, the majority of research into creativity has occurred through educational psychology research, rather than from a practical education (teaching) perspective. As the research questions relate specifically to the field of applied education, clarifying this relationship is relevant. I will also explore how creative thinking and critical thinking are related, as both have been seen to contain the same thinking element of convergent thinking.

3.8. Creative Thinking in Education

As calls for an increased focus on teaching for creativity have increased (e.g., Care et al., 2016; Torii, 2018), so has the need to measure creative thinking. In education we both value what we measure, and need to measure what we value (Foster & Schleicher, 2022; Mislevy, 2012). If we value creative thinking but are unable to measure it, educational factors which positively or negatively influence creative thinking in students cannot be accurately identified. For these reasons educational researchers both within Australia and internationally, have been developing forms of assessment to measure creative thinking. Before reviewing the assessment measures, the characteristics of creative thinking used to devise those measures must be clarified to ensure their applicability.

The most recent definitions of creative thinking coming from the OECD, (Organization for Economic Co-operation and Development), and from ACER (Australian Council for Educational Research), were developed for use in association with assessments of creative thinking in schools. The OECD definition was developed for use in the PISA 2021/2022 Creative Thinking assessment (OECD, 2020). ACER's definition was developed in association with an assessment task to assess critical thinking and creative thinking in collaborative problem solving (Scoular et al., 2020).

ACER's definition of creative thinking is:

> the capacity to generate many different kinds of ideas, manipulate ideas in unusual ways and make unconventional connections in order to outline novel possibilities that have the potential to elegantly meet a given purpose (Ramalingam et al., 2020a).

The OECD definition of creative thinking is:

> the competence to engage productively in the generation, evaluation and improvement of ideas, that can result in original and effective solutions, advances in knowledge and impactful expressions of imagination (OECD, 2020).

The two definitions were examined for relevance against the theoretical underpinnings of psychology-based creativity research and compared in Table 12. As expected, there are similarities and differences between the definitions. The OECD definition and Psychology understanding identify creative thinking as more than a single skill, while the ACER definition identifies it "as a unidimensional construct or singular skill" (Education Services Australia (ESA), 2019, p. 11). By describing it as a competency[13], the inference is that skills, knowledge etc. will combine to enable creative thinking. Both education definitions agree that creative thinking is being applied to an idea. A significant difference between the two education definitions is that the OECD definition anchors creative thinking firmly in the phases of generation and illumination by specifying the three skills of generation (divergent thinking), evaluation (convergent thinking) and improvement (convergent and divergent thinking). The ACER definition places a heavy focus on the divergent thinking skills of generating, manipulating, and making connections, with convergent thinking only really described in the evaluation or verification phase (Ramalingam et al., 2020b). The biggest difference between the education definitions and the psychology understanding is that the psychology understanding recognises the influence of other factors (e.g., Person and Place) as having an effect on creative thinking.

With its clear inclusion of both convergent and divergent skills, and its recognition of creative thinking as a competency, the OECD definition provides the best alignment of creative thinking to creativity research and will therefore be used to define creative thinking in this thesis.

[13] A competency is understood to include a set of skills, knowledge, abilities and attitudes, see Cropley, D. H. (2019a). Creativity, innovation and the future of engineering. World Engineers Convention (2019: Melbourne, Australia), , Thornhill-Miller, B., Camarda, A., Mercier, M., Burkhardt, J.-M., Morisseau, T., Bourgeois-Bougrine, S., . . . Mourey, F. (2023). Creativity, Critical Thinking, Communication, and Collaboration: Assessment, Certification, and Promotion of 21st Century Skills for the Future of Work and Education. *Journal of Intelligence*, *11*(3), 54.

Table 12

Comparison of education definitions of creative thinking (CrT) with the psychology understanding of creative thinking

Components of the definition	OECD	ACER	Psychology understanding
What is CrT	A competence	A capacity	A competency
What does CrT enable	to engage productively	to generate many different kinds	to engage in the development of
What affects CrT	—	—	external factors such as person and place
What skills are required	generation, evaluation and improvement	manipulation of ideas, making unconventional connections	divergent and convergent thinking
What are the skills applied to	ideas, also solving problems, (inferred by the production of solutions)	ideas	the development of
What are the products	original and effective solutions, advances in knowledge, impactful expressions of imagination	outlines of novel possibilities that have the potential to elegantly meet a given purpose	creative products

The relevance of investigating the relationship between creative thinking and creativity, and the significance of examining the different definitions of creative thinking, was to enable the most suitable assessments of creative thinking to be identified. As creative thinking incorporates convergent thinking, which is also a component of critical thinking, it is also necessary to clarify the relationship between critical thinking and creative thinking, before identifying a suitable measure or assessment task.

3.9. The Relationship Between Creative Thinking and Critical Thinking

Critical thinking and creative thinking are described as an inextricably linked set of skills in the Australian Curriculum (Ab Kadir, 2018), and are commonly referred to in schools as CCT (e.g., Coleman, 2022; Education Services Australia (ESA), 2019) as though they are a single concept. Although the review of the literature into critical thinking and creative thinking has clearly demonstrated the differences between them, they do, however, have in common the use of one type of thinking – convergent thinking. It is not clear from the literature whether there are forms of divergent thinking that are not used in creative thinking. Neither is it clear if there are forms of convergent thinking that are not used in critical thinking, however Figure 8 is provided as a guide to the relationship between convergent thinking and divergent thinking, and their relationships with critical and creative thinking. Although the subtleties of the relationship between these types of thinking are not demonstrated by this illustration, it enables both critical thinking and creative thinking to be represented as competencies that incorporate much more than just thinking skills. The model demonstrates the nature of:

- divergent and convergent thinking as being two distinct forms of cognitive process,
- convergent thinking as a subset of critical thinking,
- creative thinking as incorporating both convergent and divergent thinking skills.

Figure 8 also clarifies that both critical thinking and creative thinking use many of the same convergent thinking skills, but that creative thinking alone is understood to include divergent thinking.

The significance of understanding this relationship is that although *some* of the same convergent thinking skills are used in both critical and creative thinking, some critical thinking skills in the Australian Curriculum, e.g., metacognition, are not convergent thinking and are not components of creative thinking. Creative thinking assessments, although potentially measuring those convergent skills that are required for effective *creative* thinking, cannot be assumed to be

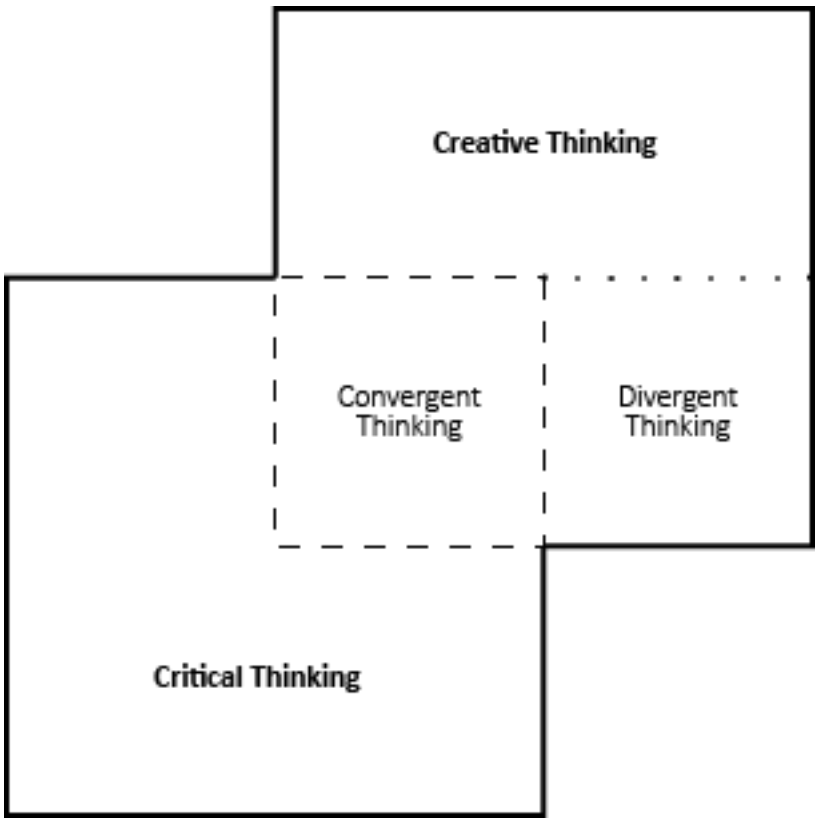


Figure 8

Relationships between the terms critical and creative thinking, divergent and convergent thinking

capable of measuring all the *critical* thinking skills. Therefore, in response to Research Question 2 we can conclude that both convergent and divergent skills must be included in a measure suitable to assess the application of creative thinking skills, and that no single assessment task is currently available that can simultaneously measure the application of both critical and creative thinking skills.

3.10. The Domain Specific Nature of Critical and Creative Thinking

The question of whether critical and creative thinking skills are domain specific is under debate (Vincent-Lancrin et al., 2019). Research evidence appears to support both that the skills in question are domain specific, or alternatively, that they are not (e.g., see Kaufman & Baer, 2005; Plucker & Beghetto, 2004; Reiter-Palmon et al., 2012; Runco, 1987a; Silvia et al., 2009). There is also a middle ground supporting the existence of both domain specific and domain general skills and abilities (Barbot et al., 2016; Kaufman et al., 2009). This relates to both the person's application of the

creative skills (If a person is creative in one domain, can it be assumed that they are able to apply their skills to be creative in other domains?) and the assessment of creative skills (If a person undergoes a domain-specific assessment, do the results relate to the person's abilities in other domains also?) (Baer & Kaufman, 2005; Barbot et al., 2016; Barbot et al., 2011).

If we acknowledge that even some of the skills are domain specific (i.e., an ability to apply a skill effectively in one domain does not guarantee that the same skill can be equally well applied in other domains), then it follows that assessments will also need to be domain specific to avoid inaccurate results (Kaufman et al., 2017; Long et al., 2022). An appropriate measure would therefore not rely, for example, on artistic skills or language skills for the student to demonstrate their use of a skill in a domain other than the languages or the arts. As the original question for investigation in the research described in this thesis specifies the application of critical and creative thinking skills in a STEM domain, it follows that the results of testing will potentially be more accurate if the measurements are taken in a STEM domain.

3.11. Measuring Creative Thinking

Measures of creative thinking were investigated in response to Research Question 2, 'How are critical and creative thinking skills measured during learning?'

The literature was searched for measures suitable to measure the types of divergent and convergent thinking skills described (albeit indirectly) in the continuum. The list of additional criteria, previously used to determine the suitability of critical thinking measures, was used with the additional requirement that the measurement task must also be suitable for use to measure skills used in a STEM-domain, as shown in Table 13.

Of the hundreds of measures that exist for measuring creativity and creative thinking, those that were prominently reported, or that appeared potentially relevant to the selection criteria, were investigated for their relevance to the research question. Although the intention was to measure creative *thinking*, very few measures were found that measured the use of creative thinking skills during learning, therefore measures claiming to measure creativity (rather than creative thinking) were also examined for relevance. To narrow down the most suitable measures for the purpose of this thesis, the measures found were classified according to two criteria: the factor of creativity which they represented (the factors of the 4Ps), and how/by whom, the factor was assessed (the method of assessment). Suitable measures were then compared for reliability and validity prior to determining the most appropriate measure to use for the purposes of this thesis.

Table 13

Revised Criteria for suitability for measures of critical and creative thinking skills, used in the primary school classroom

Number	Criteria
1	Validated for use with primary school students
2	Measures the skills identified in the C&CrT continuum
3	Identifies the level at which the skill is being used, as described in the C&CrT continuum
4	Has sufficient granularity[1] to identify change over a period of 10 weeks
5	Does not rely on specialised equipment or computers
6	Is inclusive of students with low literacy[2] skills
7	Is easy to administer in the classroom
8	Is applicable to use in STEM-domain lessons

Note. [1]Granularity is identified as the level of detail in the information (Victorian Curriculum and Assessment Authority (VCAA), 2019) which can be identified by the measurement, (Rovers et al., 2019) e.g., a coarse-grained level would measure the presence or absence of a skill, while a fine-grained measure would be capable of measuring different levels of application of the skill. [2] This stipulation was as a response to the low reading ability of some of the students (mainly boys) who appeared to show significant development in their skills during the STEM Learning Evaluation study but were unable to complete the assessment task successfully due to their low literacy levels.

All the measurement tasks were found to measure one or more of the factors of the 4Ps, that is:

- measures of creative process, i.e., creative thinking skills: both the potential to use creative thinking skills, and the levels of skills applied in genuine situations requiring creative solutions.
- measures of creative person, which includes: measures of personality, temperament, behaviour etc, which are generally reported as measures of creative potential; and level of creativity achieved, e.g., mini-c to pro-c (Kaufman & Beghetto, 2009)

- measures of creative environments (press), e.g., the characteristics of the environment which promote (or not) creativity, and the level of response from the environment to the creative act or product.
- measures of creative product, i.e., assessing the level of creativity of a product.

There was some overlap in these categories, e.g., assessing the level of creativity achieved by a person (the second category), relies on categorising a product (the fourth category). However, the area of interest for this research (cognitive processes) was found to be exclusive to the measures of creative process (the first category).

All the measures compared were found to fit within three methods of assessment: 1) self-report, 2) objective assessment by others, e.g., teacher or peers assessing against an agreed standard such as a rubric, and 3) subjective assessment by others, e.g., teacher or peers assessing the agreed skill based on their personally held beliefs or understandings about that skill.

Categorising the measures by the form of assessment was useful, as the form of measurement can determine the appropriateness of a measure in different situations, e.g., in a primary school as distinct from a tertiary education setting. The complete list of measures investigated has been categorised and included in Appendix C, and examples of existing measures have been included in Table 14. Where no example is provided, no measures were found that fit that category.

Assessments of the strand or factor 'process' are the most common approach used to measure creativity by researchers (Kaufman et al., 2007; Plucker et al., 2011; Said-Metwaly et al., 2017; Zeng et al., 2011). Long et al. (2022) reported that 51% of the 816 articles in their review measured divergent thinking, indicating that there is considerable focus on assessing the cognitive processes. Of the tasks investigated here, 19% were categorised in the 'process' strand. As creative thinking has been identified as utilising both divergent and convergent thinking, it is important to ensure that any task identified as relevant is capable of measuring both these types of thinking.

Divergent thinking measures, sometimes reported as measures of creativity, are more accurately described as tests of creative potential (Acar & Runco, 2012; Barbot et al., 2019) as they test the ability of a person to produce a creative response within a specific set of circumstances (Lubart et al., 2019). This is not evidence that a person has, or can apply these skills in genuine situations requiring a creative solution (Runco & Acar, 2012). Furthermore the most widely used tests, e.g., the Torrance Tests of Creative Thinking (TTCT: Torrance, 1966, 2008) and the Wallach-

Kogan Creativity Tests (WKCT: Wallach & Kogan, 1965), only measure divergent thinking, despite the now widely-held belief that creative thinking relies on both divergent and convergent thinking skills (Cropley, 2006). Although measures of skill potential provide general information on how likely a student is to be able to apply a skill in certain circumstances, these measures of divergent thinking do not meet the requirements of Research Question 2, i.e., to measure the skill *during learning*. Tests of creative potential would be found in Table 14, in cells A to C. It is only assessment tasks in Table 14 which fit cells D to F that are therefore of interest, as these tasks will potentially provide information on the application of the skill during learning.

Table 14

A creativity assessment comparison grid with examples of measures

Factor assessed		Method of Assessment		
		Self -report	Objective assessment by others	Subjective assessment by others
Process	Potential for use of creative thinking skills	**A** CPES[1]	**B** EPoC[2], TTCT[3]	**C**
	Application of creative thinking skills	**D**	**E** ACER[4]	**F**
Person	Potential for creative behaviours, etc.	**G** Cps[5]	**H** RAT[6]	**I**
	assessed level of creative behaviours, etc.	**J** BICB[7]	**K** ACER[4], ICAA[8]	**L**
Press	Level of support for creativity	**M** SOQ[9]	**N**	**O**
Product	Level of creativity of a product	**P** CDQ[10], CAQ[11]	**Q** CSDS[12], ICAA[8]	**R** CAT[13]

Note. [1]Zhang and Bartol (2010); [2]Lubart et al. (2011); [3]Torrance (1966); [4]Scoular et al. (2020); [5]Gough and Heilbrun (1983); [6]Mednick (1962) [7]Batey (2007); [8]Diedrich et al. (2018); [9]Isaksen et al. (1999); [10]Kaufman et al. (2009); [11]Carson et al. (2005); [12]Cropley and Kaufman (2012); [13]Amabile (1996); For a more comprehensive list of assessment tasks see Appendix C.

It is of little surprise to find that there are few creativity or creative thinking measurement tools identified as being relevant to these categories of tests. None of the reviews of creativity assessments (e.g., Barbot et al., 2019; Cropley, 2000; Henriksen et al., 2015; Kaufman, 2019; Long et al., 2022; Runco, 1999a; Said-Metwaly et al., 2017) described the availability of measures with a focus on measuring skills during learning. When categorising measures from research literature by what they measure, Table 15 shows the categories described and how they relate to the strands of the 4Ps (Clapham, 2011; Henriksen et al., 2015, p. 458; Hocevar & Bachelor, 1989). Although the category of 'ratings by peers, teachers, or supervisors' may potentially include observations of skill application, no measurement tasks describing a suitable process were found in the research literature. Teacher judgement through observations of skill application is often the primary source of information used by primary school teachers to assess their students on creative thinking skills (Machts et al., 2016). Furthermore, although some skills can be observed as a product is being created, others (e.g., cognitive skills such as considering alternatives) may not be observed unless they are documented by the student or creator. Inference, e.g., by relying on "indirect or proxy items" (Wiggins, 1990, p. 2) at best provides information on "the echoes of what we're trying to measure" (Race, 2010, p. 71). Although occasionally used as evidence in the research literature, this not a reliable form of measurement (Brown & Race, 2013; Sambell et al., 2019; Sternberg & Lubart, 1996).

One further form of measure was examined for suitability for measuring critical and creative thinking skills — problem-solving assessment tasks.

3.12. Problem-Solving Assessment Tasks

The forms of measurement which came closest to meeting the requirements for investigating Research Question 2 were creative problem-solving assessment tasks. These tasks include measures of some creative thinking skills and critical thinking skills, as well as other desirable skills such as collaboration and communication skills that were not relevant to this research. Creative problem-solving assessments are characterised by persistence and difficulty in problem formulation, novelty and unconventionality (Newell et al., 1962). It is their unconventional nature which differentiates them from non-creative problem-solving tasks. An advantage of problem-solving tasks is that they incorporate both divergent and convergent thinking, and the problem-solving tasks examined (ACER, OECD and VCAA) were all designed for use in schools.

Table 15

Classification of creativity measures

Category of test	Strand of the 4Ps assessed
Divergent thinking tests	process – potential
Attitude and interest inventories	person
Biographical inventories	person
Personality inventories	person – potential
Ratings by peers, teachers, or supervisors	product/ person/ press/ process
Ratings of eminence	product/ person
Judgments of products	product
Self-reported creative activities	person
Environmental climate inventories	press

Three creative problem-solving assessments were investigated. Two were developed in Australia: The Victorian Curriculum and Assessment Authority (VCAA) Digital Assessment Library (Victorian Curriculum and Assessment Authority (VCAA), 2019), and the Australian Council for Educational Research (ACER) 21st century skills assessment (Scoular et al., 2020). The third assessment task was developed by the OECD (Vincent-Lancrin et al., 2019), however this task incorporated none of the detailed descriptions of the skills from the C&CrT continuum, so was not a candidate to answer Research Question 2.

The ACER task was also considered to be unsuitable for a number of reasons. As previously discussed, the ACER definition of creative thinking does not identify the convergent thinking skills recognised for effective creative thinking to occur. The effect of this focus on divergent thinking was that there was insufficient coverage of convergent thinking in their task to provide sufficient coverage of the skills in the C&CrT continuum. Also, their assessment of the divergent skills focussed on how successfully the aspects of creative thinking had been applied, rather than on the level of development of the skill. Marks were given, for example, on how many different ideas were

generated, rather than on the student's approach to generating new ideas, e.g., did they come up with new ideas "by building on what they know" (Australian Curriculum Assessment and Reporting Authority (ACARA), n.d.-b; Imagine possibilities and connect ideas Level 2), or "by combining ideas in a variety of ways and from a range of sources" (Australian Curriculum Assessment and Reporting Authority (ACARA), n.d.-b; Imagine possibilities and connect ideas Level 5). Further, this task was considered unsuitable because ACER developed a simplified thinking skill development framework similar to but not identical to the C&CrT continuum, around which they structured the assessment task (Scoular et al., 2020). The task was therefore not measuring the skills as described in the C&CrT continuum, which was a major requirement of the measure being sought for use in this thesis.

The descriptions of the skills in the VCAA task, although not identical to the C&CrT continuum, were extremely close. The disadvantage of this assessment, which was common across all the problem-solving assessments, was that it was designed around learning modules with assessments embedded within the task. The student therefore would necessarily be required to complete the modules in order to be assessed and the skills would only ever be measured during purpose-designed assessment tasks. Although a student may apply many of the skills they are capable of using in an artificial scenario such as an assessment task, this still only provides an indication of the *potential* for a student to apply the skills in a genuine creative event. The VCAA assessment therefore measures the use of skills during the assessment task, rather than the use of skills applied in the learning environment.

3.13. Choice of Assessment Task to Measure Critical and Creative Thinking

Having examined the creativity, creative thinking and creative problem-solving assessment tasks, it became clear that the measurement tool I needed to investigate my research questions did not currently exist. This lack of a suitable form of measurement indicated a clear gap which had to be addressed if my research questions were to be answered, necessitating the development of a suitable task.

Critical thinking and creative thinking, while sharing some components of convergent thinking skills, are distinctly different. In clarifying the relationship between them, it became clear that although there is overlap in their use of convergent thinking skills, a measure of creative thinking would be unlikely to effectively fully measure critical thinking. In answer to Research Question 2, measures of critical thinking and creative thinking would need to include both convergent and divergent thinking skills. No suitable measures were found that were suitable for measuring either critical thinking skills or creative thinking skills during learning in an Australian

primary school. I therefore developed a measure for the purpose. Research Question 2, presented in Table 16, was refined to reflect this change in requirement for a new measure to be developed. Research Question 2a, with a focus on what should be measured, will be investigated in Chapter 7 in Part B of this thesis.

3.14. Current Gaps in Knowledge

Gaps in the literature relating to the critical and creative thinking skills and their intersection with STEM Problem-Based Learning with primary school students, are found particularly in four areas. The first area is in descriptions of the stages of development of critical and creative thinking skills used by students, at all levels of schooling. Without fully understanding the stages in the development and use of the skills, it is difficult to effectively teach and assess the skills of interest.

The second area is whether the continuum effectively represents how the critical and creative thinking skills develop and are applied during learning. As the Australian Curriculum provides guidance on what should be taught and expected from students at specific year levels, it is important that the information is accurate to support teachers effectively in their planning and teaching. This, in turn, enables teachers to support student learning effectively. The development of an accurate curriculum depends however on the level of understanding of the stages of development of the critical and creative thinking skills used by students.

The third area in which there are significant gaps in the research is in the methods of measuring critical and creative thinking skills used by primary school students during learning activities. Although some of the critical thinking skills and some of the creative thinking skills may be demonstrated visibly by students (e.g. asking questions and applying divergent thinking techniques), other skills are solely cognitive and difficult to measure through standard forms of assessments. Effective measures are required if the skills are to be accurately measured for effective research.

The fourth area in which there are significant gaps in research literature, is on the effect of STEM PBL on primary school student critical and creative thinking skills. Without a clear understanding of the stages of development of critical and creative thinking skills, and appropriate measures to identify their use, it is not possible to know whether specific methods of teaching have any effect on their use and development.

Table 16

Research questions (1st revision)

Overarching Question:	What is the impact of an industry partner in STEM PBL on the **application** of critical and creative thinking skills by primary school students?
Research Question 1	What is the theoretical basis for the Australian Curriculum Critical and Creative Thinking learning continuum?
Research Question 2	How are critical and creative thinking skills measured during learning?
Research Question 2a	**What should be measured to assess critical and creative thinking in the classroom during learning activities?**
Research Question 3	How do the characteristics of STEM PBL affect the use of skills during learning?
Research Question 4	How are student motivation and engagement measured in primary school classrooms during learning activities?

Note. Revisions to the original questions are indicated in bold

Chapter 4.

Theoretical Foundations and Measurement of the Effect of Problem-Based Learning

4.1. Introduction

The (now revised) overarching question 'What is the impact of an industry partner in STEM PBL on the application of critical and creative thinking skills in primary school students?' was inspired by my participation the 2015 DfE STEM Evaluation Study. Unanswered questions at the conclusion of the project led to the development of Research Question 3, 'How do the characteristics of STEM PBL affect the use of skills during learning?', with a focus on the use of Problem-Based Learning for teaching students. Chapter 4 begins by investigating the terminology of PBL, and the features which distinguish it from other styles of teaching. Next, it investigates and evaluates the research literature pertaining to a) PBL and the development of critical and creative thinking skills; b) PBL and its effect on motivation and engagement; and c) the use and significance of industry partners and the problem they share with students. Gaps in the literature are identified.

4.2. What is Problem-Based Learning?

The term Problem-Based Learning (PBL) is loosely used to refer to a range of teaching practices. Problem-Based Learning belongs to a pedagogical approach to teaching and learning collectively known as inquiry-based, or experiential learning approaches (Condliffe, 2017; Larmer, 2014; Savery, 2006). Three forms of inquiry-based approach used in South Australia are Inquiry Learning, Problem-Based Learning and Project-Based Learning (Department for Education (DfE) South Australia, n.d.-a, n.d.-b). The relationship between these styles of teaching is illustrated in Figure 9.

For the purpose of this thesis, PBL is defined as described by Newman et al. (2003). Characterised by what is known as the PEPBL criteria, this is based on the structure described by Barrows (1986) and Walton and Matthews (1989). It includes three main characteristics, relevant to its application at any level of education— from primary school to tertiary education (Newman et al., 2003, p. 4):

- The curriculum is organised around problems rather than disciplines, with an emphasis on cognitive skills as well as knowledge.

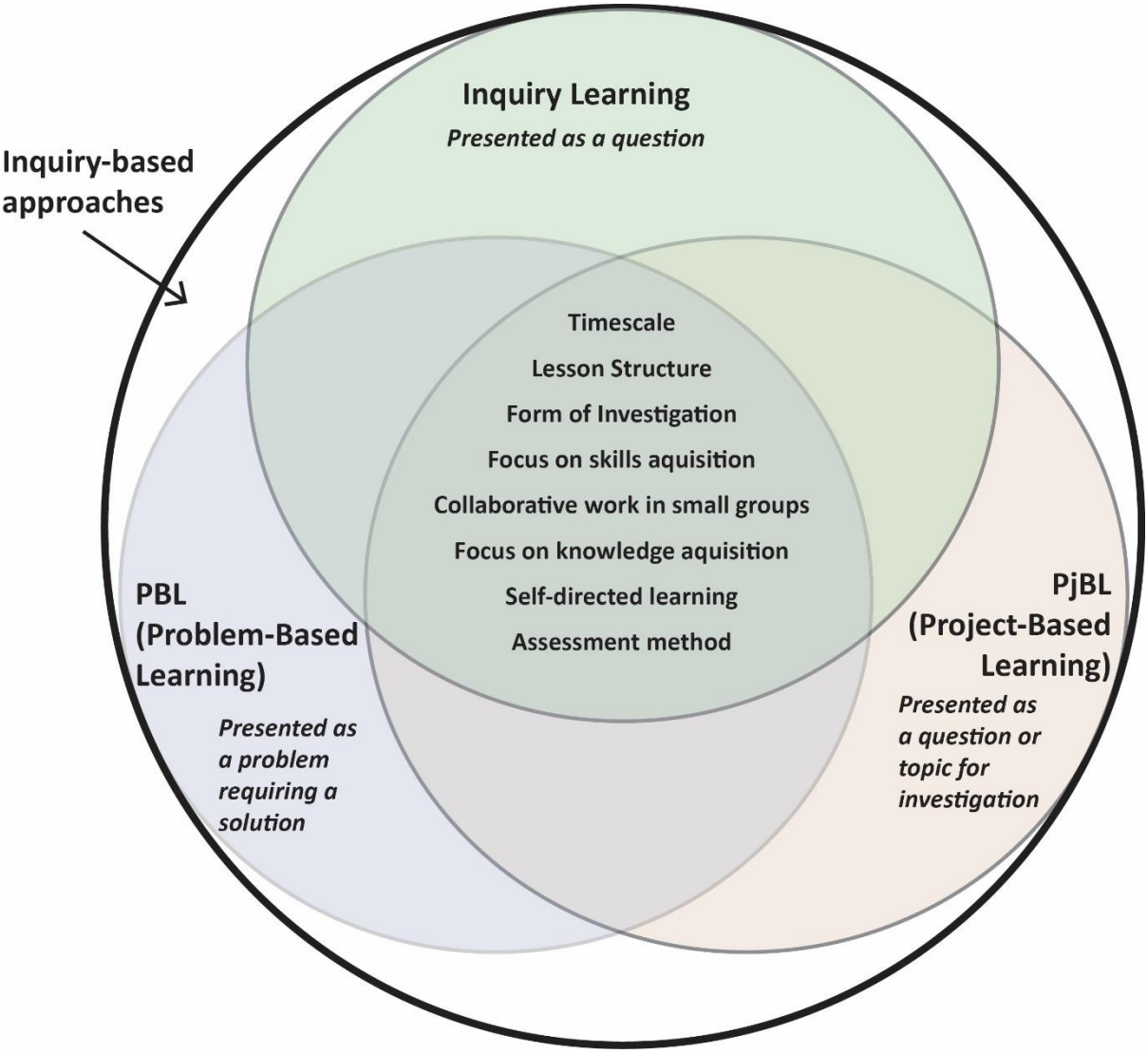


Figure 9

Relationship between three inquiry-based approaches: Inquiry Learning, Problem-Based Learning and Project-Based Learning

- The learning environment uses small groups, active learning, and independent study, and it is student centred. Teachers are facilitators, providing knowledgeable structure for the learners.
- Outcomes focus on skills development and motivation, as well as abilities, for lifelong learning.

This definition is informed by my review of PBL literature which has revealed a diverse and confusing array of definitions, terminology and understandings of PBL and related inquiry-based approaches.

Three overlapping issues contribute confusion about these styles of teaching. These issues are: similarities in the terms used to denote different styles of teaching; similarities between the characteristics of these styles of teaching; and inconsistent use of the names – particularly Problem-Based Learning and Project-Based Learning. Both Problem-Based Learning and Project-Based Learning share the acronym PBL compounding the confusion. Condliffe (2017), Albanese and Mitchell (1993) and Jensen (2015) all refer to both Problem-Based Learning and Project-Based Learning as PBL.

In Condliffe's (2017) review of literature on *Project*-Based Learning for example, the acronym PBL is used throughout. PBL is also central to Albanese and Mitchell's (1993) review of literature on *Problem*-Based Learning. Jensen (2015) refers to them both as PBL (p. 4). Elsewhere, Project-Based Learning is known as PjBL (Kimberlin Education, 2018), and in the interests of clarity, this approach will be used in this thesis. Figure 9 summarises the key features of the three inquiry-based approaches – Problem-Based Learning (PBL), Project Based Learning (PjBL) and Inquiry Learning. The characteristics that most clearly define the three approaches will be described in the following section, as a reference for the subsequent analysis of research on the various forms of inquiry-based learning.

In their purest forms, Inquiry Learning begins with a question – what Kath Murdoch describes as a 'compelling question' (Murdoch, 2015, 2019), and ends with a reflection on the learning and a 'call to action'; PBL begins with a complex, real-world problem and ends with a proposed solution (Hmelo-Silver, 2004; Jonassen & Hung, 2008); and PjBL begins with a task, action or topic, which is designed with a focus on a final product (Condliffe, 2017; Murdoch, 2015; Savery, 2006). Table 17 illustrates these characteristics.

Although described 'in their purest forms', the variations in detail possible across these approaches to teaching and learning, as showed by the shared characteristics in Figure 9, create many possible forms in the classroom. Inconsistencies in the use of the names arise where, for example, the teaching style is described as Problem-Based Learning, but instead of a problem the students are presented with a topic to research with the requirement of a poster presentation at the end, or where the teaching style is reported as Inquiry Learning, but the students are presented with a simple question and directed towards a product to demonstrate their learning at the end.

Table 17

Characteristics defining Inquiry Learning, Problem-Based Learning and Project-Based Learning

Form of inquiry-based learning	Initial focus	Desired endpoint
Inquiry Learning	Compelling Question	reflection on the learning and a 'call to action'
Problem-Based Learning (PBL)	complex, real-world problem	Problem solution
Project-Based Learning (PjBL)	task, action or topic	final product

There is also a set of contextual variables that are common across the forms of inquiry-based learning (Barrows, 1996; Duke et al., 2017). These common contextual variables include such things as the timescale of the investigation, the amount of focus on skills or knowledge acquisition, whether self-directed learning is enabled, and how the process is assessed. Some of the most obvious variables have been included in the overlapping area of the circles in Figure 9 to illustrate the commonalities in these three forms of Inquiry-Based Learning. There can however be significant variations in the emphasis on the various shared characteristics (Newman et al., 2003; Smith et al., 2022; Zumbach et al., 2004). A teacher may choose, for example:

- Timescale: to complete the process in one lesson, or over a longer period (e.g., a term),
- Lesson structure: to overlay a structure such as the Engineering Design process to provide rigour, or may allow the students to determine the ebb and flow of the lessons,
- Assessment: to assess the students on the skills they use, or on the number of facts they can report/ to allow them to determine their own way of demonstrating their learning, or to be constrained by a strictly defined endpoint,
- Form of Investigation: to support the students to develop their own research question, or provide the students with a question designed by the teacher.

This is not by any means a definitive list of the characteristics that can be varied but serves to illustrate the difficulties in distinguishing between the three styles of teaching by name alone. Researchers sometimes provide sufficient information to enable the significant characteristics of the form of inquiry-based approaches to be distinguished, however this is rare. Most researchers simply specify the name – either PBL, PjBL or Inquiry Learning. Duke et al. (2017) are an exception to this,

describing their decision process on justifying their investigations as Project-Based rather than Problem-Based by comparing the various characteristics of the teaching process (Duke et al., 2017). They determined that the approach could more accurately be described as Project-Based Learning as the projects did not all focus on a problem and were instead developed around a public event (Duke et al., 2017). They also state (erroneously I believe) that the sustained nature of their inquiry disqualified them from using the term Problem-Based for their teaching style.

In summary, the differences in the application of Problem-Based Learning (PBL) and Project-Based Learning (PjBL) in the classroom depend greatly on the choices made by the teacher. In some classrooms what is described as PjBL may have more in common with the description of 'pure' PBL, and the reverse may also be seen, where a set of lessons described as PBL is closer to the description of a 'pure' PjBL teaching approach.

This lack of clarity in terminology has been a limitation of much of the research on inquiry-based styles of teaching, and of the way in which that research is reported. Because each of the characteristics of the teaching style can be varied, this enables the characteristics to behave as variables which will potentially affect the results of the research reported (Barrows, 2000). As a result, the reliability of comparisons made between studies on these styles of teaching can be affected (Hung, 2011; Thomas, 2000), and research often produces what appear to be contradictory results (Hung, 2011).

It is therefore important when comparing the results of research on forms of inquiry-based approaches, e.g., PBL, to ensure that the main characteristics of the teaching style are similar enough to produce comparable results. It is also important to specify the main characteristics when reporting on research. Table 18 provides a resource to enable comparisons to be made between the reported characteristics of inquiry-based learning styles. The PEPBL criteria, described previously, provides clarity in its description of PBL, and will be the form of PBL used in this thesis.

PBL rather than PjBL is the teaching approach used in the research described in this thesis, for two reasons. Firstly, because it was the approach being used during the initial study prior to the commencement of this research. The observations made regarding the effect on student skills arose from teaching through PBL, so it was logical to make as few changes as possible, to determine

Table 18 (over)

Comparison between features of inquiry-based approaches

Variables	Inquiry Learning	PBL	PjBL	Details/Variations
Essential characteristics				
Initial focus	'Compelling Question'	complex, real-world problem	task, action or topic	
Desired endpoint	reflection on the learning and a 'call to action'	problem solution	final product	
Variable characteristics				
Timescale	varied	varied	sustained inquiry	can vary from a single lesson to a term or beyond
Lesson structure	clear structure, e.g., notice, think, predict, investigate, review, communicate	clear structure, e.g., define, identify, investigate, generate, critique, re-design, share	varied	overlay a structure such as the Engineering Design process to provide rigour, or may allow the students to determine the ebb and flow of the lessons
Form of investigation	students aided to develop research questions based on 'wonderings	students develop their own research questions, driven by the problem	direction of research determined by teacher	students to develop their own research question or provide the students with a question designed by the teacher.
Focus on skills acquisition	strong focus	strong focus, enabled by the problem solving process	weak focus	
Focus on knowledge acquisition	moderate focus	knowledge acquisition is prompted by the problem solving process	strong focus	
Role of the teacher	facilitator	facilitator	varied	
Collaborative work in small groups	strong focus	strong focus	varied	
Self-directed learning	strong focus	strong focus	varied	less evidence of self-directed learning where learning is more structured by teacher
Assessment method	varied	students choose the most appropriate way to present their solution (model, report, presentation, etc)	format of final product specified (poster, model, etc) and assessed against a rubric	assessment may include self-report or other format of assessment to measure skills and/or knowledge acquisition.
Final outcome of investigation	varied	varied	public presentation of final product	

whether the observed changes are measurable. Secondly, PBL provides more opportunities for student choice in the final presentation of the solution, potentially providing more opportunities for students to apply a range of skills (Jonassen 1997).

To investigate Research Question 3: 'How do the characteristics of STEM PBL affect the use of skills during learning?', literature pertaining to the results of PBL studies with primary school students was sought. Much of the research on PBL has focussed on tertiary education, with a strong emphasis on medical applications, as this was where PBL was first actively used in education (Albanese & Mitchell, 1993; Berkson, 1993; Vernon & Blake, 1993). In comparison, research on PBL across various facets of non-medical education is criticized for lacking in rigor and quantity (Hung, 2011; Savery, 2006; Strobel & Van Barneveld, 2009; Wilder, 2015). This is illustrated by the growing number of research articles which purport to report on the positive effects of PBL on Critical Thinking, but contain no detail of the teaching or assessment methods, e.g., Setyowati (2022) and Fajari et al. (2020).

In addition to concerns about the quality of the research, concerns have been expressed about the application of the body of research relating to PBL in medical fields, to non-medical fields (Hmelo-Silver, 2004). Mergendoller et al. (2006) warn against applying the results of studies on tertiary students to secondary students. They cite differences between the populations (tertiary and secondary students) such as their level of verbal and quantitative skills, intellectual development, experience and accomplishment using hypothetical-deductive reasoning, and motivation to engage in the learning. Generalizing the results of studies on tertiary and secondary students even further to primary students would invite the same comparisons. As Fredericks and McColskey (2012; also Kuhn, 1999) point out, there are also potential issues when developmental differences are not taken into account. Over a decade ago, Drake and Long (2009) and Pease and Kuhn (2011) commented on the relatively few examples of research on PBL in primary school, and little has changed in the intervening years. This gap in the literature is stretched even further when looking for research specifically relating to the effect of PBL on primary school students' skills. The majority of PBL research focusses on how PBL affects the acquisition of knowledge rather than of skills; e.g., a meta-analysis by Walker and Leary (2009), a pilot study by Drake and Long (2009), and a study by Gallagher and Gallagher (2013) on developing academic potential in primary school students. Added to the general lack of research on the effect of PBL on skill acquisition or application, is the lack of clarity around the descriptions of the skills being assessed, which makes meaningful comparisons between research results extremely difficult. As we saw in Chapter 2, this lack of clarity extends even into the Australian Curriculum.

4.3. PBL and Critical and Creative Thinking

The search for relevant literature pertaining to the impact of PBL on student skills commenced by looking for research which examined the use of the skills specified in the C&CrT continuum. Although no PBL-focussed research has been found evaluating the development of the specific skills identified in the continuum, studies have been carried out measuring more general skills (e.g., critical thinking without specifying which components of critical thinking were the focus) developed through PBL teaching.

An extensive search (see Appendix E for the Search Outline) for studies examining the effect of PBL on skills in primary school students found only three relevant studies. Studies which did not sufficiently describe the process of PBL, or described a form of PBL which varied significantly from the PEPBL criteria were not included. Studies claiming to measure skills were examined but were only included where critical thinking or creative thinking skills which shared characteristics of the skills described in the Critical and Creative Thinking learning continuum were included.

In the first of two studies from the USA, Shepherd (Shepherd, 1999) implemented a PBL intervention with fourth and fifth grade students *N* = 35). The study claimed that PBL can have a positive effect on the development of students' critical thinking skills. Shepherd describes a PBL model of teaching which he called PROBE – a hybrid Project-based and Problem-based teaching approach. Although PROBE shares only limited similarities to the PEPBL criteria, his focus on the development of student skills is highly relevant to our study. A significant increase in a number of specific skills (deduction, credibility, induction, and assumption identification), measured with the Cornell Critical Thinking Test (CCTT) was reported post-intervention in comparison to the control group. Shepherd also reported on an increase in student motivation and engagement, identified through qualitative student and teacher interviews.

Drake and Long (2009) describe the implementation of a PBL study (based on the PEPBL criteria) on electricity with 4th grade elementary students *N* = 33). Specific problem-solving skills (e.g., 'asking questions' and 'make a plan') were assessed through a problem-solving activity and were found to have increased. Motivation and engagement, measured by time-on-task and reduction of inappropriate-non-productive behaviours, were also found to have increased.

In a study from Taiwan[14], P-H Hung et al. 2014, analysed the questioning ability of fifth and sixth grade students (*N* = 43). The questioning skills assessed included: exploring environments,

[14] Assumed, not specified.

understanding contexts, organizing information, and proposing a valuable and answerable driving question. Although their definition of Questioning differs from that specified in the Australian Curriculum, some overlap of skills is evident, making this research relevant to this study. An intervention based on the particular PBL approach of ubiquitous problem-based learning system (UPBLS) was credited for the increase in skills, although the lack of a control group in the study makes this is difficult to prove. The authors state that measured increases in the students' skills are evidence of the effectiveness of the UPBLS approach in helping develop the ability to question.

As already discussed, results of studies on tertiary students are considered too different to be relatable. Studies on secondary students also cannot be assumed to be applicable in a primary school context, however examining research on secondary students provides an example of the difficulties faced with making comparisons between different studies in the area of PBL and skills, so are worth examining briefly.

Studies on secondary students provide contradictory results. As an example of the contradictions, Burris and Garton (Burris, 2005; Burris & Garton, 2007), found no increase in the critical thinking skills of students in a PBL treatment group. Students in their control group demonstrated a measurable positive change in skills. This could be taken to indicate that PBL has had a deleterious effect on the development of critical thinking in these students. Specht (2015), and Lapuz and Fulgencio (2020) on the other hand, report a significant increase in critical thinking skills in their students through PBL. One of the difficulties in unravelling the contradictions in these studies is that none of the studies provide sufficient details of the skills being measured. By using the term 'critical thinking skills', the presence of multiple skills is inferred. The use of the term 'critical thinking' may refer just to argumentation. But without clarification, insufficient detail is provided to ensure that the skills being compared are in fact the same in all three of these studies.

Not all studies relating to problem-solving are relevant in the current context. Permatasari (2020) describes PBL in maths with a focus on problem-solving. This might be more accurately described as a maths-problem investigation as it shared few of the characteristics of PBL.

No articles were found on the effect of PBL on creative thinking. This is a gap in the literature.

Shepherd (1999), Drake and Long (2009) and Hung et al. (2014) are the only studies found which provide sufficient detail of the skills measured and PBL methods used, to inform the relevance of the research question and guide the formulation of a hypothesis in response to Research

Question 3: ‘How do the characteristics of STEM PBL affect the use of skills during learning?’ Based on these studies, Research Question 3 was revised to the following two questions:

- Research Question 3a: ‘What is the impact of PBL on student critical and creative thinking skill application?’
- Research Question 3b: ‘What is the impact of using an industry partner problem in STEM PBL on student critical and creative thinking skill application?’

Tentative hypotheses for Research Questions 3a and 3b are proposed based on the limited research described above, that:

- Hypothesis H3a: using PBL as a teaching approach with primary school students will enable them to use a greater variety of critical and creative thinking skills, and at higher ‘levels’ at the end of a 10-week PBL intervention than when a more passive method of teaching is used.
- Hypothesis H3b: Students engaged in STEM PBL with an industry partner will show use of a greater variety of critical and creative thinking skills at higher ‘levels’ than students engaged in a Theory PBL problem.

This investigation of the literature also exposed the research regarding motivation and engagement, and PBL. This led to the development of Research Question 4: ‘How are student motivation and engagement measured in primary school classrooms during learning activities?’, and two sub-questions:

- Research Question 4a: ‘What is the impact of PBL on student motivation and engagement?’
- Research Question 4b: ‘What is the impact of using an industry partner problem in STEM PBL on student motivation and engagement?’

To formulate a hypothesis for these questions the literature on PBL and motivation and engagement must be further investigated.

4.4. PBL and Motivation and Engagement

There are numerous claims from both PBL research, and from motivation and engagement research, that inquiry-based learning styles have a positive effect on motivation and engagement (e.g., Belland et al., 2013; Dochy et al., 2003; Hmelo-Silver, 2004; Jerzembek & Murphy, 2013; Specht, 2015; Stepien & Gallagher, 1993). As with other areas of research however, the evidence of research with primary school students is limited. The lack of description of the methods used to measure

motivation and engagement also makes analysis of research reports extremely difficult to carry out. As motivation and engagement are two extensively researched and described attributes, this section of the Chapter 4 focusses very explicitly on the intersection of PBL and motivation and engagement in primary school students. Chapter 5 examines motivation and engagement in the broader context.

Research reports on motivation and engagement of primary school students engaged in PBL were disappointingly few in number, revealing another gap in the literature. This finding is supported by Hmelo-Silver (2004), who notes that there is little actual research to support the claims of increased intrinsic motivation and engagement in students through PBL.

In some cases research reported on the results of teaching approaches which shared characteristics in common with PBL (e.g., hands on activities, active learning classrooms and highly engaging learning activities), and their effect (if any) on motivation and engagement. Anderson (1998) reports on the increase in comprehension and conceptual knowledge in primary school students who were more intrinsically motivated with their hands-on activities compared with a control group. Rotgans and Schmidt (2011) confirm Schraw and Lehman's (Schraw & Lehman, 2001) observations that an active-learning classroom maintains students' engagement more effectively. Saeed and Zyngier (2012) confirm that primary school students who were engaged in group work, provided with choice in their learning, and engaged in interesting learning activities (characteristics of PBL teaching – and other approaches to teaching) were more motivated and engaged in their learning. Finn and Zimmer (2012) report that where student interactions encourage discussion, and support student expression through dialogue, student engagement is facilitated. These results align with Christenson's assertion (2012b) that motivation and engagement do not happen in isolation.

> The role of context cannot be ignored. Engagement is not conceptualized as an attribute of the student but rather as an alterable state of being that is highly influenced by the capacity of school, family, and peers to provide consistent expectations and supports for learning (p. v).

Classroom learning opportunities and interactions with the teacher provide opportunities for students to choose to engage with learning – "In short, both the individual and context matter." (Christenson et al., 2012b, p. vi; see also: Martin, 2012; Reschly & Christenson, 2006a, 2006b; Thomas, 2000). Despite these generalised findings in relation to motivation and engagement, research on motivation and engagement specifically referencing PBL and primary school students was conspicuously absent.

The lack of existing research makes it of particular interest to test whether providing learning through a PBL approach has a measurable effect on motivation and engagement. As the

indications are that PBL has a positive effect on critical and creative thinking, understanding the impact of with motivation and engagement may help us to better understand how to help our students develop their critical and creative thinking skills more effectively.

On the basis of the literature review findings, an additional hypothesis was tested during the empirical phase of the research that, using PBL as a teaching approach, primary school students will increase their motivation and engagement. (Research Question 4a, reported in Chapter 9)

4.5. The Problem in Problem-Based Learning

The final feature of PBL of interest in this literature review is the problem, or topic, presented to the students. As indicated in Section 4.2, this is one of the variable characteristics of inquiry-based learning styles.

Advocates of the different forms of inquiry-based teaching argue that the initial introduction (i.e., the format through which the students are introduced to the 'investigation': as a question to be answered, a problem to be solved, or a topic to be researched), has an important effect on the way the students react to or engage with the investigation (Jonassen & Hung, 2008). As indicated in Table 18, PBL uses a problem format to present the investigation to the students. The type of problem most frequently recommended for PBL is a real-world, or 'authentic' problem (see: Hung, 2016; Jonassen & Hung, 2008; Rossi et al., 2019; Stepien & Gallagher, 1993)[15]. Real-world contexts are promoted as encouraging motivation and engagement (Maestrales et al., 2022; Newmann, 1992; Rotgans & Schmidt, 2011), which are, in turn, believed to increase development of both knowledge and skills (Fredricks et al., 2004; Lei et al., 2018). Although the previous works cited support the use of real-world or authentic problems, no direct comparison of problem types had been attempted in PBL research prior to research by Jonassen and Hung in 2008 (Jonassen & Hung, 2008, p. 22). Comparisons of problem types appears to be an under-researched area of PBL (Ames, 1992; Hung, 2016), as no comparisons between problem types in PBL were forthcoming in the literature search which formed the basis of this literature review, either. This is another gap in the literature worthy of note.

Hung et al. (2016; 2013) researched the features of real-world problems that successfully engage students. Their research is on tertiary students and has not been repeated with primary

[15] An alternative to an 'authentic' problem would be, for example, a hypothetical problem generated by a teacher or one obtained from 'problem-resource materials', e.g., Jolly, A. (2016). *STEM by Design: Strategies and Activities for Grades 4-8*. Routledge. ; or Lioncrest. (2023). *Macmillan STEM Investigations*. https://lioncrest.com.au/our-range/stem-science/macmillan-stem-investigations-detail

school students, where there is a gap in the research literature. Hung et al. identified six elements relating to the perceived authenticity of problems. The six elements are: that the problem should be an unsolved real-life problem; it needs to have time proximity, i.e., it needs to be relevant at the time the students are working on it; it needs to have location proximity, i.e., it should be relevant to the students because it impacts on them (this need not exclude world problems, if the students believe that they will be personally impacted by the problem or the solution); the presence of the subject makes a difference, i.e., the person with the problem needs to be in the room with the students; the problem should be of personal interest; and finally, (this is likely to be of most relevance to older students) it should also be of career interest (Hung, 2016, p. 7).

Although no further investigations specifically analysing PBL problem-types were found in the research literature, two studies of interest were discovered. The first is a small study *N* = 12) comparing the effect of ill-structured and well-structured learning tasks on study tactics and critical thinking with elementary school students (11 and 12 year olds), in a self-regulated learning study (Malmberg et al., 2014). This study is of interest because one of the commonly described characteristics of effective problems in PBL is their apparently ill-structured character (Hmelo-Silver, 2004; Jonassen, 2000; Savery, 2015). Ill-structured in this context refers both to the possibility of multiple 'correct' solutions, and to the inclusion of potentially irrelevant information. The similarities between the study and PBL learning however ended there. Neither task was open-ended, and all support materials considered relevant to the problems were provided to the students, i.e., this was not student-directed learning. The findings of relevance to this thesis, were that students were reported as choosing to apply critical thinking skills only in the ill-structured task (reasoning, thinking and questioning), and not in the well-structured task. This would therefore suggest that the use of ill-structured tasks may be beneficial to the use of student skills.

The second study of interest examined task characteristics and engagement with elementary students (*N* = 1,348). Marks (2000) identified a positive relationship between engagement of students, and their belief that they are engaged in 'authentic instruction' (i.e., work connected to the world beyond the classroom). Although not specifically a study of authentic problems in PBL, this study does support Hung, Jonassen and their colleagues' assertions that authenticity matters.

Pianta et al. (2012) report on the positive effects on engagement of students involved in real-life and real-world experiences that make meaning in their lives. They support the findings of Hung et al. with the point that "the real-world connections must be made in ways that are meaningful as perceived by the student." (Pianta et al., 2012, p. 371; also Hofkens & Pianta, 2022). Again, Pianta et al. support the suggestions that authenticity matters, but they go further, suggesting

that no matter how genuine a problem is, it will only provide the results of an effectively 'authentic' problem when the student perceives it as so.

Hung's comments on student's ownership of problems supports this, as he says:

> …we found that many students developed a strong sense of ownership of the problem…we found that they took the problem into their own hands and hearts, not because of the problem but because of the person in the problem and where the problem occurred (Hung, 2016, p. 7).

The effectiveness of 'the person in the problem', or the real live 'stranger in the room' who appeals to students with 'I have a problem – can you help me?', is potentially the link between a genuine problem, and a problem that the students identify as authentic. This also suggests that the results of student learning may be different when a problem is presented by an industry partner, than when it is a hypothetical problem either developed by the teacher or sourced from commercial problem-resource materials. (Research Question 3b, reported in Chapter 9)

4.6. The Impact of Industry Partners in PBL

The role of an industry partner in a PBL environment is to provide the real-world context for students to investigate. This is achieved by the industry partner providing a genuine, unsolved problem – preferably one they are currently working on, for the students to investigate. This provides authenticity, which is, as previously discussed, one of the essential characteristics of PBL problems (Hung, 2016)

The promotion of industry partners to increase motivation and engagement in students in Australian schools comes from Education sources such as the South Australian Department for Education (n.d.-c; also Rossi et al., 2019) and the Queensland University of Technology (Watters et al., 2016). Research on the effect of the use of industry partners on student learning is however limited. The only research report found was produced by the Australian Council for Educational Research (Lonsdale, 2011). The report provides feedback from a study of industry partnerships with primary and secondary schools in Australia, gathered from award applications submitted by the schools[16]. 801 applications, each representing a partnership between Industry and a school, were analysed. Difficulties in the analysis were caused by the variations in reporting methods by the schools, and the variations in the forms of the partnerships in the schools and classrooms. These differences make it difficult to draw comparisons between the results from the participating schools. Schools provided quantitative and/or qualitative data, which was collected before and after the

[16] Schools First Awards program

partnership occurred. Benefits such as improved student attendance, quality of work and enhanced critical and analytical skills were reported, however almost 70% of the evidence used to identify the outcomes was anecdotal. As Lonsdale explains, “These may well be legitimate outcomes but it was not clear how these conclusions were reached or what data were collected to show improvement.” (Lonsdale, 2011, p. 17).

4.7. Current Gaps in Knowledge

The gaps in the current research literature relating to Problem-Based Learning are numerous – particularly where primary school students are concerned. Specifically, gaps were found in the following areas:

- All research with a focus of its application on primary school students,
- PBL research in which the characteristics of the PBL method used are clearly described,
- The effect of PBL on critical and creative thinking skills,
- The effect of PBL on motivation and engagement,
- The effect of problem type on motivation and engagement,
- The effect of problem type on critical and creative thinking skills,
- The effect of an industry partner in PBL.
- In addition, where research was conducted in primary schools it tended to lack rigour. The terms of investigations tended to lack precision, and control groups were not always used.

The lack of research literature investigating primary school students in all of these areas is particularly significant as Problem-Based Learning continues to be promoted as an effective pedagogical approach for teaching students. The paucity of literature on the use of PBL and industry partners on the development of critical and creative thinking skills, and on their effect on motivation and engagement in Primary school students, are gaps in the literature worthy of research. As identified in Chapter 3, no suitable assessment tasks exist to measure the critical and creative thinking skills of interest. It will therefore be necessary to develop a task for this purpose. The question now arises of whether a suitable assessment task exists to measure motivation and engagement in Australian primary school students. This will be investigated in the following chapter.

Chapter 5.

Theoretical Foundations and Measurement of Motivation and Engagement

5.1. Introduction

Motivation and engagement initially became a topic of interest in this study due to claims that the use of Problem-Based Learning (PBL) increases motivation and engagement in students (e.g., Blumenfeld et al., 1991; Hmelo-Silver, 2004; Norman & Schmidt, 1992). Motivation and engagement were also characterised by multiple commentators as dispositions which support effective critical and creative thinking, (Amabile, 2012; Csikszentmihalyi, 1996; Facione, 1990). If both these statements are true — that PBL increases motivation and engagement and that motivation and engagement increase critical and creative thinking, then motivation and engagement may be an important link between the effectiveness of PBL and the development of students' critical thinking and creative thinking skills. The question being investigated in this chapter, Research Question 4, is therefore: 'How are student motivation and engagement measured in primary school classrooms during learning activities?'

Theories of student motivation and engagement are complex and intertwined (Skinner & Raine, 2022). Understanding what is meant by the terms motivation and engagement, and which theories may be relevant to research on students engaged in STEM PBL, is necessary before their effects on the development of skills, by primary school students, can be investigated.

The terms motivation and engagement, and the relationship between them, are subject to debate (Finn & Zimmer, 2012; Reschly & Christenson, 2012; Skinner & Raine, 2022). The debates include whether motivation and engagement are separate constructs, and if so which components belong to which construct; whether motivation is a necessary condition for engagement, or vice versa (Blumenfeld et al., 2006; Mahatmya et al., 2012; Mitchell, 1982). There are clear areas of overlap and interaction between the two concepts, as described by Martin in his analysis of various theories and the constructs they identify (Martin, 1999-2022). Martin's research (e.g., Martin, 2007; Martin, 2008b; also Martin et al., 2021; Martin et al., 2017; Martin et al., 2015) has focussed on the relationship between the concepts of motivation and engagement rather than their differences. His recent work (Martin et al., 2017) appears to demonstrate a cyclical relationship in which motivation is both influenced by engagement and provides an impetus for subsequent engagement. Skinner and

Raine (2022) support this view of the intricate relationship between motivation and engagement with the suggestion that motivation and engagement offer complimentary perspectives but also that important overlap exists between them. Martin's research also indicates the presence of both positive and negative aspects of motivation and engagement, supporting Wang and Salmela-Aro's view of separate continua for these dimensions (Salmela-Aro et al., 2022; Wang et al., 2019).

It is important for researchers to be aware of the complexity of some of the issues which continue to be debated, particularly with regard to the structure and components of both motivation and engagement which have an impact on the choice of assessment method. An awareness of the debates also helps explain the varieties of definitions, and some of the anomalies in the research literature.

5.2. What is Motivation?

Motivation is widely regarded as having a crucial role in education (e.g., Deakin Crick, 2012; Richardson et al., 2012; Taylor et al., 2014). At its most simple characterization, motivation is 'wanting' (Baumeister, 2016, p. 1). It is "to be moved to do something" (Ryan & Deci, 2000, p. 54) — as distinct from a lack of motivation where there is no impetus or desire to do anything. (Ryan & Deci, 2000, p. 54)

Early theories of human motivation, although not specifically focussed on students and learning, contribute to an understanding of what motivates us at the level of our most basic needs. Freud's Life and Death Instincts (Zilbersheid, 2013) and the Humanistic theories such as Maslow's Hierarchy of Needs (Maslow, 1943), and the Arousal theory (Mitchell, 1982), as described below, are examples of this. The Arousal theory proposes that people have a preferred level of arousal[17] which motivates them to engage in behaviours to maintain an optimum arousal level—whether engaging in risk taking behaviours to increase arousal, or in calming behaviours such as reading, to reduce arousal. Certainly, some student behaviours such as those identified as classroom attention-seeking behaviours could be explained by this theory. The Hierarchy of Needs theory (Maslow, 1943) suggests that only when essential physiological and safety needs are satisfied can needs further up the hierarchy motivate a person (Lester et al., 1983). Interventions such as school breakfast programs draw on this theory, with the justification that if students haven't met their basic need for food before class begins, they will be unmotivated to learn.

[17] The term arousal here refers to a level of stimulation a person might seek. The type of stimulation might take any form, e.g. physical, emotional, mental.

While any of these general theories might partially explain some aspects of student motivation more specific theories have been developed which provide deeper insights into students' motivation to learn. Freud's and Maslow's theories don't explain a student's motivation to achieve an 'A' grade, when a 'B' grade is sufficient for a pass. The theories discussed so far describe motivation arising from a perceived need: they are needs-based models. More recent theories have been developed which describe motivation as directly influenced by outcomes rather than need (Mitchell, 1982); or that focus on specific psychological processes (Locke & Latham, 2004). Expectancy theory, operant conditioning and equity theory are examples of outcomes influenced motivation. Equity theory focusses on people's motivation by a desire for fairness (e.g. Carrell & Dittrich, 1978)); expectancy theory and operant conditioning on how the promise of rewards provides motivation (e.g. Davis & Luthans, 1980; Schwab et al., 1979). As explained below, some of these theories have a more direct relationship with education.

Self-Determination Theory (SDT: Ryan & Deci, 2000) is one of the more recent theories of motivation which provide clearer insights into how motivation influences student outcomes. STD is an example of psychological-process theory, in which the satisfaction of different psychological needs is postulated to be responsible for people to "experience ongoing growth, integrity and wellness."(Ryan & Deci, 2022). In SDT a clear distinction is made between extrinsic and intrinsic motivation. Extrinsic motivation is characterised as taking action in response to external incentives, goals or pressures (Foster & Schleicher, 2022, p. 11). Intrinsic motivation occurs when actions are perceived as inherently enjoyable; in Skinner's proposition the activity is its own reward (Ryan & Deci, 2000, p. 56). Intrinsic motivation is described as resulting in high-quality learning and creativity (Ryan & Deci, 2000; Saeed & Zyngier, 2012). Although extrinsic motivation has been typically described as a less effective form of motivation, and even as undermining intrinsic motivation e.g., when extrinsic incentives overwhelm intrinsic interest (Lepper et al., 1997, p. 26), SDT demonstrates how extrinsic motivation can be effectively utilised to achieve learning where intrinsic motivation cannot be stimulated (Hennessey, 2010; Ryan & Deci, 2000; Saeed & Zyngier, 2012). An example of the effective use of extrinsic motivation is where students accept the value of achieving a high score in a university entrance examination, and this motivates them to strive for the highest possible score, even though they may not value the learning which must be engaged in to achieve the high score. As illustrated by these examples, both intrinsic and extrinsic motivation may play a significant role in a student's appetite for working with industry partner problems; intrinsic motivation if an emotional connection is made with a plea for help or if the problem presented is in an area which interests the student, or extrinsic if the student is motivated by the desire to develop a 'winning' solution or to achieve a high mark.

Further examples of how different forms of motivation may interact and affect how students learn, are provided by the relationships shown between behaviours described by Expectancy-Value Theory and Self-Efficacy Theory. Expectancy-Value Theory is concerned with the outcomes that are anticipated as the result of an action (Meece et al., 1990), e.g., hard work gives good results. Self-Efficacy on the other hand, reflects a person's belief in whether they have sufficient ability to successfully achieve a given outcome (Bandura, 1981). While Expectancy-Value Theory and Self-Efficacy Theory each provide a specific view of forms of motivation, when viewed together they provide a deeper understanding, (Belland et al., 2013; Meece et al., 1990; Pintrich, 2003). For example, viewed through Expectancy-Value Theory, a student is motivated to prepare for their exam because they equate exam preparation with success in exams. When Self-Efficacy Theory is overlaid on this, the student may either a) have sufficient confidence in their ability to believe that they can achieve, and so remain motivated to study, or b) have insufficient confidence in their ability to believe that they can achieve, so their motivation to study is diminished – not because their expectancy-value belief is low, but because their self-efficacy is low (Schunk & DiBenedetto, 2020). It may therefore be theorised that students with low self-efficacy in STEM subjects may be less motivated by a STEM problem, while high self-efficacy in STEM would motivate a student to engage in a solving a STEM problem. Low expectancy-value may account for some students being unwilling or unable to develop solutions to a problem in PBL if they have not had previous success with developing problem solutions, while high expectancy-value would motivate them to engage with STEM problems as they had a high expectancy of success.

The discussion so far has focussed on a series of different theories of what affects motivation – whether it is stimulated by physical or psychological needs or desires (e.g., to achieve academic success), or determined by personal beliefs. Although any or all of these theories of motivation (and others not described here) can explain some of the behaviours observed in students, the plethora of theories and their focus on individual elements of a complex psychological process makes it difficult to navigate the field (Baumeister, 2016).

Reschly and Christenson (2022) have analysed the research on motivation and engagement and describe it in more holistic terms. Rather than focussing on individual theories they view the research from two perspectives (p. 19):

- A developmental perspective, which draws from developmental theories which have achievement as their primary outcome. This viewpoint explores motivation in the context of personal traits such as e.g., adolescent egocentrism which may cause increased social anxiety and introjection, reducing students' motivation.

- A motivation researcher's perspective, in which motivation is central and inextricable to a larger construct in which engagement and motivation play individual parts, with achievement as one of the many outcomes of interest. This viewpoint enables multiple theories of engagement and motivation to be drawn together to explain a wide range of motivation and engagement behaviours.

Many researchers call for theories on motivation to be integrated. Several of these also promote the integration of theories of engagement (Martin, 2007, 2009; Murphy & Alexander, 2000; Pintrich, 2003; Reschly & Christenson, 2012). Models such as Martin's (2007) Motivation and Engagement Scale, Locke and Latham's (2004) Integrated Model of Work Motivation and Belland and colleagues (2013) work on designing effective scaffolds to support multiple forms of motivation are examples of integrated models.

Martin's description of motivation as "the inclination, energy, emotion, and drive relevant to learning, working effectively, and achieving" (Martin et al., 2017, p. 150) illustrates a motivation researcher's viewpoint rather than a developmental researcher's viewpoint. Martin defines intent: an inclination, an energy or an emotion; and indicates that it should be *combined* with drive. He then specifies the behaviours which will be affected as learning, working effectively and achieving. The emphasis on want – the reasons for behaviour, or the intent, reflects the earlier needs-based models of Freud and Maslow, while the emphasis on behaviours aligns with Pintrich's (2003) emphasis on contextual-cultural features which are relevant to a more integrated theory of motivation.

This review of theories of motivation has identified theories which help explain the motivation of students engaged in STEM PBL. Theories of self-efficacy, expectancy-value, self-determination and others which we have not examined in detail – need achievement, self-worth, attribution, control, motivation orientation, goal and self-regulation – could all play a part in students' motivation during STEM PBL. Integrated theories of motivation therefore provide a useful starting point to develop a hypothesis for the proposed research.

5.3. What is Engagement?

Although motivation is generally described as a precursor to engagement (e.g., Mahatmya et al., 2012; Schunk & DiBenedetto, 2022) or as necessary but not sufficient for engagement (Appleton, 2012; Appleton et al., 2008; Kim et al., 2015), there is disagreement about which factors should be classed as motivation, and which should be classed as engagement (Christenson et al., 2012a). Thus far, researchers have not arrived at an agreed definition (Reschly & Christenson, 2022). For the purposes of this thesis, engagement will be defined as: the *quality* of a student's connection or

involvement in PBL activities (Skinner et al., 2009), while motivation is: the *degree* to which a student wants and chooses to engage with PBL activities (Mitchell, 1982) These definitions align with Martin et al. (2017) for whom, "motivation is defined as the inclination, energy, emotion, and drive relevant to learning, working effectively, and achieving; engagement is defined as the behaviours that reflect this inclination, energy, emotion and drive" (Martin et al., 2017, p. 150).

While motivation has a history dating back to the first half of the 20th Century (Ryan, 2012), theories of engagement were only brought to the fore in the 1980s. While there are general theories of motivation and learning focussed theories of motivation, theories of engagement began with a learning focus. Studies of engagement effectively commenced in the 1980s, when low rates of engagement were proposed as a cause for high rates of drop-out from school and boredom with education (Finn & Zimmer, 2012). Student engagement, (also referred to as school engagement or academic engagement: Fredricks et al., 2004; Reschly & Christenson, 2022), therefore became a focus for research (Appleton et al., 2006; Skinner & Raine, 2022).

Although the initial focus of research on student engagement was to identify the cause of students dropping out of school (Finn & Zimmer, 2012), engagement has subsequently been found to be a strong predictor of academic outcomes (Carter et al., 2012; Finn & Zimmer, 2012; Klem & Connell, 2004; Liem & Martin, 2012), which potentially also affects the acquisition of skills. Engaged students demonstrate behaviours which support classroom learning (Deakin Crick, 2012; Finn & Zimmer, 2012; Martin, 1999-2022) such as:

- More effective study skills.
- A commitment to participate in the process of learning.
- A commitment to achieve mastery of a subject at the highest level of which they are capable.

Research over the past 40 years has resulted in a diverse proliferation of student engagement theories (Skinner & Raine, 2022; Wang et al., 2019). These have been reviewed in depth by Lei et al. (2018), Reschly and Christenson (2022), and Skinner and Raine (2022), following an earlier review by (Fredricks et al., 2004). As observed with the research on motivation, having more theories, although providing a greater breadth of understanding of a concept, doesn't necessarily provide clarity. There is, accordingly, no unified theory of engagement.

In their influential literature review, Fredricks et al., (2004) described the most prevalent understanding of engagement as a multifaceted construct consisting of three distinct dimensions: behavioural engagement, cognitive engagement, and emotional engagement (re-labelled social-

emotional by Martin (2022). Although there is general consensus between researchers that engagement is multidimensional, the dimensions described vary in number and definition (Christenson et al., 2012b). Appleton et al (2006) replace emotional with academic and psychological engagement. Martin (2022) re-labels social engagement, as social-emotional engagement, Reeve and Jang (2022) add agentic engagement, and Carter et al., (2012) describe engagement as consisting of four subtypes: behavioural, cognitive, academic and affective (Finn & Zimmer, 2012).

Also contributing to the lack of agreement on a unified theory of engagement is the continued uncertainty about the relationship between motivation and engagement (Skinner & Raine, 2022). Are the terms synonymous, e.g., as represented in the National Research Council, Institute of Medicine (USA) report (2004)? Is motivation subsumed by the mega-construct of engagement (Christenson et al., 2012a; Fredricks et al., 2004; Newmann et al., 1992) ? Or does engagement create motivation (Salmela-Aro et al., 2022)? Most scholars recognise that the two share some sort of relationship, that the two are not a single construct (Guthrie & Wigfield, 2000; Saeed & Zyngier, 2012; Skinner & Raine, 2022), and that the most dominant perspective is that of a multidimensional construct (Christenson et al., 2012b; Salmela-Aro et al., 2022).

A third reason for the lack of agreement on a single theory for engagement is the more recent question of the role of the opposite of engagement (Reschly & Christenson, 2022). Labelled disengagement, disaffection or burn-out, the discussion is focussed on whether engagement and disengagement exist in a single continuum (i.e., disengagement is the complete lack of engagement) (Reschly & Christenson, 2012) or if they are two separate continua (Salmela-Aro et al., 2022; Wang et al., 2019).

The most prevalent understanding of engagement comprises the three dimensions of behavioural, cognitive and emotional engagement. Behavioural engagement is demonstrated by participation in activities and includes behaviours such as following rules and participating in activities (Fredricks et al., 2004; Mahatmya et al., 2012). Cognitive engagement includes effort, persistence, and contributing to class discussions (Fredricks et al., 2004). Emotional engagement encompasses “positive and negative reactions to teachers, classmates, academics and school … including interest, boredom, happiness, sadness and anxiety.” (Mahatmya et al., 2012, p. 47). Just as there are calls within the motivation research field to unite multiple componential theories to form a more comprehensive, rounded theory of motivation, so too are there calls to incorporate engagement within an integrated construct (Pintrich, 2003; Skinner & Raine, 2022).

Based on this review of the theories of motivation and engagement, and the recognition of the close and potentially inextricable relationship between them, the negative and positive aspects of motivation and engagement form a complex, multidimensional relationship, which impacts on educational outcomes. For the purposes of this thesis, this research needs to focus on the impact of these dispositions on the use of critical and creative thinking skills in STEM PBL.

5.4. Measures of Motivation and Engagement

Having identified a model of motivation and engagement which appears to be appropriate for our purposes – motivation and engagement as a complex, multidimensional construct including both negative and positive aspects of motivation and engagement – measures of motivation and engagement were reviewed to ascertain whether a suitable measure existed for the purposes of this study. Twenty-two motivation and/or engagement measures were compared for suitability. Those which indicated that elements of engagement and/or motivation were the focus of the measure were then scrutinised against the factors listed in Table 19. A short-list containing all the assessments considered suitable for use with primary school aged students is shown in Table 20. Appendix D contains the full list of the measures compared. Most of the measures reflect only a single dimension of motivation or engagement underpinned by a single theoretical perspective. As the MES-JS is the only measure that integrates both motivation and engagement, and both negative and positive aspects of the two constructs, it was considered to the be the most comprehensive measure. This is also the only measure to have been validated for use with Australian students. The MES and MES-JS have been used extensively by Martin in his research, including for measuring motivation and engagement during STEM learning (e.g., Green et al., 2007) . The MES-JS was therefore identified as the most suitable measure for use to assess the motivation and engagement of Australian primary school students. Reasons for the lack of suitability of the other measures include that the construct being measured is too specialised. For example, the MRP-R (Motivation to Read Profile: Malloy et al., 2013) only measures motivation with relation to reading and could not be generalised to a broader use. Also, some measures have not been validated for use in Australian schools which risks issues of cross-cultural differences, e.g., comparisons made between the MES and the SEI in an American context indicated that the construct of student engagement may not be transferrable from the Australian context without further testing and validation. The SEI has not been tested in an Australian context, so may not provide reliable results when used with Australian students (Reschly et al., 2014). Additionally, not all measures have been validated for use with the appropriate student age group, i.e., they have only been validated for use with secondary or tertiary students. Other issues which preclude the relevance of measures include the administration of the

measure being too complex for use in the classroom, or requiring specialised equipment e.g., Neuropsychological/physiological assessments or those requiring 1:1 interviews with students or observation of students. (Fulmer & Frijters, 2009; Kosovich et al., 2017). Finally, some measures were too limited in their coverage e.g., only including motivation or engagement, or only some of the components of engagement.

Table 19

Criteria for suitability for measures on motivation and engagement

Number	Criteria
1	Validated for use with primary school students
2	Measures a range of cognitive, behavioural and emotional engagement behaviours and a range of motivation behaviours
3	Has sufficient granularity[1] to identify change over a period of 10 weeks
4	Does not rely on specialised equipment or computers
5	Is inclusive of students with low literacy skills[2]
6	Is easy to administer in the classroom
7	Validated for use with Australian students
8	Is applicable to use in STEM-domain lessons

Note. [1]Granularity is identified as the level of detail in the information (Victorian Curriculum and Assessment Authority (VCAA), 2019) which can be identified by the measurement, (Rovers et al., 2019) e.g., a coarse-grained level would measure the presence or absence of a behaviour, while a fine-grained measure would be capable of measuring different strengths of the behaviour. [2]This stipulation was as a response to the low reading ability of some of the students (mainly boys) who appeared to show significant development in their skills during the STEM Learning Evaluation study but were unable to complete the assessment task successfully due to their low literacy levels.

Table 20

Motivation and engagement measures validated for application in primary school

Name of measure	Construct measured: Motivation/Engagement	Author	Focus	Meets criteria[1]
Student Engagement Instrument (SEI-E)	E	(Carter et al., 2012)	Cognitive and affective engagement	1
School Motivation Questionnaire (SchMOT)	M	(Conradty & Bogner, 2022)	Intrinsic and extrinsic motivation	1
Engagement vs. Disaffection with Learning – Student Report (EvsD)	E	(Skinner et al., 2009)	Motivational conceptualization of engagement and disaffection; measures behavioural + emotional engagement	1
Patterns of Adaptive Learning Scales (PALS)	M	(Midgley et al., 2000)	Social and cognitive aspects of motivation	1
Children's Academic Intrinsic Motivation Inventory, (CAIMI)	M	(Gottfried & Gottfried, 1996)	Academic intrinsic motivation	1
Student Motivation and Engagement Scale SMES-JS	M + E	(Martin, 1999-2022)	Adaptive and maladaptive engagement and motivation[2]. Can be used with generalized form of questions or subject specific; measures behavioural, cognitive and emotional engagement	1, 2, 3, 4, 5, 6, 7, 8
School Engagement Measure (SEM) - MacArthur	E	(Fredricks et al., 2005)	Relationship between classroom context and engagement; measures behavioural, cognitive and emotional engagement	1

Note. [1]Criteria listed in Table 19. [2]Adaptave forms of motivation and engagement produce positive outcomes, maladaptive produce negative outcomes.

The Motivation and Engagement Scale – Junior School (MES-JS) is a contemporary self-report assessment, developed in Australia (Martin, 2007) and validated by numerous Australian and international studies (e.g., Green et al., 2007; Liem & Martin, 2012; Martin, 2007, 2009; Martin et al., 2015). The scale shows an internal consistency (Cronbach's alpha) of over .70 reported (.70-.87) (Fredricks & McColskey, 2012), and incorporates a focus on cognitive and behavioural aspects of motivation and engagement (Liem & Martin, 2012). The MES-JS identifies four higher-order factors relating to positive motivation, negative (or maladaptive) motivation, positive engagement and negative (or maladaptive) engagement. Confirmatory factor analysis resulted in the four high order factors being further divided into 11 first order factor subscales: self-belief, valuing and learning focus (positive motivation), anxiety, failure avoidance and uncertain control (negative motivation), persistence, task management and planning (positive engagement), self-sabotage and disengagement (negative engagement). The MES-JS task was the only measure available that measured both engagement and motivation. As it is desirable to measure both, combining them in a single measure is an advantage for the administration of the task. The MES (high school version) had been successfully used previously to measure change in students' level of motivation and engagement following an intervention of a term, indicating that it provides a suitable level of granularity (Martin, 2007, 2008a). The MES-JS was also the only task which was validated for Australian primary school students.

5.5. Choice of Task to Measure Motivation and Engagement

In answer to Research Question 4, 'How are student motivation and engagement measured in primary school classrooms during learning activities?', the MES-JS has been assessed as suitable for this purpose. As this task measures both motivation and engagement and is validated for use with Australian primary school students, the MES-JS would be the most appropriate of the measures compared, for use in primary school classrooms. This is relevant to the overarching investigation, as motivation and engagement have been demonstrated to be enhanced by exposure to PBL (Hmelo-Silver, 2004; Jerzembek & Murphy, 2013; Specht, 2015). Chapter 9 describes the implementation of the MES-JS study to ascertain the effect of PBL on motivation and engagement in a classroom setting.

5.6. Current Gaps in Knowledge

The gaps in knowledge relating to motivation and engagement relevant to the research in this thesis, are in the intersection of motivation and engagement with STEM Problem-Based Learning in primary schools. As previously discussed, the lack of research on primary school students is significant. The research in this thesis adds to knowledge particularly in two areas.

The first area relates to the relationship between PBL and motivation and engagement in primary school. Although PBL is claimed to increase motivation and engagement in students (e.g., Blumenfeld et al., 1991; Hmelo-Silver, 2004; Norman & Schmidt, 1992), the evidence of this, particularly relating to primary school students, is lacking. Knowing what motivates and engages students, assists teachers to make informed decisions about choices they make in their teaching. This, in turn, provides for better outcomes for students.

The second gap in knowledge to be addressed by the research in this thesis, relates to the effect of problem type in PBL on motivation and engagement in primary school. By comparing the effect of the use of a hypothetical problem to motivate and engage students, with a genuine, multifaceted, unsolved problem provided by an industry partner, a greater understanding of the effects of different types of problems in PBL will be achieved.

Chapter 6.

Conclusion to Part A

This chapter briefly summarises the outcomes of the conceptual work described in Part A, with reference to Research Questions 1-4, and presents the revised research questions which underpinned the second phase of work, conducted in schools, to be presented in Part B.

Part A has presented the findings of a conceptual and theoretical exploration of a) the Australian Curriculum Critical and Creative Thinking learning continuum; b) tools for measuring the use of the skills described; c) the relationships between Problem-Based Learning and critical and creative thinking skills; d) the relationships between Problem-Based Learning and motivation and engagement; and e) the effect of industry partners on student learning in Problem-Based Learning. The outcomes offer a theoretical framework that can be used by educators and researchers to inform the design of effective and engaging learning experiences. By uncovering the theoretical underpinnings of PBL and its motivational dynamics, the research contributes to ongoing dialogue about innovative pedagogical approaches and their impact on student learning.

The conceptual research specifically examined research and other literature in the context of teaching Australian primary school students, through Problem-Based Learning with a STEM focus. Significant gaps in the literature were found in all areas of the investigation, and there was a particular dearth of literature focused on primary school students. The focus of the investigation was driven by the Research Questions developed in Chapter 1. The findings of conceptual research, summarised in Table 21 and Table 22, below, include descriptions of theory adaptations, theory synthesis and typology. The contribution of the conceptual research, recommendations and gaps found in the research literature and conceptual innovations are also detailed in Tables 21 and 22.

As a result of these investigations, Research Questions 1-4 were answered. Research Questions 2, 3 and 4 were refined and restructured as Research Questions 2a, 3a, 3b, 4a and 4b. These questions and the hypotheses, based on the conceptual research, are presented in Table 21. These five questions will now form the basis of the empirical studies described in Part B of this Thesis.

Table 21

Findings of the conceptual study — Research Questions 1 and 2

Research Question	Findings
RQ 1: What is the theoretical basis for the Australian Curriculum Critical and Creative Thinking learning continuum?	• The Critical and Creative Thinking learning continuum is loosely based on a number of different educational theories. A critical examination of the referenced theories utilising a new theoretical lens on the continuum identifies that it does not accurately represent the development of the skills described within it. The effectiveness of the structure of the continuum is thereby brought into question through its comparison to existing educational theories. • Further research on the skills and how they develop and are used by students is required if they are to be presented in a developmental continuum.
RQ 2: How are critical and creative thinking skills measured during learning?	• The C&CrT Continuum has a similar structure to the Delphi description of critical thinking skills. • An original definition of critical thinking as: *a collection of skills which enable focussed, purposeful, reflective thinking, and rely on the presence of various dispositions to be successfully enacted*, is proposed. • An original definition of creative thinking as: *a competency used to engage in the development of creative products, which incorporates divergent and convergent thinking, and is influenced by external factors such as the person and place associated with the creative act*, is proposed. • The relationships between critical thinking, creative thinking, divergent thinking and convergent thinking are clarified and a representation supplied. • Criteria are developed to identify measures suitable for use in the primary school classroom. • An original creativity assessment comparison grid is developed for use by researchers. • There are are no existing tasks that will effectively measure the critical and creative thinking skills as described in the C&CrT continuum, during student learning activities.

Table 22

Findings of the conceptual study — Research Questions 3 and 4

Research Question	Findings
RQ 3: How do the characteristics of STEM PBL affect the use of skills during learning?	• Clarity about the characteristics of the teaching style is essential if results of research are to be of use to other researchers. A visual representation of the relationship between PBL, PjBL and Inquiry Learning was developed. An original grid was developed to enable a comparison of the features of inquiry-based learning styles to be made, for use by teachers and researchers. • Very little research has been described on PBL in primary schools. • Very little research on PBL in primary schools in which the characteristics of the PBL method used are clearly described, has been reported. • PBL and critical and creative thinking skills are under-researched in in primary schools. • Research specifically on the relationship between creative thinking and PBL in primary schools is lacking. • Research on the characteristics of PBL in primary schools and motivation and engagement is lacking. • Research on motivation and engagement and critical and creative thinking skills in primary school students is lacking.
RQ 4: How are student motivation and engagement measured in primary school classrooms during learning activities?	• Motivation and engagement were identified as potentially playing an important role in the successful use of PBL, and the development of critical and creative thinking skills. • A synthesis of theories on motivation and engagement was accomplished, and measures of motivation and engagement were examined for relevance for use in the Australian primary school classroom during learning. • The MES-JS (Martin, 1999-2022) is suitable to use in response to RQ 4 to measure motivation and engagement.

Table 23

Revised research questions 2a, 3a, 3b, 4a and 4b, and associated hypotheses

Research Question 2a: What should be measured to assess critical and creative thinking in the classroom during learning activities?

Research Question 3a: What is the impact of PBL on student critical and creative thinking skill application?

Hypothesis 3a: Using PBL as a teaching approach with primary school students will enable students to apply a greater variety of critical and creative thinking skills at a variety of 'levels' at the end of a 10-week PBL intervention than when a more passive method of teaching is used.

Research Question 3b: What is the impact of using an industry partner problem in STEM PBL on student critical and creative thinking skill application?

Hypothesis 3b:
Students engaged in STEM PBL with an industry partner will show application of a greater variety of critical and creative thinking skills at a variety of 'levels' than students engaged in a theory PBL problem.

Research Question 4a: What is the impact of PBL on student motivation and engagement?

Hypothesis 4a:

Students will demonstrate an increase in positive motivation and engagement and/or a decrease in negative motivation and engagement at the end of a 10-week PBL intervention.

Research Question 4b: What is the impact of using an industry partner problem in STEM PBL on student motivation and engagement?

Hypothesis 4b: Students investigating an industry partner problem in STEM PBL will demonstrate a more statistically significant increase in positive motivation and engagement and/or decrease in negative motivation and engagement than students investigating a Theory PBL Problem.

Part B.
Investigating Problem-Based Learning and Skill Application in the Classroom

Chapter 1 posed the (now revised) overarching question, 'What is the impact of an industry partner in STEM PBL on the application of critical and creative thinking skills in primary school students?' The conceptual research described in Part A identified gaps in the literature and proposed innovative responses to some of the issues. This led to the revision of the research questions and five further questions to be investigated – Research Questions 3a, 3b, 4a and 4 b shown in Table 23, and Research Question 2a, 'What should be measured to assess critical and creative thinking in the classroom during learning activities?'

B.1. Methodology of Empirical Studies

As described in Chapter 1, the underlying assumption that the effective use of critical and creative thinking skills is a route to acquiring knowledge and achieving success in society, informed the development of the research questions and thereby also the research methodology. Further, it is assumed that all individuals possess the inherent potential for rationality and creativity, and that these traits are demonstrated by the use of critical and creative thinking skills, which are identifiable, measurable entities. It was hypothesised that critical and creative thinking skills can be enhanced through the specific context of STEM PBL, and that this assumption can be tested through empirical research.

The purpose of the investigation described in this thesis was to gather data about how students use their critical and creative thinking skills under specific conditions, i.e., a PBL approach in a STEM context, and working with an industry partner. The expectation was that the data generated from this research would provide the evidence required to enable the overarching question to be answered, 'What is the impact of an industry partner in STEM PBL on the development of critical and creative thinking skills in primary school students?' To achieve this end, an empirical study was designed (described in Study Three, Part B), however the search for an appropriate measure to assess the skills being used in the classroom during learning, required extensive investigation which led to the conceptual studies described in Part A of this thesis. The conceptual studies analysing the Australian Curriculum C&CrT continuum, and critical and creative thinking, provided the theoretical foundations for the development of a measure which was then investigated in Part B. The studies examining PBL, and motivation and engagement were unable to answer the overarching question.

However, they provided the theoretical background to point to the possible relationships between STEM PBL, motivation and engagement, and the critical and creative thinking skills, indicating a positive relationship between STEM PBL and industry partners and the use of critical and creative thinking skills.

Part A confirmed that classroom-based studies would be necessary to prove whether PBL and industry partners have an effect on students' critical and creative thinking skills during learning. These studies, described in Part B, enabled student behaviour to be captured, and student thinking to be explored while the students were in the process of using the skills under investigation. Research on primary school students takes various forms, e.g., classroom observations, questionnaires, testing, interviews, etc. (e.g., Creswell & Guetterman, 2018; Lodico et al., 2010; Mitchell, 2009). Using a Pragmatic framework through a Scientific Realism viewpoint, the research described in these studies incorporated a combination of multiple research methods (a mixed methods correlational approach in a test development framework, introspective methods and an experimental approach guided by a scientific realism viewpoint) to enable the required data to be collected (Aldridge et al., 1999; Lodico et al., 2010; Williams, 2013). By conducting new classroom-based studies the research generated up to date knowledge on skill use by primary school children, which helps fill the gaps identified in the areas of research on primary school students, PBL and motivation and engagement.

The classroom-based studies were developed as a series of empirical studies designed to incrementally develop the necessary foundations for the final study (Study Three). These incremental stages result in the development of a measure and an assessment task suitable to measure both the skills being used by students during their learning, and to measure the change in the skills being used in a longitudinal study.

B.2. Research Design and Methods

In Part B, the first of the empirical studies, Study One, investigates Research Question 2a; 'What should be measured to assess critical and creative thinking in the classroom during learning activities?'. This question was developed as a result of the conceptual study of critical and creative thinking skills described in Chapter 3. Study One is an experimental pilot study in which the first stage of the development of a novel measure is described. The measurement variables for assessing critical and creative thinking skills were identified and refined to enable a measure suitable for use with primary school students to be developed. As well as aligning the re-worded statements with the continuum, coverage of critical and creative thinking skills was assured by comparison with

classifications of skills described as convergent and divergent (see Chapter 3), critical thinking (Delphi Report: Facione, 1990) and creative thinking (see Chapter 3), and through the OECD classification of desirable skills (Vincent-Lancrin et al., 2019). An expert panel of teachers was utilized to refine and validate the descriptions of the skills to ensure their suitability for primary school students and assure content validity. A self-report measure was developed and trialled with students, the researcher presenting the measure to the students while the classroom teacher simultaneously provided an assessment on each of the students. The results were tabulated and reviewed, and as a result of this pilot study a model was developed through further conceptual research and integrative theory construction[18], with the SOLO taxonomy (Biggs & Collis, 1982, 1989). This enabled the development of a theory on the use of some critical thinking skills. Initial considerations and recommendations were also made regarding a suitable format for the design of the measure to assess the skills. These recommendations provided the structure on which the Critical and Creative Thinking Self-Report Measure, developed in Study Two, was based.

Building on the findings of Study One, Study Two addressed Research Question 2b, 'How does the method of communicating a self-report measure to students affect the accuracy of their responses?' This second empirical pilot study resulted from my observations a) from student reactions to the measures presented in Study One, and b) of the customary classroom practice of providing instruction to students simultaneously in both verbal and text formats. Study Two begins by engaging in a theoretical integration of theories of Multimedia learning and Cognitive Load Theory. This application of multimedia learning theory contributes to understanding the role of multimedia presentations in assessments. The hypothesis was that students, and their teachers who were experienced with using the Critical and Creative Thinking learning continuum, would be accurate in judging the student's use of skills. The data was analysed with Fleiss' kappa, a measure of inter-rater agreement for categorical scales, to determine the level of correlation between the student and teacher responses. The results of the analysis indicated a surprisingly low agreement level between students and teachers about the student's use of skills. This led to a thorough investigation into the possible reasons for a low rate of agreement, and recommendations for remediating this effect. Although the hypothesis was not supported that the combination of simultaneous verbal and text formats would produce a higher correlation between the student and

[18] The integration of existing theories with one's own ideas to develop a new theory. Kalmar, D. A., & Sternberg, R. J. (1988). Theory knitting: An integrative approach to theory development. *Philosophical psychology*, *1*(2), 153-170.

teacher scores, the study provided useful data on the format and presentation of the measure and the assessment task.

Study three, the final empirical study, investigated how the characteristics of STEM PBL affect the use of skills by students during learning, and their impact on motivation and engagement. The chapter investigates the four remaining research questions (3a, 3b, 4a and 4b), shown in Table 23, with their accompanying hypotheses. This quasi-experimental study[19] employed an Educational Intervention Framework by implementing a STEM PBL intervention in primary school classrooms over a 10-week term. Teachers inexperienced in PBL were trained in PBL by the researcher to implement the intervention with their students in the classroom. A control group, a theory group and an industry partner Group completed pre-, post- and follow-up intervention assessments presented by the researcher (the Critical and Creative Thinking Self-Report Measure developed in Study Two, and the MES-JS (Martin, 1999-2022) which were administered to measure the use of critical and creative thinking skills, and student motivation and engagement. The longitudinal nature of the study enabled the data to be analysed using the equivalent of a one-way repeated measures ANOVA for a dichotomous dependent variable (Cochranes's Q) allowing the change in the levels of the skills being used by the students over the period of approximately 16 weeks to be measured. This enabled comparisons to be made about the use of skills in a control group and a PBL group, and between students engaged in a hypothetical problem and a problem provided by an industry partner.

Together, these conceptual and empirical studies enabled the question of whether STEM Problem-Based Learning with industry partners helps students use their critical and creative thinking skills during everyday learning, to be answered in the affirmative. Not only did students use their critical and creative thinking skills more through Problem-Based Learning, but the use of skills was even greater when an industry partner provided the problem. Further, the longitudinal nature of the study and inclusion of a control group enabled test-retest reliability of the Critical and Creative Thinking Self-Report Measure to be confirmed. This confirms that the students are responding consistently to the statements, i.e., the measure reflects true scores rather than measurement error (Cozby & Bates, 2012, p. 98).

In the final chapter in Part B, Chapter 10, the results of the research in these chapters are discussed. Three contributions of the studies to primary education research and the contribution of

[19] A true experimental design was not possible due to a lack of ability to randomly allocate the teachers and students to the treatment conditions.

the studies to addressing nine gaps in the literature are described. Recommendations are made regarding the use of industry partners in the classroom, and of the wider applications of the Critical and Creative Thinking Self-Report Measure.

The final chapter in the Thesis, Chapter 11, provides an answer to the revised overarching question, 'What is the impact of an industry partner in STEM PBL on the application of critical and creative thinking skills in primary school students?'. The need to understand the development of the critical and creative thinking skills more deeply and the question of the relevance of a continuum of skills is broached. Wider implications for the use of industry partners in schools and the relevance of this research worldwide is discussed.

Chapter 7.

Study One: Initial Development and Pilot Study of a Measure of Critical and Creative Thinking

7.1. Introduction

The general capabilities in the Australian Curriculum (version 8.4[20]) provide guidance on the critical and creative thinking (C&CrT) skills students are expected to learn (Australian Curriculum Assessment and Reporting Authority (ACARA), 2013a), in the form of the Critical and Creative Thinking learning continuum (C&CrT continuum; Australian Curriculum Assessment and Reporting Authority, 2019). The continuum includes descriptions of the critical and creative thinking skills which students are expected to have developed by the end of specified years of schooling (Australian Curriculum Assessment and Reporting Authority (ACARA), 2013b, p. 7), but it provides no guidance on how to assess these skills (Australian Curriculum Assessment and Reporting Authority (ACARA), 2013b, p. 5). While I was working with teachers and students on STEM Problem-Based Learning in 2015, it became clear that teachers were unsure how to effectively measure and report on these skills. A new teaching approach (PBL with an industry partner) which was reputed to have a positive effect on student skills, was being trialled in the school[21]. This opened up an opportunity to assess how effectively students were using their critical and creative thinking skills. Although teachers observed and discussed changes in their students' use of skills, they lacked sufficient understanding of the skills described in the continuum to confidently measure and report on what they were observing. Additionally, although aware of the need for reliable assessments which would avoid subjective judgements (Brown & Race, 2013), teachers lacked suitable measures to assess the impact of PBL on the students.

Although a plethora of measures are available to measure a variety of skills (e.g., see Cropley, 2000; Facione, 1991; Torrance, 1998; Treffinger et al., 2002), as described in Chapter 3, the literature did not reveal a suitable measure for the purposes of this research, i.e., one which was specifically related to the skills as described in the Critical and Creative Thinking learning continuum,

[20] Version 9 of the Australian Curriculum was published in 2022 following the completion of this research. Although changes were made to the structure of the Critical and Creative Thinking Capabilities, and to the descriptions of the skills contained within, the changes did not reflect on the relevance of this research.
[21] Although PBL teaching was known of, PBL with industry partners was a new approach to teaching and learning in the school in which I was working.

and appropriate to measure the skills in use during everyday learning by students in primary school. The self-reflection measure developed for the 2015 STEM Evaluation Survey (see Appendix A: Gaardboe, 2019) was designed to capture students' beliefs about the skills they were using, rather than the lingering evidence of the thinking, as described by the socio-cognitive view of assessment (Mislevy, 2012). Although this measure provided some insight into the skills students were using, it also highlighted the need for a task that was fit-for-purpose (Brown & Race, 2013). Such a task would relate to the skills students were using during learning, it could be completed in one lesson and would be applicable to students with varying levels of literacy. In short, a task that fits the criteria developed in the course of the current research project and described in Chapter 5 (Table 19).

This chapter describes the initial development phase for creating and testing a novel assessment which incorporates a novel measure of Critical and Creative Thinking: a Lego task stimulation activity and 'brain break' activities developed by the researcher. The Critical and Creative Thinking measure was designed to address Research Question 2a, 'What should be measured to assess critical and creative thinking in the classroom during learning activities?' This chapter also provides initial considerations and recommendations for how to assess the skills, which is addressed in Research Question 2b.

The effects of the COVID-19 pandemic began to be felt in Adelaide in early 2020, just at the time that Study One was planned for implementation. Public gatherings were cancelled, and the Department for Education in Adelaide implemented a lock-down policy restricting school access to essential staff and students only. The lack of knowledge about the long-term implications on access to students in class made it difficult to plan the most effective data collection activities to support the planned research. The initial eleven schools with up to 400 students dwindled to a last-minute approval to collect data in 3 classes across two schools and three year-groups, in the final weeks of the school year. Having no confidence that access to classes would improve in the short term, the decision was made to test two answer formats in the anticipation that sufficient responses would be collected to enable some insights which would inform the later studies. Despite the limited number of students, which had implications for the reliability of data analysis, useful insights were gained through this study, which influenced the design of the two later studies described in Chapter 8 and 9.

7.2. Research Design and Methods

Scope of Research

The purpose of this study was twofold. The first purpose was to develop a self-report task which would enable primary school students to report on the levels of critical and creative thinking skills they were using in STEM Problem-Based Learning (PBL). The second purpose was to investigate how primary school students and their teachers responded to two different response formats of the measure. This phase of the research was not intended to present conclusions of a statistical nature or to draw inferences for a wider population. Rather, this was a pilot with a focus on developing a task that was fit-for-purpose and identifying and solving administrative and logistical issues relating to working with and assessing primary school students in the classroom.

Research Design

The research design was shaped by both the results of the conceptual studies described in Part A and the limitations of the COVID-19 pandemic. Figure 10 illustrates the main steps of the multi-step process (DeVellis, 2012; Downing, 2006) used to develop the assessment described in Study One. These can be further expanded on as follows:

- Step one: Development of statements: Statements in student-friendly language were developed, based on the descriptions of skills specified in the C&CrT continuum and other relevant literature.
- Step two: Expert panel review and validation. An expert panel (primary school teachers each with a minimum of 8 years teaching experience, a familiarity with the AC C&CrT learning continuum, and confidence to assess students against this continuum) was engaged to review the statements to ensure they accurately described skills that students in primary school classes would be expected to use at the described year levels. This provided content validity for the measure.
- Step three: Presentation and Response Format Development. Two formats of the Critical and Creative Thinking measure were developed for testing to determine which type of response format would more effectively capture the information being sought.
- Step four: Data Collection. A pilot study was run to gauge the student response to the two response formats and to gauge the teacher response to the format of a teacher version of the Critical and Creative Thinking measure.

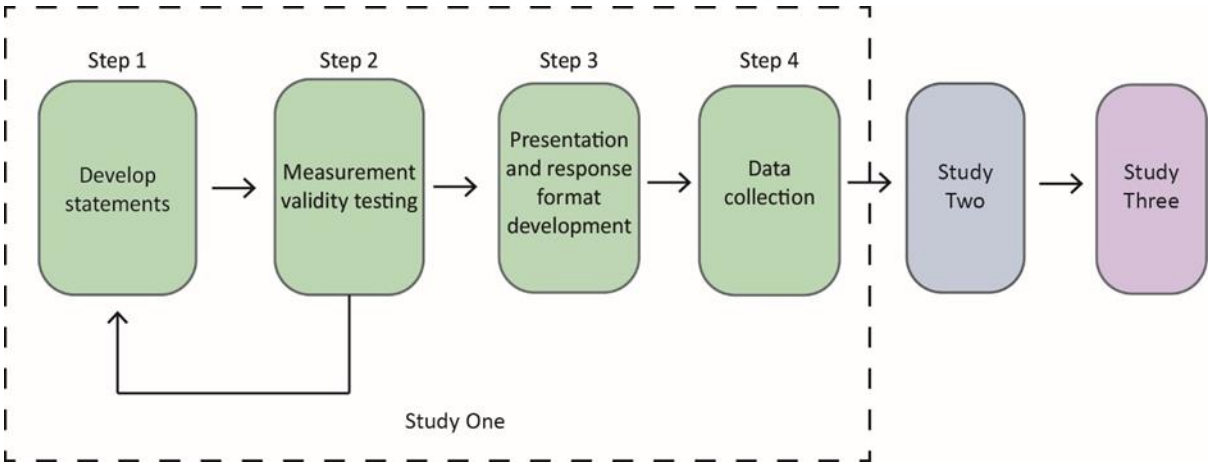


Figure 10

Research design for Study One

7.3. Step One: Development of Statements

In the Australian Curriculum Critical and Creative Thinking (C&CrT) learning continuum (V8.4) the skills are arranged under four categories or 'elements' and then into 'sub-elements', as shown in Figure 11. A statement relating to the level of skill students are expected to be able to achieve by the end of the specified year level ('level' statement) is provided for each sub-element covering expectations for students from Foundation to Year 10 (Australian Curriculum Assessment and Reporting Authority (ACARA), 2013b, p. 7). This provides 72 statements in total, describing the skills and their various stages. As discussed in Chapter 2, the description of the C&CrT continuum as a learning continuum created the expectation that statements at each consecutive level within a sub-element would describe progressive steps in each skill (Victorian Curriculum and Assessment Authority (VCAA), 2019, p. 4). The conceptual research described in Chapter 2 identified shortcomings in the continuum structure. The decision was therefore taken to analyse the continuum in more detail prior to developing a measure.

In order to develop a measure based on the C&CrT continuum, it was necessary to dissect the sub-elements to ensure that only one skill was described at each level. Where statements existed in the C&CrT continuum, they were incorporated. As no research could be found in the literature that described the levels of the skills in question at this level of detail, where there were

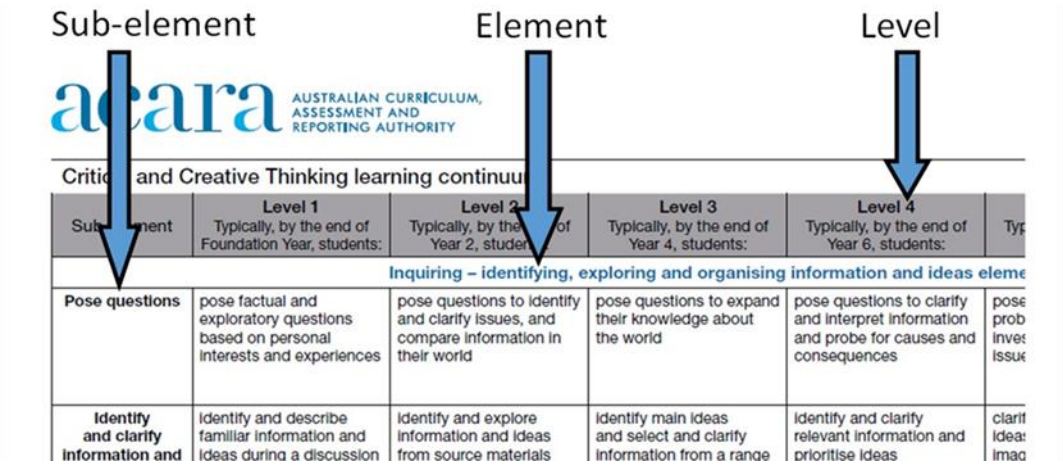


Figure 11

Terminology of the structure of the Critical and Creative Thinking learning continuum (V8.4)

gaps, additional stages were created. The descriptions of these stages were based on: the observations of students by the researcher; discussions with teaching colleagues; feedback from the expert panel. This process produced an excessive number of statements, which could not all be utilised in a single measure. The number of skills was therefore culled using the process described below.

Although the primary aim of this next step was to ensure there was sufficient coverage of the skills identified in the C&CrT continuum while still reducing the number of statements, it was also desirable to maintain coverage of the breadth of critical thinking and creative thinking skills. A broader analysis was therefore carried out which related the categories of skills to other international skill classification categories.

As discussed in Chapter 3, skills can be classified in various ways. Four classification methods were used to guide the coverage of skills in the measure and to ensure the domains of both critical and creative thinking skills were adequately sampled. The four methods were: 1) the classifications of the elements in the C&CrT learning continuum (Australian Curriculum Assessment and Reporting Authority (ACARA), n.d.-b); 2) critical and creative thinking (Baker et al., 2001; Ellerton, n.d.; Facione, 2020); 3) divergent/convergent thinking (Acar & Runco, 2012; Cropley, 2006); and 4) the four

Table 24 (over)

Critical and creative thinking measure statements and their categories

Australian Curriculum Skill / description	Statement	Year	Category
Asking Questions/ Gathering information	I ask questions about things I'm interested in.	F	Critical Thinking, Inquiring, Convergent, AC: 1
	I ask questions to compare information about things around me.	2	
	I ask questions to understand about things I notice in the world.	4	
	I ask questions to understand why and how things happen.	6	
	I ask questions to help me learn about more complex (complicated) issues.	8	
	I ask questions to understand and analyse very complicated issues and abstract ideas (theories).	10	
Finding and organising information/ Selecting and organising information	I find and organise information by collecting and sorting things that are the same or similar.	F	Critical Thinking, Inquiring, Convergent/Divergent, AC: 1
	I find and organise information by identifying and ordering it.	2	
	I find and organise information by collecting, comparing and putting it into categories.	4	
	I find and organise information by choosing the relevant part to use and knowing which are the facts.	6	
	I find and organise information by carefully selecting and analysing it against criteria I have been given.	8	
	I find and organise information by comparing it against criteria I have developed to make sure that it is reliable.	10	

Note. Categories: Critical Thinking/ Creative Thinking; OECD: Inquiring/ Imagining/ Doing/ Reflecting; Divergent Thinking/ Convergent Thinking; Alignment to the AC continuum elements: AC: 1 = Inquiring - identifying, exploring and organising information and ideas element; AC: 2 = Generating ideas, possibilities and actions element; AC: 3 = Reflecting on Thinking and Processes element; AC: 4 = Analysing, synthesising and evaluating reasoning and procedures element.

Table 24 (continued)

Australian Curriculum Skill / description	Statement	Year	Category
Connecting information/ Idea Generation	I can connect information about different things.	F	Creative Thinking, Imagining, Convergent, AC: 3
	I can connect information I learned from doing something else, and use it to come up with a new idea.	2	
	I can connect information from different places, even when it is about something completely different.	4	
	I can connect information from different places, and use it to help me learn and understand something completely different.	6	
	I can connect information I learned somewhere else, and I can explain why I choose to use it when I'm working in a different subject area.	8	
	I can connect information from different places, I plan to do it, and I can explain why it improves my ideas.	10	
Evaluating a plan or idea/ Idea evaluation, Applying logic and reasoning (metacognition)	When I've made a plan to do something I can tell you what decisions and choices I had to make.	F	Critical Thinking, Reflecting, Divergent /convergent, AC: 3 & 4
	When I've made a plan to do something I can explain the decisions and choices I had to make, and why I made them.	2	
	When I've made a plan to do something I can compare my plan to someone else's, and explain which will be the best plan.	4	
	When I've made a plan to do something I can change my plan if the situation changes or if I find new information.	6	
	When I've made a plan to do something I can see where more information is needed, or where parts of the plan may need more thought.	8	
	When I've made a plan to do something I can discuss how reliable my information is and tell you where there might be mistakes in my reasoning.	10	

Table 24 (continued)

Australian Curriculum Skill / description	Statement	Year	Category
Coming up with ideas/ Idea Generation, Idea Evaluation	When I've come up with an idea I just go with it.	F	Creative & Critical Thinking, Doing /Reflecting, Divergent /convergent, AC: 2
	When I've come up with an idea I see if I can make it better by coming up with more ideas.	2	
	When I've come up with an idea I compare it with other people's ideas and sometimes add some of theirs to mine to make it better.	4	
	When I've come up with an idea I choose the best option after I've carefully tested a few.	6	
	When I've come up with an idea I can describe what the problems might be, and how I might change my idea to avoid them.	8	
	When I've come up with an idea I test options, analyse solutions and justify choices.	10	
Creating new ideas/ Idea generation	I create new ideas by looking closely at things around me to give me ideas.	F	Creative Thinking, Imagining, Divergent /convergent, AC: 2
	I create new ideas by adding to things I already know.	2	
	I create new ideas by coming up with more than one idea and then choosing the best.	4	
	I create new ideas by combining ideas from more than one place, and in more than one way.	6	
	I create new ideas by seeing how my ideas are related to existing ideas, and making them more imaginative.	8	
	I create new ideas by coming up with complex solutions that combine ideas and are imaginative and unusual.	10	

Note. Categories: Critical Thinking/ Creative Thinking; OECD: Inquiring/ Imagining/ Doing/ Reflecting; Divergent Thinking/ Convergent Thinking; Alignment to the AC continuum elements: AC: 1 = Inquiring - identifying, exploring and organising information and ideas element; AC: 2 = Generating ideas, possibilities and actions element; AC: 3 = Reflecting on Thinking and Processes element; AC: 4 = Analysing, synthesising and evaluating reasoning and procedures element.

categories described in the OECD analysis of Critical and Creative Thinking (Vincent-Lancrin et al., 2019). Table 24 sets out the reduced set of skills chosen for use in the Critical and Creative Thinking measure (column one), and the statements identified to represent the appropriate levels as described in the continuum (column two), along with the category relating to the four chosen classification systems (column three).

Previous versions of a measure of critical and creative thinking, based on the AC C&CrT continuum, had been trialled in the classroom during the STEM Learning Evaluation Study described in Chapter 1. The format used was considered appropriate for use with middle to upper primary school students, and this format therefore formed the basis for the measures trialled in Study One.

7.4. Step Two: Measurement Validity Testing

A measure such as the Critical and Creative Thinking measure would usually be expected to undergo testing for validity and reliability to ensure that it is in fact measuring what is intended (Cozby & Bates, 2012, p. 102; DeVellis, 2012, p. 95). Validity pertains to the precision of the conclusions drawn from a measure, whereas reliability describes the constancy and durability of the measure over time or in various conditions (Cozby & Bates, 2012).

Having ensured there was sufficient coverage of the areas identified in the C&CrT continuum, and that statements covering a range of levels for each skill had been developed, feedback was sought from a panel of experts to validate the statements. The expert panel consisted of five primary school teachers with experience from Foundation to Year 7 in both the private and public sectors and with experience of teaching in schools from a range of socio-economic levels. Further, the teachers were all familiar with the AC C&CrT continuum, and were confident to assess students against the continuum. The panel members were individually asked to evaluate how well the statements reflected the capabilities of students at the target year level, and whether the statements represented a progressive sequence in their opinion. Finally, they were asked to suggest changes to the statements to improve clarity and ensure appropriateness of the language for upper primary students e.g., ‘Students justify reasons for decisions when transferring information to similar and different contexts’ was adjusted to ‘I make connections between information I learned in different places, subjects or times. I can explain the connections and why they improve my ideas’.

Measurement scales are designed to measure phenomena that are believed to exist, but are not easily observed (DeVellis, 2012, p. 15). This may be because behaviour cannot be assumed to be a reliable indicator of the phenomena, or because there is no access to behavioural information. Such

measures or scales consist of a series of items whose values are caused by an underlying construct (the latent variable or 'effect indicator': Bollen, 1989, p. 65). Each of the items within the scale is expected to evoke a similar response. The Motivation and Engagement Scale, for example, includes four questions on self-belief: 'If I try hard, I believe I can do my schoolwork well'; 'If I don't give up, I believe I can do schoolwork that is hard'; 'If I have enough time, I believe I can do well in my schoolwork'; and 'If I try hard enough, I believe I can do all my schoolwork'. A respondent would be expected to respond to each of these four statements in a similar way.

An index, on the other hand, is designed to determine the level of a construct. Each item in the index is a "cause indicator" rather than an "effect indicator" (DeVellis, 2012, p. 17). Items that share certain characteristics are grouped together under a common category heading such as 'Asking Questions', but the items are not expected to evoke a similar response, and grouping them together does not infer any causal linkage (DeVellis, 2012). As each item (the statements) in the measure developed in Study One are supposed to relate to different levels of the same skill, it was not expected that the students would respond to all six of the questions in the same way. Each group of six questions therefore works as an index or sub-scale rather than as a factor within a scale. Factor analysis was therefore not applicable as a method of validation for the measure being developed.

In addition to being unable to carry out a factor analysis, due to the nature of the Critical and Creative Thinking Measure, the lack of research literature on the individual skills caused some uncertainty about how students would respond to the items in each construct. Although the majority of students were expected to respond to the statement equating to their year level with a positive response, they may also have indicated that descriptions of skills used at less advanced levels were still being used. If students only indicated one statement in an index as being representative of their use of a skill, and no other statements in the index were relevant to them, this would demonstrate an 'unfolding' scale or Thurstone Scale (Andrich, 1985). If, on the other hand, students indicate a number of consecutive statements as being representative of their use of a skill, this would more closely demonstrate a 'cumulative' scale, or Guttman scale (DeVellis, 2012, p. 119). In this case, multiple items within the index will be appropriate, and the point of transition from a positive response to a negative response will mark the highest level of achievement within that index.

In the case of both a Thurstone scale and a Guttman scale, the latent variable will not demonstrate an equally strong causal relationship with each item in the index. Standard measures of reliability and validity therefore do not all apply to these scales (DeVellis, 2012). Content validity (DeVellis, 2012, p. 102), accomplished by seeking the opinions of experts in the field of primary school education, was therefore chosen as the most suitable option. Other tests of reliability such as inter-rater

reliability and parallel or alternative forms reliability (Creswell & Guetterman, 2018, p. 159) could not be carried out due to the nature of the measurement task (one teacher and one student judging the student), and the lack of similar instruments with which to compare this novel measure. Test-retest reliability (Creswell & Guetterman, 2018, p. 158) was not possible in this study due to the timeframe of the studies being unavoidably shortened due to the constraints of the COVID-19 pandemic.

As with other tests of validity, convergent or criterion validity could not be tested as this requires a comparison of results with other similar measures (Leedy & Ormrod, 2015, p. 115). In this case, no other measures exist that measure the skills as described in the Critical and Creative Thinking measure. Although it was hoped to test for concurrent validity (Cozby & Bates, 2012, p. 104) this was not possible due to the reduced access to participants forced by the pandemic.

The reading level of the statements was calculated according to Fry (1977; DeVellis, 2012). These levels are only reported at grade levels, and no indication is provided of the average age for grade levels. It is assumed that these approximate to Australian school levels[22]. Although the average level of the statements in the measure registered at a grade 9 level, the expert panel believed this was not above the capacity of the students' understanding if the statements were read to them. This was based on the assumption by the panel of teachers that listening comprehension is greater than reading comprehension for the majority of primary school students. This assumption is neither fully supported nor disproved by research. The early influential study by Sticht et al. (1974) certainly supported this assumption, but more recent studies (e.g., Horowitz & Samuels, 1985; Royer et al., 1990; Verlaan et al., 2017), indicate that proficient readers demonstrate a greater level of comprehension from reading than listening. Findings for the level of comprehension for nonproficient readers are also under some dissent, with Horowitz and Samuels (1985) finding listening comprehension exceeding reading comprehension at all levels of primary school including year 6, while Verlaan et al. (2017) reports findings of reading comprehension exceeding listening comprehension at all year levels.

7.5. Step Three: Presentation and Response Format

The following considerations were taken into account while considering the layout and form of presentation of the Critical and Creative Thinking measure. It was important that all students received the same measure for the results to be comparable. Providing a simplified form of the measure which only contained the Foundation to Year 6 statements (i.e., leaving out the Year 8 and 10 statements for

[22] The average age for a student in Year 6 is 11 years in Australia, USA, Canada and India. In the UK students are 10 years old in year 6. In Northern Ireland they are 9 years old. *Comparing Class and Grade Level across Countries*. (n.d.). http://seandelaney.com/wp-content/uploads/2018/07/Comparing-Class-and-Grade-Levels-across-Countries.pdf, *Grade Equivalents*. (n.d.). https://www.asmilan.org/admissions/grade-equivalents

each of the sub-elements) was considered for use with younger students (e.g., Year 4). Although this would make the measure less daunting for younger students by reducing the size of the measure from 72 to 48 statements, it was considered possible that even younger students might be able to use skills at some of the higher levels. It was also not desirable to have one measure for younger students, and an alternative one for older students. Therefore, only one form of the statements was produced for use in both forms of the measure. Statements were designed to be as short as possible while retaining as much of the intent of the Critical and Creative Thinking learning continuum as feasible.

Two alternative response formats and styles of presentation were chosen for testing: a) using a dichotomous (yes/no) scale response and a verbal only presentation, and b) using a 5-point Likert style scale and a verbal combined with a text presentation format. Because oral comprehension is often more advanced than written comprehension for primary school students (Diakidoy et al., 2005), the presentation method took this into account, incorporating a verbal format with both tasks. It was hypothesised that this would enable the inclusion of students with lower literacy levels and ensure they would still be able to respond to the statements in the measures. It was desirable to include these students, as prior experience in the classroom had demonstrated that students with poor literacy skills are capable of using advanced critical and creative thinking skills, and it was desirable to capture their responses in this measure of critical and creative thinking. Table 25 sets out the main features of the critical and creative thinking measure.

Table 25

Main features of the components of the critical and creative thinking measure

Component	Features
Lesson format	Researcher to instruct and supervise students Length reduced to one 50 min lesson
Response Type	Assessment Task 1: dichotomous (yes/no) scale response Assessment Task 2: 5-point Likert style scale
Order of statements	Statements in random order
Presentation format	Assessment Task 1: Verbal only Assessment Task 2: Verbal + text

Student Assessment Task 1

The first version of the Critical and Creative Thinking Measure, included in Student Assessment Task 1 (Appendix F-2), was a tick-box format with dichotomous scale response in which students were asked to respond to each statement with the choice of response: 'yes' (I do this), or 'no' (I do not do this). This measure was designed with a purely verbal presentation. Students were to be provided with a response sheet consisting of tick boxes, and all statements would be read to them twice. A verbal only presentation format is not without precedent, e.g., with the MRP assessment of reading motivation (Malloy et al., 2013).

Student Assessment Task 2

The second version of the Critical and Creative Thinking measure, included in Student Assessment Task 2 (Appendix F-3) had the students responding to the same statements but being offered a choice of responses in a five-point Likert style scale specifying frequency of use of each level of skill for each of the statements. The choice of response in the Assessment Task 2 measure was: 'never' (I never do this); 'occasionally' (I occasionally do this); 'sometimes' (I sometimes do this); 'often' (I often do this); or 'always' (I always do this). Although both versions of the measure were asking if the students used the skills as described, the change in format of responses effectively changed the focus of the response from what the students did (the Assessment Task 1 measure), to how often they did it (the Assessment Task 2 measure). This change in the response format enabled *what* was being measured to be varied between the two measures, even though the statements read to the students were identical. The Assessment Task 2 measure incorporated a combined verbal and text format by providing the statements for students to read as they were read aloud by the researcher. The verbal combined with text format was chosen to support students in the class with lower reading levels, and to support those with difficulties in auditory processing. Using dual modes of presentation in this way is considered standard classroom practice in schools in which I have worked. The likelihood of students mis-reading the statements was reduced by asking them to read at the same pace as the verbal presentation. This was also expected to slow some of their response times, encouraging students to consider their response before committing to an answer.

A presentation was developed to be run during the testing session with both presentation formats to help demonstrate which part of the measure students should be on at any time.

The Teacher Task

A third measure was developed to be used by teachers (Appendix F-4). This measure was identical to the Student Assessment Task 2 measure but also included the question, 'How confident are you to give this

score?', for each of the six sub-elements. The response to this question was in the form of a 3-point categorical scale (a little/moderately/very). The ultimate purpose of the teacher version of the Critical and Creative Thinking measure, was to enable a comparison to be made between the teacher's assessment of each student, and each student's self-report. A copy of the teacher task was expected to be completed by the teacher for each of the students in the class, during the presentation of the assessment task to the students by the researcher.

Presentation of the Measure

To reduce possible variables the Critical and Creative Thinking measure was designed to be presented by the researcher to each class of students. The time to complete the measure was limited to within a 50-minute lesson to fit with most school timetables.

The next consideration related to the order in which the statements were presented. From my personal experience in the classroom and that of my teaching colleagues, it was considered likely that rather than carefully considering the meaning of each statement and choosing the one that best represented their level of skill, some students would be inclined to identify their level depending on where they thought they were in the class (I'm a middle achieving student so I probably do the middle statement). The statements were therefore presented in a random order, and the students were instructed that this was the case. It was hoped this would encourage students to identify what they actually did, rather than what they thought they should be able to do.

As previously described, each assessment consisted of a measure (the Critical and Creative Thinking Measure), and various activities. A pre-test activity was devised for the commencement of the assessment to help the students' recall of some of the skills they might have used in different situations, before they were asked to respond to the statements about the skills in the measure. The activity was to consist of students in groups of two or three being presented with a small bag of Lego pieces and a pictorial instruction sheet. The instruction sheet was from one of two commercial Lego kits (11006 and 11007) suitable for ages four+. Each bag contained a) the original instruction sheet for a single model (a choice of four sets of instructions were used to ensure variety across the class); b) some, but not all the blocks from the kit; and c) a random selection of some additional Lego blocks. The students were to be warned of this and asked to 'Do the best you can to make the thing in the picture'. The students were also to be encouraged to talk and ask questions during the five-minute activity, and at the conclusion of the activity were to be asked questions about the sorts of questions they had asked, and the sorts of conversations they had taken part in. The Lego activity was to provide an opportunity for the students to use divergent and convergent thinking skills as they analysed the equipment and instructions, asked questions, collaborated with their peers, drew on previous experiences with Lego, applied creative

solutions to their problems and changed their plans when necessary. It was hypothesised that class discussion around the activity would help students connect the skills they were using for their Lego activity with the skills being asked about in the measure. Explicit connections to skills were to be suggested by the researcher during the discussion, e.g., by asking the students about different types of questions they might have asked each other, and by having students describe connections they had made to the use of Lego outside school, to help them solve the problem they were now faced with. Students were clearly instructed that the responses to the questions in the measure were to be about their learning generally, not just what they did in the Lego task.

Opportunities were also planned for during the presentation of the Critical and Creative Thinking Self-Report measure, for the students to consider whether there were times during their regular school lessons when they had used the skills they were now being asked about in the measure. This was to be done by interrupting the presentation of statements with questions for the class to which an answer was required, e.g., 'What have you had to find information about recently?'; 'Did you have to do something with the information (find the facts, sort it under topic headings)?'; or 'Did you just copy it into your book?'

The insertion of 'brain breaks' during the presentation of the measure was also planned, as they have been shown to positively affect both achievement and engagement levels in primary school students (Gernes, 2021; Ostroff, 2014). These took the form of simple physical activities ('Wriggle your fingers, pat your head and rub your tummy' etc) and small games (e.g., Rock Paper Scissors) as described in the script in Appendix F-1. Teacher expectations are that a ten year olds attention can be maintained for between approximately 10 – 30 minutes (Shakibaie, n.d.). These activities were therefore spaced throughout the paper-and-pencil component of the task to break the tedium of sitting still, concentrating, and listening to a presenter's voice for up to an hour.

7.6. Step Four: Data Collection

Having developed the critical and creative thinking skills measure and designed the assessment procedure to present the measure to the students, Step Four was to carry out the pilot study to gauge the student response to the two response formats of the Critical and Creative Thinking Measure and to gauge the teacher response to the format of the teacher version of the Critical and Creative Thinking Measure.

Participants

To recruit schools, the research was publicised through various teacher forums in the South Australian Department for Education, and through word of mouth and personal contacts. For this study teachers

were required to have some confidence assessing the critical and creative thinking skills in the Australian Curriculum, have five years or more of teaching experience, and have been teaching their current year level for at least three years. Eleven schools initially indicated interest in the project, and teachers who indicated they were interested in taking part were sent a link to an online screening survey (Appendix F-5). Teachers who indicated that they had never used the C&CrT continuum and were not confident using the C&CrT continuum to assess critical and creative thinking skills were screened out of the research.

This study was initially designed to trial the tasks with students from Year 4 to Year 7, and in a range of schools. However, before the selection of participants was completed, the COVID-19 Pandemic arrived in Australia, and schools were advised to restrict access to essential staff only. The researcher was therefore unable to enter any schools between April and November 2020. Finally, in December 2020, two public schools in metropolitan Adelaide provided consent for data collection to occur, resulting in data being collected from three classes in the final two weeks of the school year – a Year 4/5 class from School 1 (*n* = 22) and two year 6/7 classes from School 2 (*n* = 40). Table 26 shows the breakdown of year and sex of the participants.

Approval was obtained from the University of South Australia's Human Research Ethics Committee; The Assistant Director, Department for Education Data Reporting and Analytics (ethics); and the principals of the schools, prior to research commencing. Information about the research was provided to each school principal, teacher and the families of student participants, and active informed consent was received from all participants. Information and consent forms are included in Appendix F-6.

Personal data about each student (date of birth, sex, English language level, diagnosed learning conditions) was provided by each school to enable students with English language levels lower than Year 4 to be screened out, and to identify potential patterns in the data.

Measures

The two measures presented to the students were Student Assessment Task 1 and Student Assessment Task 2, described above. The teachers were presented with the Teacher Task, also described above.

7.7. Data Collection Approach

The researcher presented the assessment task to the students in the three classes, as described above. The Assessment Task 1 measure or Assessment Task 2 measure was then presented to the students as shown in Figure 12. The choice of which measure was to be presented to which class was

made based on the number of students in the class. This ensured that a balance of the two measures was presented across the year groups as much as possible. All students in a class completed the same

Table 26

Participants in Study One, Pilot Study

Year Level	Assessment Task 1[a]	Assessment Task 2[b]
Year 4	10	0
Year 5	12	0
Year 6	11	16
Year 7	5	8
Male	15	13
Female	23	11

Note. [a]*n* = 38. [b]*n* = 24

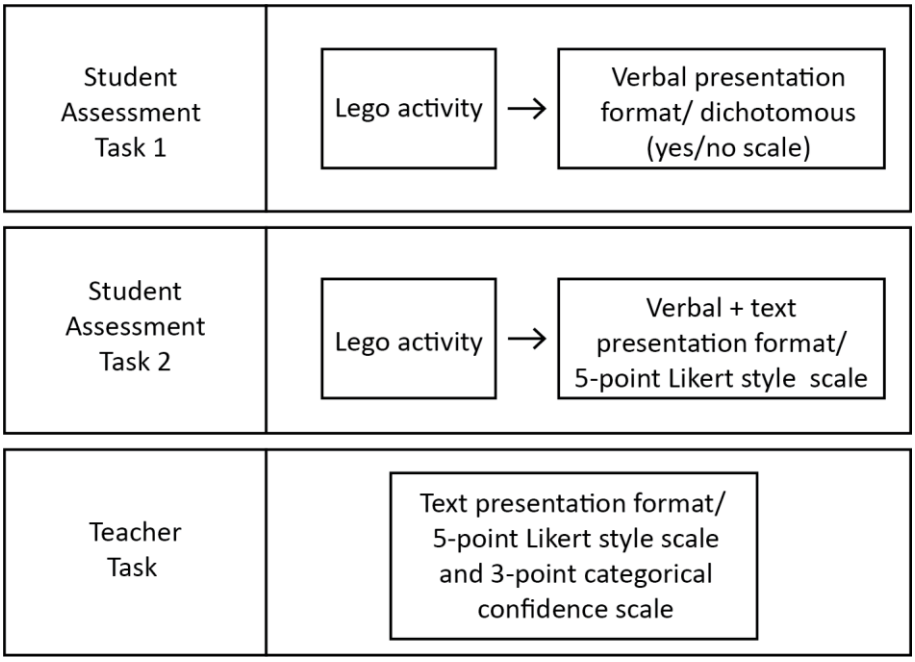


Figure 12

Data collection activities: Study One

task. Two classes, one year 4/5 and one year 6/7, completed Assessment Task 1 and one year 6/7 class completed Assessment Task 2. The two measures were completed by 62 students (Assessment Task 1: *n* = 38, Assessment Task 2: *n* = 24), providing sufficient data for observations to be made, although not for reliable statistical analysis. Due to the small number of classes in the study (*n* = 3) and because all students in any class were required to complete the same measure, Assessment Task 2 was not trialled with students in years 4 or 5.

Classroom teachers were provided with hard copies of the Teacher Task to provide their judgement of the skills used by each student in their class. These were mostly completed by the teachers during the assessment task presentation by the researcher. Where teachers had completed their tasks, these were collected by the researcher at the end of the session. Any outstanding copies were collected by the researcher later, on completion.

Method for Statistical Analysis

The small numbers of responses in this study increases the risk of Type I and Type II errors, limits generalizability of the results, risks unreliable results, and results that are not robust enough to draw meaningful conclusions. Despite these drawbacks, an analysis of the data is desirable to provide some indications of the effectiveness of the two assessment tasks. Therefore, statistical analysis was undertaken, and the following discussion takes account of the limitations caused by the small numbers of participants. All analysis of the data was carried out using IBM SPSS Statistics (Version 28). Preliminary assumption testing was conducted to check for normality. As this data was not normally distributed, the choice of data analysis techniques reflected this. Further as the responses to the measure in Assessment Task 1 were categorical (dichotomous yes/no response), modal responses were used to enable comparisons of the data from the measures in both assessment tasks.

The results of the Assessment Task 2 measure were in the same format as the Teacher Task responses (5-point Likert), therefore a comparison was carried out to compare the differences in the teacher scores and the student scores. A Shapiro-Wilk's test confirmed that the data was nonparametric ($p < .05$ in all cases). In such cases the Wilcoxon signed-rank test is often used as the nonparametric alternative, however as the distribution of differences were not symmetrically shaped the data was considered unsuitable to analyse with this test. The sign test was therefore run as the nonparametric alternative. This test examines the scores which are not an exact match (i.e., the non-tied results), comparing the number of negative and positive 'difference' scores. Although it is most common not to include the number of ties in the analysis (Gibbons & Chakraborti, 2011), the small number of

participants, and proportions of ties in this data affects how the results should be interpreted. The number of ties has therefore been included in the reporting of the results.

7.8. Results of Data Collection and Analysis

The results of the data collection and analysis are presented below in two parts. The first part describes the statistical results. These commence by comparing the modal responses for the two measures, and the two year-level groups (Year 4/5 and Year 6/7). This is followed by the results of the Student Assessment Task 2 measure, compared with the Teacher Task assessment of the students. The second part of the results reports on observations of how primary school students and their teachers responded to two different response formats of a self-report measure.

Statistical Analysis of Measures

Student Assessment Task 1. The measure used in Student Assessment Task 1 required the students to provide a response of 'Yes' or 'No', to the question type, 'Do you (do the skill)?' Modal responses are reported as the variables are categorical. The informal hypothesis was that the majority of students would give a 'yes' response to the skill which equated to their year level, and possibly to the levels expected for younger students. If some of the less complex levels of skill were subsumed or superseded by more complex levels, then a 'no' response might be observed at the lower levels. The number of the participating primary school students using skills at the Year 8 or Year 10 level was expected to be extremely low. As can be seen in Table 27, these hypotheses were not supported by the results from the measures in Assessment Task 1 or 2.

The results of the analysis of the Assessment Task 1 measure showed that the predominant response from Year 4/5 students to each statement was 'yes', except for two out of the 36 statements. With the Year 6/7 students the responses to two of the sub-elements followed the expected pattern of lower numbers of responses at the extreme ends of the level statements – 'Asking Questions' and 'Coming Up with Ideas'. The consistent 'Yes' response to 77% of the statements did not however support the hypothesis that the 'yes' responses would be clustered around the statement corresponding to the year level of the students.

It is not surprising that the option of 'always' was not chosen by students, however it was expected that 'never' and 'occasionally' would be chosen for the skills at the extreme ends of the level statements. This was not observed.

Table 27

Student responses to Assessment Tasks 1 and 2

Sub-element	Year Level Statement	Assessment Task 1		Assessment Task 2
		Year 4/5[a] Mode	Year 6/7[b] Mode	Year 6/7[c] Mode
Asking Questions	Foundation	Yes	Yes/ No	Often
	Year 2	Yes	Yes	Sometimes/ Often
	Year 4	Yes	Yes	Often
	Year 6	Yes	Yes	Often
	Year 8	Yes	Yes	Often
	Year 10	Yes	No	Often
Finding and Organising Information	Foundation	Yes	Yes	Often
	Year 2	Yes	Yes	Sometimes/ Often
	Year 4	Yes	Yes	Often
	Year 6	Yes	Yes	Often
	Year 8	Yes	Yes	Sometimes
	Year 10	Yes	Yes/ No	Sometimes
Connecting Information	Foundation	Yes	Yes	Often
	Year 2	Yes	Yes	Often
	Year 4	Yes	Yes/ No	Sometimes
	Year 6	Yes	Yes	Often
	Year 8	Yes	Yes	Sometimes/ Often
	Year 10	No	Yes/ No	Sometimes

Note. [a] *n* = 22. [b]*n* = 16. [c]*n* = 24

Table 27 (continued)

Sub-element	Year Level Statement	Assessment Task 1		Assessment Task 2
		Year 4/5[a] Mode	Year 6/7[b] Mode	Year 6/7[c] Mode
Creating New Ideas	Foundation	Yes	Yes	Often
	Year 2	Yes	Yes	Often
	Year 4	Yes	Yes	Often
	Year 6	Yes	Yes	Sometimes
	Year 8	Yes	Yes	Often
	Year 10	Yes	No	Sometimes
Coming Up with Ideas	Foundation	No	No	Sometimes
	Year 2	Yes	Yes	Often
	Year 4	Yes	Yes	Often
	Year 6	Yes	Yes	Often
	Year 8	Yes	Yes	Often
	Year 10	Yes	No	Sometimes
Thinking about a Plan	Foundation	Yes	Yes	Sometimes/ Often
	Year 2	Yes	Yes	Often
	Year 4	Yes	Yes	Often
	Year 6	Yes	Yes	Often
	Year 8	Yes	Yes	Often
	Year 10	Yes	Yes	Sometimes

Note. [a] $n = 22$. [b] $n = 16$. [c] $n = 24$

No patterns were observed in the responses from the students to indicate central tendency bias, supporting the random ordering of the statements. This appeared to have avoided a patterned response of choosing the 'middle' level statement rather than choosing the ones that best described the skills the students used. Likewise, the central response of 'sometimes' was not the modal response in the Likert style responses, indicating that students were intentionally choosing their response.

Student Assessment Task 2. The measure used in Student Assessment Task 2 required the students to indicate how often they used each skill, on a 5-point Likert style scale (never, occasionally, sometimes, often, always). Modal responses are reported here to provide a comparison to the results from the Student Assessment Task 1 measure. Further statistical analysis is described below. The informal hypothesis was that the majority of students would give an 'often', or 'always', response to the skill which equated to their year level, and a lower-level response to the extreme ends of the range of statements (the Foundation and Year 10 statements). Again, as can be seen in Table 27, this was not the case. The results of the Assessment Task 2 analysis showed that the predominant response from Year 6/7 students to each statement was either 'sometimes' or 'often'. 'Sometimes' was observed mostly as a response to the Foundation and Year 10 level questions, however, a response of 'sometimes' still indicates a greater level of use than expected.

A second statistical analysis, an exact sign test, was used to compare the differences in the teacher scores and the student scores from the Student Assessment Task 2 measure. The results of this analysis are shown in Table 28. In all circumstances where the differences between the student and teacher results indicated a significant difference in score, the teacher scored the student higher than they scored themselves. There were only four occasions where the ties for a statement were 50% or greater, i.e., where at least 50% of the students' responses for that statement agreed with the teacher responses. The three sub-elements which demonstrated the lowest levels of agreement between the teacher and students were 'Connecting Information' and 'Creating New Ideas'. These two sub-elements had statistically significant differences between the teacher and student responses for 100% of the statements, and only achieved ties for an average of 26% of each of the student responses. 'Thinking about a Plan' did not fare much better, with five out of the six statements indicating statistically significant differences between the teacher and student's responses. The interpretation of these results must all be viewed with caution however, as they are the result of only a small group of students (n = 24) and a single teacher.

As discussed in more detail below, despite the greater range of possible responses (5 levels in the measure in Assessment Task 2 compared with a dichotomous response in the Assessment Task 1

Table 28

Study One. Results of exact sign test, Assessment Task 2

Sub-element	Year Level Statement	Year 6 and 7[a]			
		Exact Sig. (2-sided test)	Ties	No. of Pos. Diff.[b]	No. of Neg. Diff.[c]
Asking Questions	F	1.000	11	6	7
	2	.022	11	11	2
	4	.815	6	10	8
	6	1.000	11	7	6
	8	.791	10	6	8
	10	1.000	9	7	7
Finding and Organising Information	F	1.000	13	5	6
	2	.549	13	7	4
	4	.791	10	8	6
	6	.180	15	7	2
	8	.021	8	13	3
	10	.096	6	13	5
Connecting Information	F	< .001	12	12	0
	2	.001	6	16	2
	4	< .001	6	18	0
	6	< .001	6	18	0
	8	< .001	5	17	1
	10	.003	4	17	3

Note. [a]*n* =24. [b]Number of times the teacher gave a higher score. [c]Number of times the student gave a higher score.

Table 28 (continued)

Sub-element	Year Level Statement	Year 6 and 7[a]			
		Exact Sig. (2-sided test)	Ties	No. of Pos. Diff.[b]	No. of Neg. Diff.[c]
Creating New Ideas	F	< .001	6	17	1
	2	< .001	5	17	2
	4	.019	5	15	4
	6	< .001	6	17	1
	8	< .001	5	17	2
	10	.007	9	13	2
Coming up with Ideas	F	.267	11	4	9
	2	.031	6	14	4
	4	.039	12	10	2
	6	.013	7	14	3
	8	.118	9	11	4
	10	.180	10	10	4
Thinking about a Plan	F	.096	6	13	5
	2	.049	7	13	4
	4	.004	8	14	2
	6	.013	10	12	2
	8	.022	11	11	2
	10	< .001	9	14	1

Note. [a]*n* =24. [b]Number of times the teacher gave a higher score. [c]Number of times the student gave a higher score.

measure), the information provided by the responses to the Assessment 2 measure did not enable effective conclusions to be drawn about which skills the students were using.

Teacher Assessment Task. The predominant response from the teachers for all six sub-elements indicated that they were very confident to score all six sub-elements. As explained above, this confidence may have been misplaced if the results of the statistical analysis are a true reflection of the results that can be expected from a wider population (limitations previously discussed must be taken into account here). Assuming that the students understood the statements and were responding truthfully, a level of less than 50% agreement would indicate that the teacher was only aware of the students' use of less than 50% of the specific skills described.

Response by Students to Characteristics of the Measures

Reading Level. During the lesson, students in the Year 4/5 class (Assessment Task 1) asked for some statements to be repeated and explained to them, indicating they may have had some difficulties understanding the language in the statements. Their teacher also commented that she thought the language was too complex for them. The reading level of a sample of statements from the measure were tested according to Fry (1977), rating at 5.8 sentences/ 100 words and 152 syllables/ 100 words. This equates to approximately a grade 9 level.

Timing of Lesson. The measures were successfully presented within a 50-minute lesson, providing sufficient time for all students to complete the task. In classes where the students were focussed and quick to respond to instructions the timing of the lesson worked well, providing sufficient opportunity to have 'brain break' activities and discussions to re-focus the student's thinking during the task. Where the students were slower to respond to instructions however, the 'brain break' activities had to be reduced to ensure the lesson fell within the 50-minute lesson time.

Inclusion of Lego Activity. The format of the assessment with the initial Lego activity was very effective at firstly breaking down initial barriers between the students and the researcher, and secondly at providing relevant discussion points for the use of skills. The students were all fully engaged in the Lego task, and students in each class were willing to respond to the questions designed to stimulate their thinking.

Response by Students to Presentation of the Measure

Students appeared to be engaged and to listen carefully during the presentations. This was judged by their compliance during the sessions and the rate of completion of the measures. There was no obvious

difference in response from the students to the two different styles of presentation (verbal only/ verbal + text), based on their reactions during the sessions.

Response by Students to the Response Formats

Student Assessment Task 1. Some students indicated verbally that for some questions they were not sure if they actually did the skill as described. They believed they understood what the statement meant and thought they might have applied the skill at that level, but they were not sure, so did not know if they should answer 'Yes' or 'No'. During Assessment Task 1 (verbal presentation only) students in the year 4/5 class occasionally asked for statements to be repeated for a third time.

Student Assessment Task 2. In Assessment Task 2 the students had a text copy of the statements being read, but they were asked to keep pace with the verbal presentation, and all students appeared to comply with this instruction.

Teacher Assessment Task. Verbal feedback from the teachers was that they found it harder to score their students on some of the skills than others, mostly because some skills were easier to observe (Asking Questions) than others (Connecting Information) in the classroom setting. This appears to have been supported by the results of the analysis which indicated closer results between the student and teacher responses for Asking Questions, and significantly different results for Connecting Information. The modal response from the teachers for all six sub-elements of question however, indicated that they were very confident to score all six sub-elements. Teachers also reported that there were some students who were much harder to score (the quiet ones). On the whole, however, they reported that the task was not difficult to complete. Teacher feedback indicated that it was reasonable to expect teachers to complete the measures for approximately 20–25 students during the 50-minute lesson.

7.9. Study One Discussion

This study was designed to develop and test a measure for assessing the student use of critical and creative thinking skills. The data collected included the qualitative data obtained from the results of the two measures, observations of students during the assessment process, and informal discussions held with students and teachers.

Despite the small number of participants in the study it was possible to draw conclusions from the data about the successful features and alterations required to improve the measure. These observations are presented here, followed by a table summarising the recommendations.

Although the response by students to both styles of presentation of the measure was similar during the lessons, the small number of participants did not enable adequate analysis to draw conclusions as to the most suitable method of presentation to use for the measure. A further study in which two presentation styles (verbal only and verbal + text) are utilised, with the same format of question and response in both tasks, would enable a comparison.

Response Type

The six statements in each sub-element category cover the skills expected to be demonstrated by students from Foundation to Year 10. It was expected that the following indicators could be used to signal whether the measure would be fit for purpose:

- Level of skills identified by students: At this point in their schooling (Year 4 to Year 7) the majority of students (demonstrated by the mode) in each class were expected to indicate that they were using skills up to and including those described for their current year level. If the less complex levels of the skills were subsumed within the more complex levels, there may have been a reduction observed in the frequency of use of lower-level skills.
- Variation across skill sub-elements: It was expected that students would indicate more skill acquisition in some sub-elements of skill (e.g., Asking Questions) than in others. This would depend on whether the students had developed particular skills by being explicitly taught them and whether they were provided with the opportunity to use them.

Interestingly, neither of these indicators showed the expected results. The reason for this may be due to the way the critical and creative thinking skills are used by students. This is discussed further below.

Teacher Task

Overall, teachers reported that completing the Teacher Task, responding about their students, was not too difficult. What was harder to judge, particularly due to the low number of participants (n = 3), was whether all teachers could be relied upon to provide a reliable assessment of each student's level of skill. Assuming that not all teachers can observe all things in all students (part of the justification for using self-report format), it seems reasonable to also assume that some teachers are better judges of their students than others.

Although the teachers had been screened to ensure they had some understanding of the skills in the C&CrT continuum and they were all reasonably confident assessing from it, it was not easy to estimate how accurate the teachers would be at judging their students' level of ability. The screening

process enabled the reasonable assumption that the teachers' levels of knowledge and skill in assessing students would be sufficient for the purposes of this research. Other factors which could not be accounted for were such things as:

- The teaching approach which could affect the types of opportunities students had to develop or demonstrate their skills;
- The personality of the teacher which could affect how open the teacher was to observing particular behaviours in students;
- The personalities of individual students which could affect how openly they exhibited particular behaviours in class;
- The culture of the school or classroom (not fully accounted for by the teaching style in use), which could affect how effectively the students demonstrated their skills to the teacher.

The intention of having the teachers and students both complete the same assessment of the student's skills was to provide a comparison against which to judge the accuracy of the student responses. This method effectively presents the teachers assessment as a 'gold standard' against which to compare or judge the student self-report responses. Although the results above indicate that teacher observations may not be a completely reliable method for assessing student skills, the limitations of the data collected must be kept in mind. Screening teachers for familiarity with the critical and creative thinking skills is clearly essential to provide the highest level of confidence in teacher judgements.

Student Assessment Task 1

In the Assessment Task 1 measure, the majority of students of all ages indicated they believed they already used the skills at almost all levels described, and in all the areas of skills covered. This is problematic as the measure is therefore not providing an opportunity for students to show if there has been any development of their skills at a later stage. Where responses to the Year 10 level statements and the Foundation level statements indicated students were not applying the skills as described, this was expected – the Year 10 level because these descriptions of skills were expected to be beyond the level of these students, and the Foundation level question because the majority of students were expected to have moved beyond the skill described at this level. Alternatively, the skill at this level may have been subsumed within the higher-level descriptions of the skill. The discrepancy between the Year 4/5 results (students thought they were using skills at all but one of the Year 10 level questions) and Year 6/7 results (students thought they were not using skills at five out of the six of the Year 10 level questions), indicates the possibility that the Year 4/5 students either were not sufficiently self-aware to know their own abilities, did not understand the question properly (the language may have been too complex), the description of the skill does not match what students actually do, or they did not

understand the skill that was being described to them ('they didn't know what they didn't know' – the Dunning-Kruger effect: Dunning, 2011). These issues all indicate that further review is required of the wording of the statements, and that the task may be too complex for the younger age group of students.

As there was some indecision reported by students as to their choice of 'Yes' or 'No' as the more accurate response, it would seem appropriate to include a third option – 'unsure' – for students to indicate when they are not sure if they do use the skill at this level. This will provide greater opportunities for students to demonstrate change over time in the level of skills they use.

Student Assessment Task 2

In the Assessment Task 2 measure the 5-point Likert style response enabled a greater range of responses, however the students still indicated a very optimistic view of their use of skills, by the predominant responses of either 'Sometimes' or 'Often'. The expectation here had been that students would use the skill described at their current year level most often, with reduced numbers of students using earlier year level skills and more advanced year level skills. This is most closely representative of a Guttman style scale, where multiple items within the index would be appropriate, and the point of transition from a positive response to a negative response would mark the highest level of achievement within that index. In this case, however, no clear transition point was indicated for any of the sub-elements. These results are not conclusive, indicating the need for further study to understand the nature of the skills as they are developed and applied by students.

Trying to understand these results led to a re-conceptualisation of the way different skills are used. Some skills seem to be used at all levels by everyone, no matter what stage of life or education they are at, depending on the topic to which they are being applied. Asking Questions is an example of this. If the topic is new to the student, they will need to ask beginner-level questions first, until their knowledge is such that more advanced-level questions become relevant. This process could potentially be repeated every time a new topic or subject is encountered, so the use of beginner-level questions (e.g., simple questions that only ask about objects or ideas in the immediate vicinity) would be a skill that would continue to be used, even after the skill of asking more advanced-level questions (e.g., asking about complicated issues and abstract theories) has been mastered. It would be realistic then to see the response to all levels of the Asking Questions statements to be 'Sometimes', once you have developed the full range of levels of this particular skill.

Biggs and Collis describe a model of learning which is reminiscent of this description of the use of skills, in the SOLO (Structure of the Observed Learning Outcome) taxonomy (Biggs, 1999; Biggs &

Collis, 1982, 1989; Collis & Biggs, 1986). The taxonomy contains two 'scales'. The first is described as 'Modes', or 'Kinds of Knowledge'. These describe a *global* framework with stages of development occurring as an individual grows, and their physical abilities and capacities develop. These age-linked descriptions are reminiscent of Piaget's stages, where each subsumes its predecessors (see Pegg, 2018; Pegg & Tall, 2005). In SOLO, contrastingly, previous Modes can be returned to throughout life as appropriate to the occasion. Primary school students are mostly operating within the concrete-symbolic stage, as described in Table 29. It should be noted however that the modes are not restricted to the age ranges given — these are indicators of the stage which best describes the highest level most people of that age are capable of operating within.

The second of the 'scales' is described as 'Levels of Understanding', and it is this scale which is of particular relevance here. The levels, explained in Table 30, describe a local framework of understanding as it develops and is applied to learning. Biggs and Collis (1989) describe five levels of student response which may be applied, demonstrating a learning cycle when a student moves to a new topic or reaches a more advanced stage of development. It is this characteristic that is in contrast to the description of skill development through the C&CrT continuum. Instead of the expectation that once students master one level of achievement they will be unlikely to revert to using a less sophisticated level, Biggs and Collis suggest that at times, although a student may be working on a more complex topic, they will revert back to a less complex form of thinking or responding than that which they were previously using.

Although the SOLO taxonomy is usually applied to the acquisition of knowledge, (e.g., Adeniji et al., 2022; Jimoyiannis, 2011; Pegg, 2010; Svensäter & Rohlin, 2023) applying it to the use of skills may help to explain some of the results observed in the students use of some of the skills. Figure 13 illustrates the possible learning cycle when new topics are introduced at increasing year levels. The skill of Asking Questions can be used as an example. Three levels of the skill of Asking Questions can be described to roughly compare with the SOLO Levels of Understanding. These three levels of question, drawn from the Critical and Creative Thinking Self-Report Measure described in this chapter, are shown as the generic question examples in the top row of Table 31, which demonstrate unistructural, multistructural and relational levels of questions. Below this (also in Table 31) are examples of questions that students might ask, relating to topics relevant for their year level. Ideally, over a series of lessons, students would be expected to build their understanding of the topic, which would activate the desire for deeper knowledge. This would require the use of more complex questioning. It is certainly reasonable however for the unistructural questions to be asked in each of these situations, as they enable a basic level of knowledge to be gained when a new topic is introduced. This is an example of where the C&CrT continuum fails to recognise the nature of this skill – that it does not work as a simple

Table 29

SOLO's developmental modes

Approximate age	Mode	Kinds of knowledge	Description of developmental stages
Up to 1 ½ years	Sensori-motor	tacit	Learning by giving a motor response to a sensory stimulus. Although this way of learning is also used in later life (e.g., learning a sporting skill) this is the only way children of this age can learn
1 ½ to 6 years	Ikonic	intuitive	This is the extended abstract version of the sensori-motor stage, where pictoral representations (ikons) are used to describe internalised information
6 to 16 years	Concrete-symbolic	Declarative/ procedural	Symbolic systems such as number and written language are used to symbolise events in the real world
16 years and above	Formal (1 and 2)	Theoretical/ conditional	Levels of abstraction demonstrated by theorising – questioning how things are and developing hypotheses

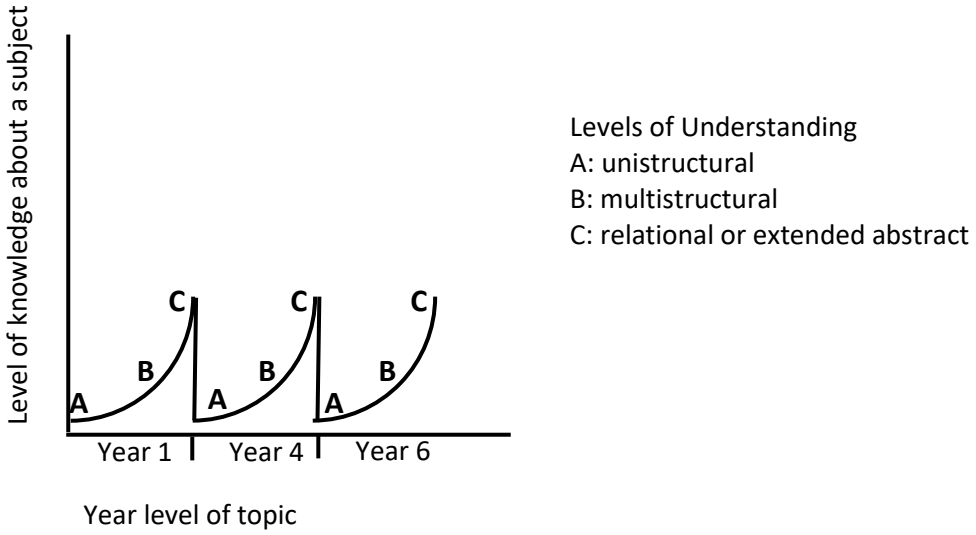
Note. Adapted from Biggs (1999); Biggs and Collis (1989)

Table 30

SOLO's levels of understanding

Level of Understanding	Description of Levels of Understanding
Prestructural	Little evidence of relevant learning is demonstrated
Unistructural	Responses deal with terminology, getting on track but little more. An individual idea may be grasped.
Multistructural	Multiple ideas are grasped, but presented as lists of facts and details. Effective connections are not made between them.
Relational	Relationships between ideas are recognised leading to a greater understanding of the bigger picture.
Extended Abstract	Understanding of a concept is taken beyond what is recognised at a *relational* level, showing new insight or perspectives.

Note. Adapted from Biggs (1999, pp. 38-39)



Note. Adapted from Biggs and Collis (1989)

Figure 13

Illustrating the effect of introducing new topics at different year levels

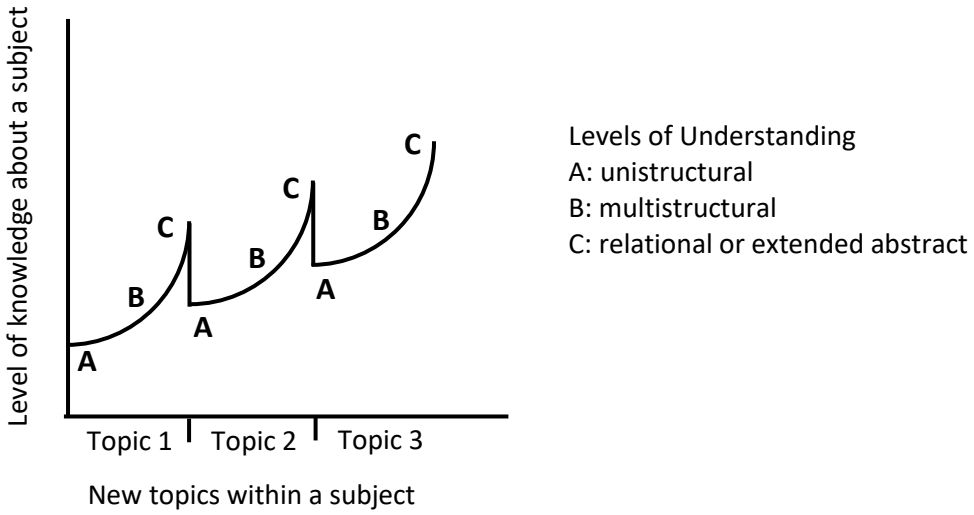
Table 31

Examples of SOLO questioning applied to year level topics

	Unistructural Question	Multistructural Question	Relational Question
Generic question example	I ask questions about things I'm interested in (Foundation level question)	I ask questions to understand more about things I notice in the world (Year 4 level question)	I ask questions to help me learn about complex or complicated issues (Year 8 level question)
Year 1 Topic on movement using toy cars	Can your car move on carpet like mine?	Why doesn't my car roll on grass, but it rolls on the pavement?	Why do the toy cars roll better on some surfaces than other surfaces?
Year 4 Topic on climate induced animal adaptations	What food does this desert lizard eat?	Do lizards in the desert eat the same foods as lizards in colder places?	What adaptations would be required for a desert lizard to survive if climate change causes the deserts to be colder or wetter?
Year 6 Topic on ancient civilizations	Did the Ancient Romans wear the same sort of clothes as us?	Did slaves wear the same type of clothes as their owners in Ancient Rome?	How did the laws about clothing affect the different social classes in Ancient Rome?

progression where only students at lower year levels use unistructural questions, and only students at higher year levels use relational questions.

This can also be demonstrated for students building on knowledge progressively as they build subject knowledge by working through a series of topics in a subject. Figure 14 illustrates how this might look. Table 32 demonstrates how the skill of Asking Questions might be demonstrated by a student asking questions in the subject of Space. Although the subject could be covered over the period of a year, with each topic building directly on the knowledge gained from the topic before it, as each new topic is broached, initial understanding is gained by the use of simple, unistructural questioning. This demonstrates again that these 'lower level' questions are not limited to use in the younger years of schooling, but are essential to use at all levels of education — indeed, of life — when new learning is happening.



Note. Adapted from Biggs and Collis (1989)

Figure 14

Illustrating the effect of introducing new topics building on previous knowledge

Table 32

Examples of SOLO questioning applied to a subject

	Unistructural Question	Multistructural Question	Relational Question
Generic question example	I ask questions about things I'm interested in (Foundation level question)	I ask questions to understand more about things I notice in the world (Year 4 level question)	I ask questions to help me learn about complex or complicated issues (Year 8 level question)
Topic 1 The solar system	How many planets are in our solar system?	Do other planets have a moon like ours?	What effect does the moon have on our life on earth?
Topic 2 Our galaxy	How many solar systems are in our galaxy?	How different are the suns in other solar systems in our galaxy, to our sun?	What effect do different types of sun have on the liveability of different planets for humans?
Topic 3 Space exploration	How far has a human travelled in space?	What conditions would a human need to survive on another planet?	What sort of life could exist on other planets which have a different type of sun and moon etc?

Rather than aligning these levels with the age of the student or the number of years the student has been at school, the structure of the SOLO taxonomy recognises that these levels of skill are demonstrated by students as they develop their knowledge in a particular area. Asking questions at a unistructural or multistructural level is not a reflection on a person's ability to ask more complex questions. It reflects on the level of knowledge they have developed in the particular subject area or topic in which they are working, and their ability to apply a more complex level of the skill to enable their knowledge and understanding to progress.

An expert in biochemistry, capable of asking extended abstract level questions in their field will still use unistructural level questions when faced with an unknown topic to investigate. If the statements relating to asking questions are understood in this way, i.e., that the use of the different levels of question relates to the depth of subject knowledge rather than to the level of skill the students is capable of applying, then two questions must be asked: 1) whether it is appropriate to align the descriptions of the skills against a school year level, and 2) whether all the skills operate in this way.

Other skills, such as Coming Up with Ideas, appear to be structured or perhaps behave differently, although this was not tested through the research described here. Coming Up with Ideas, for example, has levels of skill described at the beginner-level that you would expect to be superseded by more advanced-levels of the skill as the skill developed, e.g., the statement, 'When I've come up with an idea I just go with it', describes a skill at the Foundation level, demonstrating the impetuosity of inexperience. You could reasonably expect this to develop into the more advanced skill of, 'When I've come up with an idea I see if I can make it better by coming up with more ideas.' Once this more advanced level of skill has been experienced and mastered, you would not expect a person to return to working at the earlier level of skill except, perhaps, very occasionally. Unlike the skill of Asking Questions, the skill of Coming up with Ideas would therefore not be a good fit to the model of the SOLO taxonomy.

If, as described in the examples above, skills work in different ways then a one-size-fits-all continuum may not be the most appropriate way to describe them. Further research is required to understand the way the skills develop and are used, to enable teachers to more effectively support students to develop these essential skills.

This examination of the skills led to the conclusion that frequency of use was not what mattered. If the work you were engaged in did not require the use of more complex skills then you would not use them often. How frequently you use a skill at a certain level is perhaps more a reflection of the type of activity you are engaged in, than your ability to use a skill. What matters is whether or not

you are able to use a skill at the described level, not how often you use it. This would indicate that the response type associated with the Assessment Task 1 measure 'Do you do this?' would be more appropriate for the purposes of this research than that used in the Assessment Task 2 measure 'How often do you do this?'

7.10. Limitations of Study One

Due to restrictions placed on schools because of the COVID 19 pandemic, it was only possible for data to be collected from three classes, in the last two weeks of the school year of 2020. The number of participants was low (Task one n= 38, Task two n= 24), risking that the conclusions drawn from the statistical analyses are not robust, and Type I and II errors may occur. It was still possible, however, to make useful observations about the students and teachers and their response to the assessments.

The restricted testing of the measure other than through content validity can be seen as a limitation. However, this pilot study was focussed more on testing the presentation process of an assessment task, and identifying what should be measured, than on assuring validity at this stage. This issue of testing for validity and reliability will be addressed further in the studies following.

The results, discussions and conclusions drawn from the studies take these limitations into account.

7.11. Conclusions from Study One

Despite the small number of participants this pilot study enabled observations to be made about the way teachers and students responded to the assessments and the Critical and Creative Thinking Self-report measure, and about the structure of the C&CrT skills continuum. Together, this enabled a response to Research Question 2a, 'What should be measured to assess critical and creative thinking in the classroom during learning activities?' by identifying the most relevant critical and creative thinking skills, and the appropriate question format.

Although the two assessments were successfully run in the 50-minute time slots, some alterations to the Critical and Creative Thinking Measure emerged for consideration. Recommendations for further investigation are set out in Table 33.

Observations about the misaligned structure of the continuum, and the different ways the skills appear to be used, points strongly to the need for more research into how skills are used and developed. Examining some of the skills through a lens such as the SOLO taxonomy, may provide a more nuanced structure than the more conventional unistructural progression. Further research is also

required to enable teachers to more easily recognise the skills as they are applied at various levels. Version 9.0 of the Critical and Creative Thinking capability and the accompanying continuum (Australian Curriculum Assessment and Reporting Authority (ACARA), 2022), published after this research was completed, has reframed the sub-elements and re-worded the descriptions for clarity and simplicity. The focus on skill levels however has retained its connection to expectations at specific year levels, which may require re-evaluation with a deeper understanding of the way skills are developed and used by students.

Having carried out this pilot study to determine the most suitable form for the measure to take, the next stage in its development was to test the effectiveness of the two presentation styles for the task. In Study Two the measure, i.e. the 'scale' which enables the use of critical and creative thinking skills to be measured, is trialled with two forms of the assessment task, i.e. the task in which the 'scale is applied'. This assessment task includes the Lego activity and the presentation of the measure to the students. The comparison of two presentation styles is described in Study Two (Chapter 8).

Table 33

Recommendations for further investigation

Category	Issues	Suggestions for consideration
Lesson format	Timing is tight for a 50-minute lesson	• Consider reducing length of measure e.g., number of statements, to ensure sufficient time for discussion with students
Presentation format	Effective comparison has not occurred	• Test Assessment Task 1 with both forms of presentation (Verbal only and Verbal + text) with larger sample size to determine the most suitable presentation style. (Study Two)
Response type	Students believe they are already using all skills at almost all levels.	• Ensure the skills described match skills students actually use • Consider if the skills should be described at additional, more complex levels (beyond the Year 10 level) • Expand response type to provide more information e.g., in Assessment Task 1: no/unsure/yes • Frequency type question does not provide useful information
	Type of question asked ('Do you do this?' or 'How often do you do this?')	• Preferred question type is 'Do you do this?' (Assessment Task 1) • Carry out study to compare presentation format (Verbal/ Verbal + text) using Assessment Task 1 style question
Teacher/student agreement	Acceptable level of agreement is unknown	• Investigate method for determining acceptable level of agreement between teachers and their students
Connecting with skills	Effectively highlight connections between the statements and actual use of skills for students	• Extend these opportunities if time allows
Language Review	Language is still too complex for Year 4 and Year 5 students	• Re-assess the language in all statements, for understanding, clarity and reading level

Chapter 8.

Study Two: Two Alternatives for a Novel Assessment of Critical and Creative Thinking

8.1. Introduction

In Chapter 7, the initial development phase was described for creating a novel assessment which included the Critical and Creative Thinking Measure. The novel assessment, which includes a Lego activity, the Critical and Creative Thinking Measure and a presentation for students, was designed to enable the measurement of critical and creative thinking skills used during their everyday learning, by primary school students. The measure provides the means to begin to answer the overarching question (now revised) asked in Chapter 1: What is the impact of an industry partner in STEM PBL, on the application of critical and creative thinking skills in primary school students?

The pilot study described in Chapter 7 explored *what* was to be measured. Although the Critical and Creative Thinking learning (C&CrT) continuum does not present an accurate description of the development of the skills, a representative group of the skills described in the C&CrT continuum was used for the measure that was developed. Focussing on these skills specifically is essential because they represent the skills that are being taught in schools in Australia. If these are the skills being taught, then it follows that these will be skills that are more likely to be observed in students. The recommendations from Study One included that the manner in which the Critical and Creative Thinking Self-report measure is presented to participants should be examined, as students engaged in the verbal only presentation format appeared to have some difficulties understanding the questions. To further refine the Critical and Creative Thinking measure, this study therefore examines *how* the skills should be measured, through Research Question RQ 2b 'How does the method of communicating a self-report measure to students affect the accuracy of their responses?'. This second pilot study seeks to address the question of whether there is a difference in having the measure presented in two modes, i.e., in both verbal and text format simultaneously, versus in a single mode, i.e., in verbal format alone. While no significant difference was found in the effectiveness of the two presentation formats, there was a low level of agreement between the perceived use of skills reported by the teachers and their students.

8.2. Impact of Presentation Formats

During the pilot study described in Chapter 7 it was observed that the younger group of students who were engaged in Assessment Task 1 (verbal only format) required statements to be repeated and explained. The students engaged in Assessment Task 2 (verbal + text format) were all from the older age group and did not require the same level of support to understand the statements.

Teachers in mainstream classrooms are encouraged to use multiple modes of presentation simultaneously e.g. verbal and written, to support the perceived needs of different students in their class, particularly where the class includes students with learning complexities such as Autism Spectrum and Auditory Processing disorders (Main, n.d.; Montali & Lewandowski, 1996). In my experience as a classroom teacher, this is not unusual in mainstream classrooms in public schools in South Australia. Although the cognitive processes used in learning and testing differ, understanding the theories of multimedia learning helps us to understand the likely effects of using multiple modes of media in testing (Lindner, 2022). It is possible, that rather than supporting students, the use of two forms of presentation simultaneously may cause added cognitive load, impeding comprehension of the assessment instructions.

To identify whether having the text of the statement available for students to read while it is read aloud to them is advantageous, it was determined to test the preferred form of the Critical and Creative Thinking Measure (Assessment Task 1) using two alternative presentation formats. If there was little or no discernible difference in the responses of students to the different presentation formats, then it was most likely the level of comprehension of the students in Study One that was responsible for affecting their ability to engage with the task, not the way in which the task was presented to them. If there was a significant difference between the responses of the students to the two formats, then the most effective presentation format would be utilised for Study 3 (described in Ch 9).

The effects of using different formats for the presentation of assessments was therefore investigated initially through a conceptual exploration and theoretical synthesis, prior to being tested through empirical research methods.

8.3. Theories of Multimedia Learning

The idea of presenting information simultaneously through the two modes of text and voice stems from current school practice and research on multimedia learning. The basis of the theory of multimedia learning is that learning is enhanced when information is presented in multiple modes, including words, pictures and audio (Arslan, 2012; Mayer, 1997, 2022).

Several theories and principles describe the possible effects of multiple modes of presentation on student learning, including Mayer's Cognitive Theory of Multimedia Learning (Mayer, 2005a), the modality principle, (Low & Sweller, 2005; Mayer, 2005b), Sweller's theory of cognitive load (Sweller, 2005a), the transient information principle (Jiang & Sweller, 2022) the redundancy principle (Sweller, 2005b), and Paivio's theory of dual coding (Clark & Paivio, 1991; Sadoski & Paivio, 2013). Recent research has explored how these concepts can be applied to the use of multimedia in testing (Kirschner et al., 2017; Lindner, 2022). The effect of using multiple modes in tests using self-report assessment tasks (the form taken for the assessment task in this study) does not appear to have been the focus of any research to date and is therefore a novel contribution offered by this thesis.

Three main concepts underpin the overall idea of multimedia learning. The first, cognitive load theory (Sweller, 2005a) provides an understanding of how information is stored and used in the brain. The second, the multimedia principle, theorises that 'People learn more deeply from words and pictures than from words alone' (Mayer, 2005a, p. 31). The third theory, the dual coding theory (Clark & Paivio, 1991), helps us to understand how cognitive load theory interacts with the multimedia principle to explain why receiving information in more than one mode can be an effective aid to learning.

Cognitive load theory (CLT) has had a significant influence on theories of multimedia learning (Ginns, 2005), as it suggests that there are two types of memory – long-term memory which has an infinite capacity and working memory which has a very limited capacity (Mousavi et al., 1995). Information that is being actively used is held in the working memory, but the limited capacity of the working memory significantly affects how much information we can have in there at any time. The theory develops the idea that learning requires us to transfer information between our working and long-term memories as new memories are formed and then stored for future reference. Critically, CLT suggests that when we reduce the load on working memory, we learn better, particularly if the combination of intrinsic and extraneous load[23] are close to the capacity of working memory (Paas & Sweller, 2022). CLT and the multimedia principle have historically been applied to research on instructional design where learners are balancing information high in element interactivity (e.g., pictures or diagrams presented with words in written or verbal form: Sweller, 2005a). I have observed that the language of 'cognitive load' and 'working memory' have, however,

[23] The intrinsic load reflects the intellectual complexity of the information. The extraneous load is affected by the way the information is presented. Cerpa, N., Chandler, P., & Sweller, J. (1996). Some conditions under which integrated computer-based training software can facilitate learning. *Journal of Educational Computing Research*, *15*(4), 345-367.

entered the classroom and are used (perhaps erroneously) to describe the difficulties students have in recalling previous learning and applying it in new situations. An examination of this theory in the application of assessments therefore has implications for practical application.

Paivio's theory of dual coding suggests that some forms of information can be processed in separate modes, enabling simultaneous processing. This reduces the load on working memory to facilitate learning. In effect, this enables multiple channels of working memory to be activated thereby increasing the amount of working memory available, which facilitates learning by spreading the load.

Both Sweller's and Mayer's work on multimedia learning seek to exploit this feature of dual coding, however it is at this point that the various theories begin to diverge, as Sweller describes the two channels as based on presentation modes, and Mayer and Paivio describe the two channels based on sensory modes (Mayer, 2005a).

Proponents of the presentation mode approach (including Mousavi, Kalyuga and Sweller) classify the two channels by the mode in which the information is *presented* to the subject i.e. as having either verbal (e.g., speech) or non-verbal formats (e.g., images) (Mousavi et al., 1995). Proponents of the sensory modality on the other hand, focus on the senses used to *interpret* the information, i.e., on auditory (through the ears) or visual channels (through the eyes). (Mayer, 2005a). These differences may help explain some of the contradictory results seen in research on the various uses of text, image and speech in learning materials (Cuevas & Dawson, 2018; Moreno & Mayer, 2002; Sombatteera & Kalyuga, 2012). For the research described in this study though, the significant point of difference between the theories is in the way they classify printed text.

In the *presentation mode* approach, written text is processed as internalised speech in the verbal channel along with spoken words. If this is the case, then the presentation of written text and speech simultaneously would compete for working memory capacity. This would potentially bring into play the redundancy effect (Kalyuga et al., 1999; Kalyuga & Sweller, 2014, 2021) where redundant information causes an increased cognitive load that inhibits learning by overloading the working memory. The format of presentation of the assessment task in this study including both text and the identical verbal presentation, would potentially (according to this theory) cause a greater cognitive load on the students, which would have the further potential to reduce their ability to provide considered responses to the statements in the assessment task.

In the *sensory modality* approach, written text is processed in the visual channel, thus written text and spoken text can be processed separately, with two channels of working memory

operating simultaneously. This activates the dual coding effect described by Paivio, where the opportunity for the brain to make meaning of the words is increased without increasing the load on working memory. According to this theory, the presentation format of the assessment task in this study including both written text and the identical verbal presentation would potentially increase their ability to provide considered responses to the statements in the assessment task, without increasing the load on their working memory. It is expected that by enabling more considered responses from the students their responses would be more accurate and, assuming the teacher assessments are accurate, this would increase the similarity between the teacher assessment and the student's self-report.

As a result of the conceptual study described above, study two will test whether students assessed with the self-report Critical and Creative Thinking Measure under the condition of simultaneous text and verbal presentation (Presentation Format 2), have a stronger correlation with the teacher's assessment of their critical and creative thinking skills, than students assessed under the condition of verbal only presentation (Presentation Format 1). The working hypothesis is that the multimedia presentation format will increase the accuracy of student self-reporting, as they are anticipated to be able to operate in sensory mode, as described above.

8.4. Study Two Research Design and Methods

Validity and Reliability

As described in Chapter 7, during its development in Study One the Critical and Creative Thinking Measure underwent one assessment for validity — content validity. The same form of validity testing was repeated in Step One of Study Two through use of an expert panel, during the process of re-assessing the statements. Other forms of validity and reliability testing could not be applied due to factors such as the small number of participants and the lack of similar measures with which to make comparisons. No further tests of validity or reliability were suitable therefore to be carried out in the study described in this chapter.

Measures

The preferred measure developed in Study One (Assessment Task One) formed the basis for the Critical and Creative Thinking Self-report measure used in this study. The original task consisted of six sets of six statements (36 statements in total), each set relating to a specific sub-element in the C&CrT continuum (e.g. Asking Questions). The statements relate directly to the levels described in the C&CrT continuum at different year level expectations, e.g. the statement 'I ask questions to compare information about things around me' relates to an expectation of students at a Year 2 level,

while the statement 'I ask questions to understand why and how things happen' relates to an expectation of students at a Year 6 level. A full set of the questions is provided in Table 24. The response format is dichotomous with the two options 'yes' and 'no'.

Procedure

As shown in Figure 15, the design of Study Two consists of three steps:

- Step one: The statements and format of the Critical and Creative Thinking measure (Assessment Task 1) were reviewed. The findings from Study One were reviewed to determine which statements would be appropriate and achievable, (i.e. review of statement wording and response type), and to select which formats would be used to present the Critical and Creative Thinking measure to the students.
- Step two: Data collection – testing of presentation formats with teachers (*n* = 24) and students (*n* = 407) from 24 classes in years 4 to 7, across ten public and private schools in South Australia. This study is a between-groups, quasi-experimental design, as a true random assignment of the participants is not possible due to the nature of the participants (classroom groups).

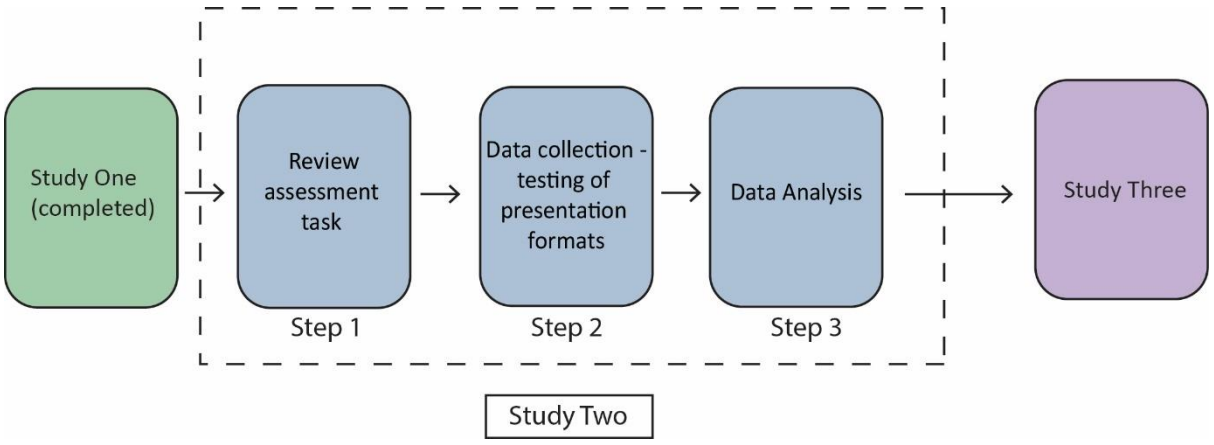


Figure 15

Research design: Study Two

- Step three: Data analysis. To identify the preferred presentation format, subject to the characteristics of the data, a measure of interrater agreement was used. The presence of outliers was assessed and outliers removed. The levels of agreement between teacher and student were then compared to the two presentation formats. Presentation Format 1 was identified as the preferred measure for use in Study Three.

Step One: Review Assessment Task

The recommendations made following the Study One pilot study were first assessed for feasibility. As shown in Table 34 the recommendations fell broadly into three categories – the need to further review the statements in the task (Statement Review), recommendations about the type of response that was most informative (Response Type), and a recommendation to investigate the levels of teacher/student agreement to enable the data to be effectively analysed (Agreement).

The review identified repetition and overlap in statements in the categories 'Creating New Ideas', 'Coming up with Ideas' and 'Thinking about a Plan'. These three categories were reduced to two, and statements were re-worded for clarity. The re-worded statements, with changes in wording indicated in bold, are shown in Table 35. These changes in wording were reviewed and approved by a panel of three teachers. This review addressed the first three recommendations from Table 34. The final recommendation in relation to statement review was to consider if the skills should be described at additional, more complex levels (beyond the year 10 level). This was deemed out of scope for this research project which is focussed on the skills described in the AC continuum. Adding an additional level would take the assessment beyond what is described in the AC continuum.

The second category of recommendation arising from Study One related to the response type. The recommendation was to expand the response type in Assessment Task 1 to provide more information. The preferred assessment task had a dichotomous (yes/no) response. Incorporating a third response – unsure, would enable a more nuanced response from students. This option was therefore added, and it was emphasised in the instructions to both students and teachers that *Yes* meant: 'Yes, I do this thing'; *No* meant: 'I do not do this thing'; *Unsure* meant: 'I know what you mean but I am unsure if I do it.'

Table 34

Recommendations from Study One

Category	Recommendations
Statement Review	• Reduce length of task e.g., number of statements, to ensure sufficient time for discussion with students • Re-assess the language in all statements, for understanding and clarity • Ensure the skills described match skills students actually use. • Consider if the skills should be described at additional, more complex levels (beyond the Year 10 level)
Response type	• Expand response type in Assessment Task 1 to provide more information e.g. no/unsure/yes
Agreement	• Investigate method for determining acceptable level of agreement between teachers and their students

Following the changes in wording and statements, the two presentation formats were developed (Appendices G-2 and G-3). The appearance remained essentially unchanged from the measures used in Study One. Presentation Format 1 (verbal presentation only) was a single page, with numbered response boxes for the student to indicate their response. It was also given the addition of a visual example of how the three different response options were to be indicated (a cross for *No*, tick for *Yes*, circle around the box for *Unsure*). Presentation format 2 (verbal and written presentation) was a four-page document, with the 30 statements grouped by category and presented on a grid with headings indicating the column for the response. Students were asked to put a tick or cross in the appropriate column to indicate their response. An example was also provided with each presentation format, consisting of 6 boxes in a row for Presentation Format 1 and 6 statements laid out in the same manner as the body of the task for Presentation Format 2.

The teacher task was similar in format to the student task Presentation Format 2, with modified response columns. The columns were headed 'Score' and 'Confidence' and teachers were instructed to score the students in the following way: 1 = *No, I have never seen them do this*; 2 = *I am unsure if they do this*; 3 = *Yes, I have seen them do this*. In the confidence column they were asked to

Table 35

Reworded statements in assessment task

Year	Study One wording	Study Two wording
	Asking Questions	
2	I ask questions to compare information about things around me	I ask questions **about things I'm interested in**
4	I ask questions to understand about things I notice in the world	I ask questions to understand **why and how things happen**
6	I ask questions to understand why and how things happen	I ask questions to understand **more about things I notice in the wider world**
8	I ask questions to help me learn about more complex (complicated) issues	I ask questions to help me learn about more complex **or** complicated issues
	Finding and organising information	
2	I find and organise information by identifying and ordering it	**When I find and organise information I identify the parts I need and sort them in order of importance**
4	I find and organise information by collecting, comparing and putting it into categories	**When I find and organise information I put it into categories, like 'facts' and 'opinions'**
6	I find and organise information by choosing the relevant part to use and knowing which are the facts	**When I find and organise information I use more than one source** and choose the relevant parts to use
10	I find and organise information by comparing it against criteria I have developed to make sure that it is reliable	**When I find and organise information I** compare it against criteria I have developed to make sure that the **information** is reliable

Note. Changed statements are indicated in bold.

Table 35 (continued)

Year	Study One wording	Study Two wording
	Connecting information	
F	I can connect information about different things	**I make connections between information I learned in different places (e.g. at home and at school)**
2	I can connect information I learned from doing something else, and use it to come up with a new idea	**I make connections between information I learned at different times (e.g. lessons last year or last week, and a lesson today)**
4	I can connect information from different places, even when it is about something completely different	**I make connections between information I learned in different places or times and use this to help me understand things**
6	I can connect information from different places, and use it to help me learn and understand something completely different	**I make connections between information I learned in different subjects, and use this to help me understand things in new ways.**
8	I can connect information I learned somewhere else, and I can explain why I choose to use it when I'm working in a different subject area	**I make connections between information I learned in different subjects. I can explain the connections and why they are useful.**
10	I can connect information from different places, I plan to do it, and I can explain why it improves my ideas	**I make connections between information I learned in different subjects, places and times. I can explain the connections and why they improve my ideas.**
	Creating New Ideas + Coming up with Ideas	***Creating New Ideas***
F	I create new ideas by looking closely at things around me to give me ideas + When I've come up with an idea I just go with it	I create new ideas by **copying things I'm familiar with**
2	I create new ideas by adding to things I already know + When I've come up with an idea I see if I can make it better by coming up with more ideas	I create new ideas by adding to things I already know

Note. Changed statements are indicated in bold.

Table 35 (continued)

Year	Study One wording	Study Two wording
	Creating New Ideas + Coming up with Ideas	***Creating New Ideas***
4	I create new ideas by coming up with more than one idea and then choosing the best + When I've come up with an idea I compare it with other people's ideas and sometimes add some of theirs to mine to make it better	I create new ideas by coming up with more than one idea and then choosing the best
6	I create new ideas by combining ideas from more than one place, and in more than one way + When I've come up with an idea I choose the best option after I've carefully tested a few	I create new ideas by combining ideas **that come from different places, and using them in new and different ways**
8	I create new ideas by seeing how my ideas are related to existing ideas, and making them more imaginative	I create new ideas by seeing how my ideas are related to existing ideas, **and changing them** to make them more imaginative
	Thinking about a Plan + Coming up with Ideas	***Evaluating a plan***
8	When I've made a plan to do something I can see where more information is needed, or where parts of the plan may need more thought + When I've come up with an idea I can describe what the problems might be, and how I might change my idea to avoid them	**When I've come up with a plan or an idea I am able to analyse my plan to see where parts of the plan may need more thought**
10	When I've made a plan to do something I can discuss how reliable my information is and tell you where there might be mistakes in my reasoning + When I've come up with an idea I test options, analyse solutions and justify choices	**When I've come up with a plan or an idea I am able to analyse my plan to tell you where there might be mistakes in my reasoning or in my information**

Note. Changed statements are indicated in bold

indicate for each category (e.g., 'Asking Questions') how confident they were to score the skill for the student. Their response options were L = *a little*, M = *moderately*, V = *very*.

The third category of task for revision was to investigate a method of determining the level of agreement between teachers and their students. Agreement may be calculated by using an interrater agreement, which may be quantified by assessing the correlation between two different assessments. Fleiss' Kappa is a statistical measure used to assess the level of agreement among multiple raters when evaluating categorical data. In this study, only the student themselves and the classroom teacher were able to provide a score for each student, classifying them as 'non-unique judges'. The use of multiple raters, or 'non-unique judges', is an unusual circumstance with testing, as multiple judges are often used to provide a score for all participants (Laerd Statistics, 2017b). In addition, each student was only judged twice (once by themselves and once by the teacher). For this reason Fleiss' Kappa was used, as described below.

Step Two: Data Collection

The data collection process commenced with recruitment of participants through word of mouth, presentations at conferences and the publication of an article in Teacher Magazine (Gaardboe, 2020). Potential participants were then screened for experience and confidence using critical and creative thinking by use of an online questionnaire.

The participants were teachers (*N* = 24) from a mixture of public and private primary schools in South Australia, and students (*N* = 407) from their classes. Teachers were screened *out* of the data collection if they indicated in the screening survey that they had no experience using the Critical and Creative Thinking continuum to assess students AND were not confident using the Critical and Creative Thinking continuum to assess students AND were the only teacher in the school who was interested in taking part in the research. The expectation was that teachers with greater experience both in teaching and with using the AC continuum would be more accurate in assessing their students against these skills. As this had not been previously tested, teachers with less experience and confidence were *not* screened out where they were in a school with other teachers taking part in the research. It was expected that any teachers who did not have good agreement with their students about their use of skills would be screened out in the data analysis phase as outliers. However, there was not a good level of agreement and therefore outliers were the teachers with the highest level of agreement. The numbers of teachers who were included in the study, categorised by experience and confidence in using the C&CrT continuum, is shown in Table 36.

Table 36

Teacher participants in Study Two

Variable	Category of response			*N*
	First year teaching	2 – 4 years teaching	5 or more years teaching	
Years of experience	1	3	20	24
Experience teaching current year level	4	3	17	24
Experience using the C&CrT continuum to assess Critical and Creative Thinking	I have not used it	I am unsure if I have used it	I have used it	
	9	2	13	24
Confidence using the C&CrT continuum to assess Critical and Creative Thinking	Not confident	Somewhat confident	Reasonably confident	
	5	14	5	24

Table 37

Student participants in Study Two

Year group	Presentation Format 1		Presentation Format 2		Total
	Male	Female	Male	Female	
Year 4	19 (4.7%)	27 (6.6%)	18 (4.4%)	44 (10.8%)	108 (26.5%)
Year 5	3 (0.7%)	8 (2%)	11 (2.7%)	13 (3.2%)	35 (8.6%)
Year 6	29 (7.1%)	57 (14%)	39 (9.6%)	33 (8.1%)	158 (38.8%)
Year 7	23 (5.7%)	18 (4.4%)	32 (7.9%)	33 (8.1%)	106 (26.0%)
Total	74 (18.2%)	110 (27%)	100 (24.6%)	123 (30.2%)	407 (100%)

Student participants were all members of the class taught by a participating teacher who consented to participate in the research. The students ranged from Year 4 to Year 7 and were all working at an English Language level of Year 4 or above. The break-down of students by assessment task, sex and year level can be seen in Table 37.

Prior to research commencing, approval was obtained from the University of South Australia's Human Research Ethics Committee; The Assistant Director, Department for Education, Data Reporting and Analytics (ethics); and the principals of all schools involved in the research.

Information about the research was provided to each school principal, teacher and the families of student participants, and active informed consent was received from all participants. Personal data about each student (date of birth, sex, English language level, diagnosed learning complexities) was provided by each school to enable screening of students with low English language levels for exclusion, and to identify potential patterns in the data.

The two presentation formats of the assessment task – Presentation Format 1 and Presentation Format 2 — were selected for presentation to the classes based on the number of students in the class and number of assessments already completed. In this way it was possible to obtain a reasonable balance of the two assessment tasks across the year groups. Classroom teachers

were provided with copies of the Assessment Task for Teachers (Appendix G-4) to complete for each student in their class in hard copy or digital format (their preference).

The data collection lesson followed the same format developed in Study One. The lesson commenced with the researcher introducing herself, discussing research in general terms, and explaining the purpose of collecting this data (Script in Appendix G-1). This was followed by a short Lego building activity, as in Study One in which pieces of a model were missing. The Lego activity was presented by the researcher as a challenge to see which of the five different models would be completed first. Working in groups of two or three, the students were required to use strategies to solve the problem of an incomplete model and how to 'do the best job they could' to build the model using alternative pieces. Students were timed, and the completion times were recorded on the board. A brief discussion followed about the patterns that could be seen in the data. The Lego activity was then used as a prompt to discuss topics in the assessment tasks e.g. what sort of questions one might ask at different times; whether knowledge and experience from an earlier experience was helpful in solving their current problem with the Lego, etc. This activity provided opportunities to discuss the skills used and demonstrate some of the language used in the Critical and Creative Thinking measure. It also acted as an ice-breaker between the students and the researcher, who was previously unknown to most of the students. The measure was then presented to the students with the researcher reading each statement twice – once for the students to hear the whole statement and a second time to confirm what had been said. The students then indicated their response in the numbered boxes on their answer sheet.

A digital presentation was used to re-enforce the instructions provided by the researcher, and multiple prompts were given for students to ask questions throughout the lesson. Again, as in Study One, 'Brain breaks' in the form of stretches and physical games were also inserted into the lesson to ensure the students had opportunities to refresh themselves and remain stimulated (Gernes, 2021).

Teachers were asked to undertake the teacher task during the lesson, reporting on their beliefs about the skills they had observed each student in their class using in the past. These were not limited to observations based on the activities during the assessment task. Although teachers were encouraged to complete their assessment tasks within the 50-minute lesson time, some were unable or unwilling to do this. Most teachers were able to complete measures for up to 22 students within the lesson time. Some of those that took longer reported that they took considerably longer, and some reported having considerable difficulty deciding on a suitable score for their students.

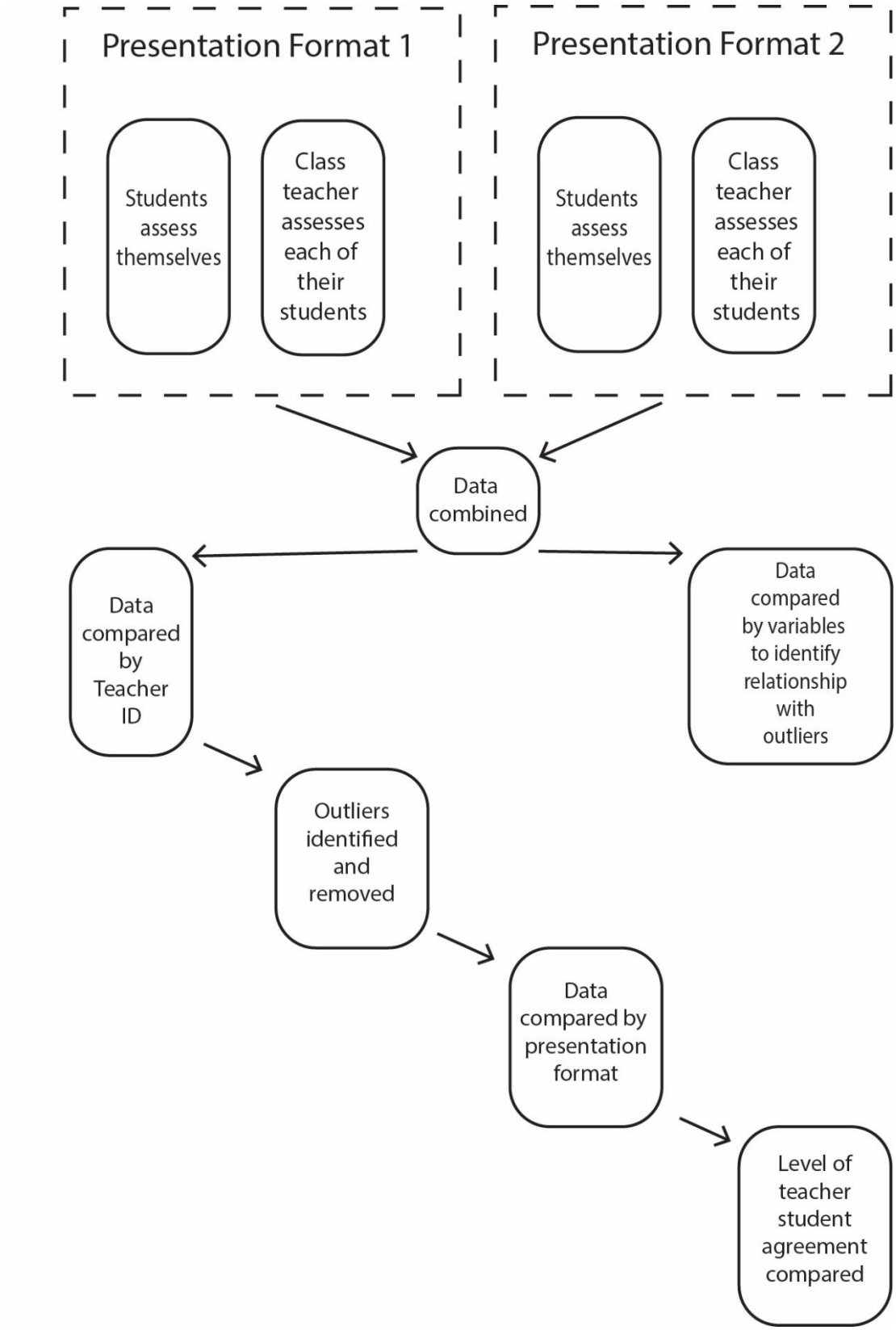


Figure 16

Process of data analysis

Where teachers had completed their tasks, these were collected by the researcher at the end of the session. If the teacher was using the digital format and had not completed all students, the teacher emailed the results to the researcher on completion. Where the teacher was completing paper copies, all completed copies were taken by the researcher, and any outstanding copies were collected by the researcher later, on completion.

Step Three: Data Analysis

To address how presenting a self-report measure in the two modes of text and verbal presentation, versus verbal only, affects the correlation of student and teacher assessments of the student, the student self-assessments were compared with the teacher assessments of the students for the two modes of presentation. Exact agreement between observers of an activity might be expected in some circumstances – such as judging (objectively) whether a goal has been scored in football. Other circumstances are less clear cut, such as judging (subjectively) whether a player has been tackled with excessive force. Identifying evidence of student skill use combines objective and subjective judgements. Although exact agreement was not expected between each student and their teacher, it was expected that a reasonable rate of agreement would be observed.

It was not easy to gauge what level of agreement between the teacher and their students would be 'reasonable'. Studies vary widely, from reporting a high positive correspondence (Südkamp et al., 2012) to reports of 75% of teacher judgements being inaccurate (Bates & Nettelbeck, 2001), even reports of teacher judgement being negatively correlated with student test scores (Machts et al., 2016). Certainly, teacher judgement appears more accurate when the task is a performance test, e.g. maths test, or oral presentation (Hung et al., 2016) than when the concept being measured is less specific, e.g. creativity (Machts et al., 2016; Spinath, 2005) or anything that cannot be easily observed (Gagné & St Père, 2001). This was confirmed in Gralewski and Karwowski's meta-analysis for example (2019), which reported that teacher ratings of student creativity are poorly related to the psychometric measures of divergent thinking. Importantly, this meta-analysis confirmed Machts (2016) earlier findings that the schooling level of the students did not affect the rate of agreement. This ran contrary to the findings of Rost (2009), who claimed teacher judgement was more accurate when assessing primary school students than secondary and tertiary students. Studies such as that by Chang and colleagues (Chang et al., 2012), reporting on the consistent results of teacher assessments for secondary students in a web-based portfolio assessment environment were considered too different to have relevance to this study.

As this study concerned cognitive skills which cannot be easily observed, based on the literature described above, it was recognised that there was risk of a low correlation occurring. However, the teachers in this study were primary school teachers who had been observing and assessing their student for almost a full school year, so they were expected to have a high level of knowledge about their students. Furthermore, 79% of the teachers in the study reported a moderate to high level of confidence with using the C&CrT continuum to assess students, therefore the level of agreement was conservatively expected to be between 25% to 65%. The teachers who reported a lack of familiarity or confidence using the C&CrT continuum were expected to register as outliers.

Data Analysis Approach

As the scale in the measure being used in this study only provides three levels – no, unsure, or yes – exact agreement is desirable. Fleiss' Kappa was used to calculate the interrater agreement as it is suitable for use with non-parametric data. The more commonly used Cohen's Kappa coefficient was not used, as this method relies on the use of unique raters. In this study, the raters are considered non-unique.

To determine the level of agreement between each teacher and their students, three measurements were considered: the strength of agreement or Kappa (*K*), the statistical significance (*p*), and the number of times an agreement of appropriate level is seen. Together, these three enable a judgement to be made as to the *total level of agreement* between teacher and student.

Another way to consider the total level of agreement could be to describe it as the *quality* of agreement (a reflection of the *combination* of statistical significance (*p*) and strength of agreement (*K*)), and *quantity* of agreement (the number of statements (n) on which the teacher and student agree).

Quality = significance (*p*) + strength (*K*)

Quantity = number of statements (*n*)

The level of statistical significance most often used in educational research is 95% or $p < .05$ (Field, 2018). This is often quoted as a cut off level to determine whether an analysis can be considered statistically significant. This level indicates that 5 out of 100 times the result of the analysis will be due to chance (Creswell & Guetterman, 2018), however the level of statistical significance can be used at other levels with good reason (Cumming & Finch, 2005). Indeed, Lancrin (2019) in his assessment of creativity and critical thinking in primary school students uses a level of *p*

< .20. Reasons for using a different p-value, or not reporting it, include small sample sizes and imprecise measurements as these are capable of producing distorted *p*-values (Cozby & Bates, 2012; Greenland et al., 2016; Wasserstein & Lazar, 2016). Small sample size and imprecise measurements were both factors potentially affecting the data in this study, so the use of *p*-values to determine statistical significance required consideration with each round of statistical analysis. Where the sample size was deemed to be low, i.e. significantly lower than the desired sample size of 65 (calculated at a statistical significance (alpha) of .05, effect size of .50, with power at .80: from Lipsey, 1990, as cited in Creswell & Guetterman, 2018, p. 614), a judgement on the quality of agreement was determined on the strength of agreement (*K*) alone.

The strength of agreement is measured by effect size or Kappa (*K*). Table 38 shows Cohen's classifications of effect size, adapted from Laerd (adapted from: Laerd Statistics, 2017b; Landis & Koch, 1977). A Kappa of 1.00 denotes perfect agreement, poor agreement scores between .01 and .20 and a negative value indicates that agreement is no better than chance. The cut off level for a noticeable effect size was therefore determined at $K > .20$ for this study.

A Kappa provides a measure of interrater agreement between two judges. A characteristic of Fleiss' Kappa is that where both judges give the same score for everything (e.g., both teacher and student give a score of three for the two statements being compared) then the interrater agreement cannot be calculated as it registers as a constant. Although the value of Kappa is not generated, in effect this is a perfect agreement between the judges. Where the interrater agreement could not be calculated this is clearly indicated in the tables in this chapter, but it is treated as having a significance of $p < .01$ and a positive, almost perfect agreement.

Outliers are values that are substantially higher or lower than the other values in a data set (Pallant, 2013). When determining if there were outliers in this study, the total level of agreement between teachers and their students was classified by the total percentage of statements on which the teacher and students had an acceptable strength of agreement with their students i.e., quantity of agreements and strength of agreements (*K)*. The limits of what was considered an acceptable strength of agreement is described with the results of the analysis.

Initial Test Analysis. The multi-step process that was followed to analyse the data is shown in Figure 15. The data was initially combined, and a test analysis run with Fleiss' Kappa to determine if a difference between Presentation Format 1 and Presentation Format 2 was discernible. Significant agreement was determined by levels of $K > .20$ and $p \leq .05$.

Table 38

Description of Kappa levels

Value of *K*	Strength of agreement
≤0	Chance agreement
0.01-0.20	Poor agreement
0.21-0.40	Fair agreement
0.41-0.60	Moderate agreement
0.61-0.80	Substantial or good agreement
0.81-1.00	Almost perfect agreement

Note: adapted from Laerd Statistics (2017b); and Landis and Koch (1977)

Identification of Outliers. Following the initial analysis, the data was again analysed with Fleiss' Kappa comparing the level of correlation between each teacher and their students. This enabled the teachers to be compared, to identify any outliers, i.e. those values that are well below or well above the other scores (Pallant, 2013).

Investigation of Reasons for Outliers. A series of analyses were carried out with Fleiss' Kappa on the combined data set, examining individual factors which might have an effect on the presence of outliers. The factors tested for were: (for teachers) number of years teaching, number of years teaching the current year level, whether they had previously used the AC C&CrT continuum to assess students, and their confidence to use the AC C&CrT continuum to assess students, (for students) their year level, their sex. Significant agreement was determined by levels of $K > .20$ and $p \leq .05$.

Two comparative analyses using Fleiss' Kappa were run to determine if there was any advantage with either of the presentation formats, for students with specific learning complexities, compared with students not diagnosed or identified with learning complexities. The conditions tested for were: Autistic Spectrum Disorder (ASD), Auditory Processing Disorder (AP) or Dyslexia in one analysis (these students were grouped together as there was some overlap in diagnosis of these conditions); EALD (English as an Additional Language or Dialect) students in a second analysis; and

finally students identified as being hearing impaired, having an unspecified intellectual disability, or being on a One Plan (an unspecified requirement for teaching to be modified). Significant agreement was determined by levels of $K > .20$ and $p \leq .05$.

Comparison of Presentation Formats. The final analysis of the data using Fleiss' Kappa was carried out following removal of the outliers. This enabled a comparison between presentation formats to be performed, and a preferred format to be identified. Significant agreement was determined by levels of $K > .20$ and $p \leq .05$.

All analysis was carried out using IBM SPSS Statistics (Version 28). Preliminary assumption testing was conducted to check for normality.

8.5. Results

The outcomes of analysis were not as expected. There was a low level of agreement between the student self-assessment and the teacher assessment, and this lack of correlation was reinforced at each stage of the analysis. Contrary to expectations, the outlying data related to higher levels of agreement in the assessment by two teachers, for which no evident cause was identified.

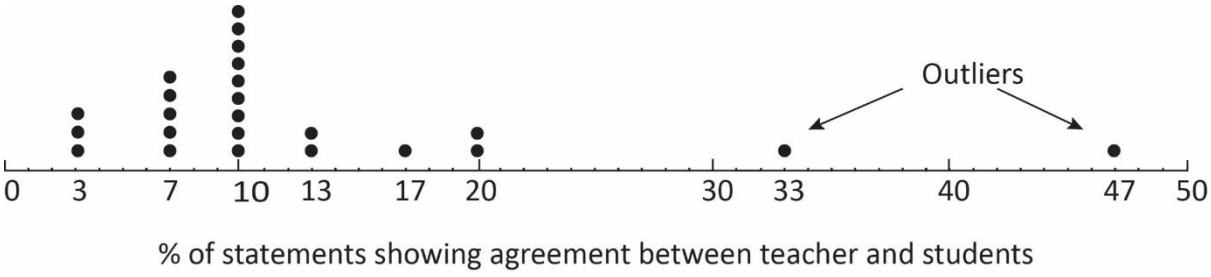
A full set of the results is included in Appendices H and I. A substantial number of the results of the analysis were not normally distributed. As non-parametric data was therefore indicated, all analysis methods were required to account for this. Levels of skewness and kurtosis are demonstrated in the following examples: e.g., for the student response for Presentation Format 1 Asking Questions Foundation level (AQF) skewness was -3.990 (SE = .180) and kurtosis was 15.278 (SE = .357) and for Presentation Format 2 AQF skewness was -3.468 (SE = .163) and kurtosis was 11.535 (SE = .324). For the teacher response for Presentation Format 1 Creating New Ideas Foundation level (CRF) skewness was -4.328 (SE = .179) and kurtosis was 19.826 (SE = .356) and for Presentation Format 2 CRF skewness was -1.741 (SE = .163) and kurtosis was 2.195 (SE = .324).

Preliminary Analysis

Using Fleiss' Kappa to compare the level of agreement between the students and the teachers results, a very low level of agreement was evident for both presentation formats. This preliminary analysis indicated statistically significant results ($p \leq .05$) in four out of the 30 statements in Presentation Format 1, and five of the 30 statements in Presentation Format 2, however none of the statements with a statistically significant result achieved a strength of agreement above the level of 'poor' ($K > .20$). The full set of results of this analysis can be found in Appendix I Table I-1. These results were considered inconclusive, so the data was checked for outliers.

Identification of Outliers

A full set of the results is found in Appendix I, Table I-2. A summary of the results of the analysis of the individual teachers and their students is shown in Table 39. The number of statements for each sub-element (e.g., Asking Questions) which achieved a positive Kappa are shown. The number of cells with $K > .20$ and $p \leq .05$ were counted and the totals and percentage they represent are shown in the final row of the table. As shown in Figure 17, the 2 teachers who clearly identified as outliers were those with a high level of agreement, the remaining 22 teachers having an agreement level of 20% and below. Teachers 16 and 21 and their students (one class of year 4/5, and one of year 7) were therefore removed from further analysis.



Note. The agreement between teachers and students is based on the number of statements for which $K > .20$ for each individual teacher, presented as a percentage of the total number of statements. Teachers 16 and 21 identify as outliers.

Figure 17

Individual teachers identified as outliers

Table 39

Summary of Fleiss' Kappa results showing number of results with positive Kappa for individual teachers (Part 1)

Sub-element	Teacher ID	1	2	3	4	5	6	7	8
	***n* (Students) =**	23	18	12	24	21	6	11	10
	Presentation Format	2	2	1	2	1	2	2	2
Asking Questions		0	1	1	0	4	2	3	1
Finding and Organising Information		2	1	1	1*	1	3	3	0
Connecting Information		4*	1*	4	1*	0	2	0	1
Creating New Ideas		1	1	2	2	1	2	1	0
Evaluating a Plan or Idea		4	2	5*	3	3	5	3	1
	No of cells *K* >.200	3	1	6	3	2	6	3	2
	% of cells *K* >.200	10%	3%	20%	10%	7%	20%	10%	7%

Note. *$p \leq .050$ for one or more results

Table 39 (part 2)

Sub-element	Teacher ID	9	10	11	12	13	14	15	16
	***n* (Students) =**	11	21	22	15	21	17	20	22
	Presentation Format	2	2	1	[a]	2	2	1	1
Asking Questions		2	1	2	0	2*	3	2	5*
Finding and Organising Information		2	0	4*	1	1	4	4	5*
Connecting Information		1	2#*	2	0	2*	4*	2	6*
Creating New Ideas		1*	2	3	0	3	2*	0	4*
Evaluating a Plan or Idea		3	0	2	0	3	2	4	4
	No of cells *K* >.200	2	2	3	1	3	3	1	10
	% of cells *K* >.200	7%	7%	10%	3%	10%	10%	3%	33%

Note. # indicates the Kappa was not calculated due to perfect agreement between the judges. This is reported as having a significance of $p < .010$ and a Kappa at or near 1.00 (perfect, or almost perfect agreement). *$p \leq .050$ for one or more results

Table 39 (part 3)

	Teacher ID	17	18	19	20	21	22	23	24
	***n* (Students) =**	15	17	22	17	11	6	22	23
Sub-element	**Presentation Format**	2	1	1	2	1	1	2	1
Asking Questions		2	1*	1	1	2#*	2#*	2	2
Finding and Organising Information		1	2	2	1	6	2	4*	1
Connecting Information		5*	0	0	2	4*	1	3	1
Creating New Ideas		0	0	1	2	4	2	0	2
Evaluating a Plan or Idea		3	0	3*	1	6*	1	0	1
	No of cells *K* >.20	5	1	3	2	14	4	4	1
	% of cells *K* >.20	17%	3%	10%	7%	47%	13%	13%	3%

Note. # indicates the Kappa was not calculated due to perfect agreement between the judges. This is reported as having a significance of $p < .010$ and a Kappa at or near 1.000 (perfect, or almost perfect agreement). *$p \leq .050$ for one or more results

8.6. Investigation of Reasons for Outliers

Fleiss' Kappa results were compared for teachers based on the responses to the screening questionnaire to which they had initially responded to determine whether any patterns were obvious based on their responses. The data was also compared based on student sex and year level. The full results of the analysis can be found in Appendix I Tables I-3 to I-8, and a summary of the results can be seen in Table 40. None of the variables indicated a significantly lower or higher total level of agreement between the teacher and student, so none of the variables investigated were considered likely to be the cause of any outliers. Characteristics of the two outliers were also compared. One of the teachers had been teaching less than five years, had never used the continuum to assess students and was only somewhat confident to use the continuum to assess students. The second teacher had been teaching for over 5 years, had used the continuum previously to assess students and was reasonably confident using the continuum to assess students. These investigations did not provide any insight into the reasons for the two outliers having achieved significantly higher levels of agreement with their students than the majority of the teachers.

Investigating Student Variables for Effect on Outliers. For students with no diagnosed complexities, one statement was both statistically significant and just higher than 'poor' in terms of strength of agreement, (Connecting Ideas Year 4 level in Presentation Format 2; $p \leq .001$, $K = .209$). For the students identified with ASD (Autism Spectrum Disorder), AP (Auditory Processing) and Dyslexia, two statements were both statistically significant and above 'poor' in terms of strength of agreement (Asking Questions Reception level in Presentation Format 1 with a moderate strength of agreement ($p \leq .050$, $K = .542$), and Creating New Ideas Reception level in Presentation Format 2 with a good strength of agreement ($p \leq .050$, $K = .619$). There were no significant results for the other students tested, in either presentation format. The significant results comparing students with no diagnosed learning complexity; students with ASD, AP and Dyslexia; and EALD students, are shown in Table 41 – full results can be found in Appendix I Table I-9 and I-10. None of the variables indicated a significantly lower or higher total level of agreement between the teacher and student for this group of students. This investigation was inconclusive, so none of the factors investigated were considered to have had an effect on the presence of outliers.

Table 40

Summary of results of analysis with all teachers and students, comparing a variety of variables

Variables	Categories	*n*	No. of statements with significant agreement *K* > .200 and *p* ≤ .050
Number of years teaching	First year of teaching	1	1
	Teaching >1 year and < 5 years	3	3
	Teaching 5 years or more	20	0
Number of years teaching the current year level	First year teaching this year level	4	1
	Taught this year level for the first time last year	3	3
	Third year (or more) teaching this year level	17	0
Has previously used the AC C& CT continuum to assess students	No	9	1
	Unsure	2	0
	Yes	13	0
Confidence to use the AC C& CT continuum to assess students	Not confident	5	0
	Somewhat confident	14	0
	Reasonably confident	5	4
Year level of students	Year 4	108	2
	Year 5	35	2
	Year 6	157	0
	Year 7	106	0
Sex of students	Female	232	0
	Male	174	0

Note. Full results of the analysis can be found in Appendix I Tables I-3 to I-8

8.7. Comparison of Presentation Formats

Following removal of the outliers, the two presentation formats were compared. As can be seen in the summary of results in Table 42 (full results can be found in Appendix I Table I-11), there were two statements in Presentation Format 1 which indicated significant correlations between teachers and their students ($p \leq .050$), with no statements achieving a level of agreement higher than 'poor' ($K \leq .200$). In Presentation Format 2 there were four statements which achieved a statistically significant result, but again, the level of agreement was 'poor' ($K \leq .200$). The hypothesis that the presentation format would have an effect on the level of correlation between the student and teacher assessments was therefore not supported.

Table 41

Summary of results of analysis comparing students with and without learning complexities

Presentation Format		*n*	No. of statements with significant agreement *K* > .200 and *p* ≤ .050
Presentation Format 1	No condition	156	0
	ASD, AP, Dyslexia	15	1
	EALD	11	0
	Hearing, Intellectual, One Plan	7	Too few complete subjects
Presentation Format 2	No condition	195	1
	ASD, AP, Dyslexia	12	1
	EALD	9	0
	Hearing, Intellectual, One Plan	7	0

Table 42

Summary showing statistically significant results of analysis comparing Presentation Format 1 and Presentation Format 2 following removal of outliers

Presentation Format	*n*	Sub-element	*K*	*p*
Presentation Format 1	151	Finding and Organising Information Year 8	.130	.031
		Finding and Organising Information Year 10	.120	.050
Presentation Format 2	222	Connecting Information Reception	.170	.004
		Connecting Information Year 2	.100	.047
		Connecting Information Year 4	.200	< .001
		Creating New Ideas Year 6	.130	.012

8.8. Discussion

The purpose of this study was to compare two alternative formats of a novel self-report assessment task, to investigate Research Question 2b: 'How does the method of communicating a self-report measure to students affect the accuracy of their responses?' Answering this question determines the presentation format to be used for presenting the assessment task to students in Study Three. The two presentation formats were tested on students and their teachers in mainstream classrooms - one format presenting the information to the students in a verbal form, the other combining a verbal presentation with text. The hypothesis was that students assessed with the self-report critical and creative thinking task under the condition of simultaneous text and verbal presentation would have a stronger correlation with the teacher's assessment of their critical and creative thinking skills, than students assessed under the condition of verbal only presentation. This hypothesis was consistent with the sensory modality approach to the theory of dual coding (Mayer, 2005a; Paivio & Clark, 2006), overlaid on Cognitive Load Theory (Sweller, 2005a) and mainstream classroom teaching practice (Montali & Lewandowski, 1996).

Initial Test analysis

The initial analysis results indicated no statistically significant difference in the correlations of the student responses with the teacher responses for the two presentation formats under consideration. The analysis also revealed a remarkably low level of agreement between the student and teacher responses, the reasons for which called for more detailed analysis of the study data.

Investigating the Possible Causes of Outliers

The process of investigating potential outliers brought to light the relatively low level of agreement between the majority of the teachers and their students. The two teachers identified as outliers were those with the highest level of agreement with their students. This came as a surprise, as it was expected that the initial screening process would have removed teachers who would have a low level of agreement with their students.

There was certainly no expectation of teachers having a 100% agreement with their students. This would indicate that both teachers and students were in full agreement about the level of skills students use. This would be unlikely, as the premise of using a self-report was that teachers are not always able to observe student skills, affecting the accuracy of their assessments. It would also have indicated that all the students were fully capable of assessing their use of skills accurately, which is also unlikely. It was however expected that the teachers in the study would have a reasonable idea of the levels of skills used by their students during the year, as the majority of students had been taught solely by the teacher for almost a full year.

In the study described in this chapter, the percentage agreement rate between teachers and students varied from 3% to 47%, with the greatest proportion of teachers not exceeding a rate of 10% of the statements in agreement with their students. Teacher #2 with 3% agreement was only in agreement with their students for one of the 30 statements in the task. At 47% teacher #21 and their students agreed about the level of skills of the student for just less than half of the questions they were asked. As teacher judgement is the norm for assessing students, it is noteworthy that the rate of agreement between what the students think they are doing and what the teachers think they are doing is so different.

The agreement levels found in this study between the teachers and their students echo the results of previous studies by Bates and Nettelbeck (2001), Gralewski and Karwowski (2019) and Machts et al. (2016). Sudkamp et al. reports of a high positive correspondence (Südkamp et al., 2012), however their report did not distinguish between the age or year group of the students and all studies reported were correlated with performance test results rather than student self-reports.

As reported by Hung, and others, (Gagné & St Père, 2001; Hung et al., 2016; Spinath, 2005) teacher judgement appears more accurate when the task is a performance test. As noted generally in this thesis, studies assessing critical and creative thinking skills (rather than creativity more generally) are scarce, and this was again found when investigating the accuracy of teacher judgement assessing these skills. Gralewski and Karwowski (2019), however, report the strength of the relationship between students' scores on creativity tests, and their teachers' assessment of their creativity, is usually weak.

Further analysis was carried out to determine whether there were any factors that accounted for the low levels of agreement between the students and their teachers in this study. As in other studies, the results were inconclusive (Machts et al., 2016; Südkamp et al., 2012). When considering the student variables, neither the year level of the students nor their sex appeared to influence results. Research on gender as a moderating factor reports either a very small effect or no effect (Gralewski & Karwowski, 2019; Urhahne & Wijnia, 2021) but in general there is a significant lack of research on the effect of student factors on student/teacher agreement. (Südkamp et al., 2012; Urhahne & Wijnia, 2021)

When investigating the teacher variables it must be noted that the number of teachers was small (n = 24). Caution is recommended when drawing conclusions from data relying on small numbers of participants as small numbers attract a higher risk of sampling error (Creswell & Guetterman, 2018, p. 144). In this study it was observed that teachers with the highest proportion of statements with significant agreement were those who had been teaching for more than one year but less than five years, teachers who were teaching the current year level for the second year in a row, and teachers who had claimed to be reasonably confident to use the AC C&CrT continuum to teach students. This may reflect a stage of development of teaching skills through which these teachers are progressing, e.g., less invested in older modes of practice and more epistemically fluid. As with research on student factors, little research was found on the effect of teacher characteristics on the reliability of their assessments. Diagnostic competence and job experience have been reported as having a weak association or no association (Machts et al., 2016). Familiarity with the assessment is reported by Leucht et al. (2012) as having a positive effect on teacher's accuracy, as is teacher confidence (Praetorius et al., 2013). In respect of these factors the teachers in this study had very little familiarity with the measure, but a majority had moderate to high confidence.

Causes for Low Rate of Agreement

The question remains – why were the rates of agreement so low? As the teachers had been screened before the research to remove the majority of inexperienced teachers and those who were not confident in using the Critical and Creative Thinking learning continuum, it was reasonable to consider the teachers as 'experts' in the field, and to expect them to be able to judge their students with reasonable accuracy, as they do at every reporting period. There are three possible categories of cause – the measurement instrument may have been faulty, teachers may have made errors, and/or students may have made errors. Factors behind teacher errors may include: a lack of familiarity with observing the skills in question; the gradations from one level of a skill to the next being too small for observation; not consciously observing the skills as they were being used; or being unfamiliar with the levels of skill above and below their current classroom year level. Student errors may have been caused by: not understanding the statements; mis-aligned interpretations of the statements; being unfamiliar with self-reporting; lack of metacognitive skills; or not being honest. Issues of accuracy of the measurement instrument are addressed below.

Level of Students' Language Comprehension. The statements underwent examination by an expert panel of teachers ranging in experience from lower primary to secondary school, to ensure the statements were appropriately worded for student comprehension. Furthermore, the statements in the earlier version of the measure, described in the STEM Learning Evaluation Grant Report in Appendix A, had undergone a form of cognitive pre-testing (Woolley et al., 2004) with Year 6 and 7 students. This testing took the form of asking students to provide an example of when they had used a skill described in a statement. Students demonstrated that they could complete this task, indicating that they understood the meaning of the statements. Although the statements the students were asked to respond to in the final task were not identical to the original statements, they had been simplified by reducing the concepts described in each statement and by simplifying the language. It was therefore expected that this would make the statements easier for the students to understand. Unfortunately, access to the students at this stage of the development of the measure was limited due to COVID-19. This level of testing with the students could not therefore be carried out, so the depth of understanding by the students – and the teachers – remains to be explored more thoroughly.

Misaligned Statement Interpretation. There are three possible components to this problem. The first, is that students and teachers may not have been interpreting the statements in the same way. Teachers, with their wealth of experience, would be expected to approach the statements with an awareness of what the skills could look like when being used by a highly competent student at the end of secondary school. They should be aware that all or most of their primary school students

would not measure up to this level of application of the skills. Many students on the other hand will not have had the opportunity to see the skills in use at that level of competency, so the Dunning-Kruger effect would come into play (when people think they know because they don't know what they don't know) (Dunning, 2011). This might have been alleviated to some extent if a comment had been included in the instructions about the statements being designed to describe the use of skills from Reception to Year 10, but this was not said to either the students or the teachers as it was desirable to allow them all to respond to each statement on its own merit, not based on potentially false expectations. Some teachers may therefore have assumed that all the statements were relevant to their students, while other teachers may have assumed that some statements were describing levels of the skills that would be unobtainable by their students. It was anticipated that students thought that the statements all related to skills they might be expected to demonstrate.

Later in the research, during Study Three (reported in Chapter 9) the assessment task was presented to students on multiple occasions with the goal of measuring any change in the use of the skills over a period of time. It is quite possible that the Dunning-Kruger effect became obvious to some students, when they reflected on how they used the skill at a later stage during the research period, compared to how they used it previously. I envisage that (some) students will have moments of awareness when they complete the assessment for the second or subsequent times, where they say 'I thought I was asking complex questions before, but now I can see that I'm asking about much more complex ideas than I was before.' This new awareness of the extent of use of the skills by the students is expected as they are learning and therefore developing their skills. This illustrates the impossible task of ensuring all respondents share a common understanding of the meaning of each of the statements presented in the assessment task.

The second issue was possible variations in teacher interpretations of the statements due to some statements being applicable to more than one year level. For example, one of the statements was, 'I ask questions to help me learn about complex or complicated issues'. This is the level of questioning expected by Year 10 students in the continuum. Some 8-year-old (Year 3) students ask questions about things that are very complex and complicated for them. What is complex and complicated to an 8-year-old though, is not likely to be complex and complicated to a 15-year-old. Does the teacher recognise that the 8-year-old is tackling issues that to them are complex and complicated, and therefore assess the student as demonstrating this skill at this level? Or do they assume that the questions asked by an 8-year-old cannot relate to issues that are complex and complicated enough to apply to this statement? While we cannot know if some teachers interpreted some of the statements in this way, it is reasonable to assume variation in how the teachers

interpreted the statements. The participants were intentionally not instructed that the statements related to Reception to Year 10 level statements but were instructed that the statements were not in any order. This was done to avoid both teachers and students making assumptions about what they must be able to do just because of the year level they were at. Although the lack of instruction kept possibilities open, it also undoubtedly allowed inconsistencies to occur in the reasoning behind some of the responses.

If teachers assume that the questions asked by an 8-year-old cannot relate to issues that are complex and complicated enough to apply to the statement 'I ask questions to help me learn about complex or complicated issues' then we also have to ask what the teacher is actually measuring or reporting on. Is the teacher measuring the application of the skill (Does this student ask about complex and complicated issues?) or are they measuring the level of knowledge to which the skill is being applied? (Are the *topics* this student asks questions about, complex and complicated issues?) To an 8-year-old asking questions about the use of 'effect' or 'affect' may be complicated, but the teacher may not view it as a complex or complicated issue. To a 15-year-old (and many teachers) asking about the ethics of purchasing Chinese solar panels to power a farm may be the level of questioning required to be considered a complex and complicated issue, but the question itself might be a simple enough question, e.g., "Why do people say we shouldn't buy Chinese solar panels?" versus "What are the ethical implications of Chinese solar panel manufacturers flooding the Australian market with inexpensive solar panels?". The first question may be considered a simple question about a complex issue, while the second question is a complex question about a complex issue. These different ways of viewing the question could certainly cause inconsistencies in the responses from the participants. This comment also further reflects on the lack of relevance of the continuum to some of the skills e.g., Asking Questions, as discussed in Chapter 7.

Graduations Between Levels of Skills. Statements in some of the sub-elements may be too similar, e.g., statements in 'Connecting Information' (significant differences between the statements are shown in bold): Year 4 statement – 'I make connections between information I learned in different **places or times** and use this to help me understand things': Year 6 statement – 'I make connections between information learned in different **subjects**, and use this to help me understand things in **new ways**'.

Similarly, the statements may not allow for enough of a graduation between the individual levels of skill for students and teachers to observe the difference, e.g., Year 8 statement – 'When I come up with a plan or an idea, I am able to analyse my plan to see where **parts of the plan may need more thought**'; Year 10 statement – 'When I come up with a plan or an idea I am able to

analyse my plan to **tell you** where there **might be mistakes in my reasoning**'. This may be making it difficult for both the students and the teachers to accurately report on which levels of skill apply to the student.

Lack of Familiarity with Metacognition and Self-Reporting. Being aware of one's own learning skills and behaviours is not something we tend to see in students without them having been explicitly taught to do it. Although there was not the opportunity to run any metacognition training for the students in this research, it would undoubtedly have helped them to be more aware of the skills they use and don't use, and the extent to which they use them. The training sessions described by Yu-Ju Hung and colleagues in their research on sixth graders in Taiwan being trained in peer and self-review (Hung et al., 2016) serve as an example of what is possible given the time. Likewise, providing some information to students on how to effectively self-report, might have helped them to be more accurate in their assessments of their own use of skills. It is not ideal to use self-report as the sole source of assessment, which is why we attempted to use teacher assessments to validate the student self-reports. Self-reports are however indispensable for capturing the internal thoughts of the participants (Pekrun, 2020) which are not on display for the teacher to observe.

Lack of Familiarity with the Skills. When asked to complete the teacher assessment on the sub-element 'Evaluating a plan or idea', one of the teachers commented that it was not something she ever did with her students, so she would find it challenging – and so would they. In fact, those students told me that they used those skills in other classes, so they had less difficulty assessing themselves than the teacher had. Although the expectation is that teachers know what they are doing, teachers are individuals with varying skill sets of their own. Teachers cannot be expected to be experts at observing and teaching all things, including skills. Even when skills are being worked on, the teacher may not be explicit about the skills students are using and how they can recognise their own developing competency. This would potentially affect how accurately the students would be able to report on their use of the skills.

Lack of Familiarity with Skill Levels. It is possible that some of the teachers were not familiar enough with the various levels of skills described in the year levels well below and well above the level they were teaching at, to recognise them effectively in their students. Although they could be expected to have demonstrated the use of the skills at the highest level themselves during their lives, they may not be familiar with observing them in others. This may account for some of the times when students scored themselves as a '*Yes*', but teachers scored them as a '*No*'.

Comparing the Results of Study Two with Literature on Multimedia Learning. The assessment task being analysed in Study Two differed significantly from the assessment tasks described in the literature on multimedia learning, which may account to some extent for the difference in the results. Firstly, all the participants in the studies described in the literature had been assessed on how effectively they were able to recall information they were being taught, with the various modes of presentation as the variables (e.g., Cerpa et al., 1996; Wieland, 2014). It was possible to assess the effectiveness of the students' learning by scoring their task. The students in Study Two were not completing an assessment of learning, and although the teacher responses were expected to provide a guide as to the 'correctness' of the students' responses, this turned out to be an ineffective method of scoring the students (judged by the level of agreement found between students and teachers).

Secondly, the purpose for using the two modes of presentation in Study Two was to enhance understanding, not to enable learning demonstrated by recall. This was a new application of the theories of multimedia learning. It is not clear whether either format achieved that, but as the results were so similar, it appears that the mode of presentation had no positive effect, nor did it appear to have the sort of negative effect described by the redundancy principle (Kalyuga & Sweller, 2014). Future research on the effect of presentation formats used for understanding rather than for learning would be of value.

Thirdly, it is not clear from Study Two whether text is more accurately described as a visually processed form of information (Mousavi et al., 1995), or whether it should be viewed, as Mayer, Paivio and Sadoski assert (Mayer, 2009; Sadoski & Paivio, 2013), as a verbal form of information (inner speech). If it acts as a visually processed form of information, it should enable twice the capacity of working memory to be operating, and therefore enable students more capacity to consider the skill being described in terms of their own experiences. This would have the potential of increasing the accuracy of the student self-reports. If, on the other hand, text should be considered as the equivalent of a verbal form of information, the dual mode of two forms of speech would have potentially over-burdened the working memories of the students, providing little space to consider how they had acted in the past regarding their use of skills. If this was so, the redundancy effect would have possibly come into play. As the two forms of the task did not demonstrate significantly different results, it is impossible to say whether either of the effects mentioned played any part in the success or otherwise of the presentation formats.

Fourthly, applying CLT to explain the effect of doing an assessment low in element interactivity is a new application of this theory, which may not be an appropriate application,

particularly as the assessment is not focussed on the recall of facts, but on stimulating memories of actions the students have previously taken.

Finally, the majority of participants in the research described in the literature on multimedia learning were adults in specialised studies at university, college or technical college, or high-performing students from high school secondary classes (Cuevas & Dawson, 2018; Mayer, 2005b; Mousavi et al., 1995; Sombatteera & Kalyuga, 2012). It is not known how much the age and other characteristics of the students affected the results of that research, or of Study Two.

Preferred Assessment Task

The results of the investigation of Research Question 2b 'How does the method of communicating a self-report measure to students affect the accuracy of their responses?', were inconclusive. Both presentation formats achieved a low rate of correlation, and the rates of correlation were almost identical.

In response to Hypothesis 2b that students assessed with the self-report critical and creative thinking task under the condition of simultaneous text and verbal presentation (Presentation Format 2), will have a stronger correlation with the teacher's assessment of their critical and creative thinking skills, than students assessed under the condition of verbal only presentation (Presentation Format 1), this hypothesis was not supported. Rather than the presentation of identical text and speech presented simultaneously causing the redundancy effect, this has not been observed. It is possible that this is because the two modes were so similar that intrinsic and extraneous loads when combined were not a problem for the capacity of the working memory. This may be more of the reason for no difference in result, rather than the classification of the mode of information as described by the sensory modality or the presentation approach playing a part.

The inconclusive results in Study Two meant that additional factors had to be considered in selecting the presentation format to be used in study three. **T**he decision to use Presentation Format 1 in Study Three was based on the following factors:

- The visual impact of Presentation Format 1 is less intimidating for students, thereby creating a less stressful test environment.
- Presentation Format 1 supports students with lower levels of reading comprehension by removing the requirement to read text.
- Presentation Format 1 ensures all students work at the same pace, enabling the measure to be completed within the 50-minute lesson time.

8.9. Limitations, Constraints and Recommendations

An issue with this study has been the difficulty in confirming the accuracy of the student self-report scores. As discussed above, this may be due to the apparent inability of teachers to determine the level of skills their students are using, or inaccuracies of student reporting. The choice of a self-report task to measure the skills used by students was to provide the opportunity for the person 'doing' the skill, to report on what they actually do. Assessments are usually made by relying on the best estimate of the teacher who is ultimately making an 'educated guess' based on their observations in the classroom. The lack of correlation between these assessments deserves further investigation.

An obvious issue with this measure of critical and creative thinking is that the only form of validation so far is content validity (DeVellis, 2012, p. 102). How can we be sure we are measuring what we claim to be measuring (construct validity) or that the statements accurately represent skills appropriate to the students at specific year levels (content validity)? The short answer is, we can not be sure the task is measuring the development of skills, as research has not yet been done which provides evidence of how the skills develop, or the levels of skills students are capable of at different ages, stages, or year levels. This is not necessarily an issue at this point in the research and both issues are beyond the scope of this thesis but would be of immense interest for further studies. The measure is reporting on whether the students believe they have used the skills at the levels described in each statement. Content validity has been tested by the use of an expert panel, which is the only method currently able to be applied. For the purposes of this research this was deemed to be sufficient, as the requirement is to have a task that measures the skills being used by students during learning on multiple occasions, to indicate whether there is change demonstrated over time. The measure used in this study appears to be suitable for this purpose.

Future work with a focus on validating this measure would be valuable. Test-retest validity for example (Cozby & Bates, 2012, p. 98), or improving the psychometric properties of this measure through cognitive testing with students to ensure comprehension (Woolley et al., 2004) would both improve this measure. If the question that initially piqued my interest in researching the development of critical and creative thinking ('Can you prove it?") is to be effectively answered, robust measurement techniques will be required. The present study however provides a solid foundation on which to build.

8.10. Conclusions from Study Two

The purpose of this study was to compare the effect of the method of communication of a self-report assessment task to students. The results of the tests to determine which of the presentation methods was the most effective for assessing student critical and creative thinking skills were indeterminate, however trialling both presentations of the task and investigating the results informed the decision to use Presentation Format 1 in future studies. This assessment task will enable the critical and creative thinking skills to be effectively measured during learning activities.

Furthermore, the analysis brought out numerous issues with the task and with the statements derived from the C&CrT continuum, pointing to the need for significant further research if we are to understand how our students develop and use their skills, and how we can best support them to do that. It also brought to light the significant gap between teachers' views of what students are doing, and the student's understanding of what they are doing – another area which requires further investigation to enable effective teaching of the skills to our students.

Chapter 9.

Study Three: The Impact of Problem-Based Learning and Industry Partners on Critical and Creative Thinking

9.1. Introduction

The need to increase the use of critical and creative thinking abilities in students of all ages has been identified across the Western world (Gamse et al., 2017; Kelley & Knowles, 2016; Torii, 2018; World Economic Forum, 2020). This has inevitably translated into pressure on schools to use teaching methods that help students effectively develop critical and creative thinking skills (Department of Education, 2022; Office of the Chief Scientist, 2014). These skills are recognised as essential for effective learning generally, and for success in the workplace and STEM careers (Ministerial Council on Education, 2008; Office of the Chief Scientist, 2014; World Economic Forum, 2020). To determine whether teaching methods are effective in aiding students to use these essential skills requires the assessment of student learning (Mislevy, 2012). The research questions in this thesis further direct our focus to the skills used by students during learning. As discussed in Chapter 3, the lack of existing measures has made it difficult to compare the effect of different types of teaching on the use of critical and creative thinking skills during learning. This difficulty is further amplified by the use of a prescriptive Critical and Creative Thinking learning continuum in Australian schools, for which no form of measurement has been developed (Wyatt-Smith & Adie, 2018). The measure described in Chapter 7 and assessment task described in Chapter 8 were therefore developed specifically to measure the application of the critical and creative thinking skills described in the Australian Curriculum, Critical and Creative Thinking learning continuum, by primary school students during learning.

Effective methods of assessing students' skills enable judgements and comparisons to be made on the effectiveness of teaching methods. Problem-Based Learning (PBL) is a particular pedagogical approach to teaching and learning, currently being promoted as being effective for developing students' skills (Department for Education (DfE) South Australia, n.d.-b; Education Council, 2015). Although PBL has been used in schools since the 1960s (Barrows, 2000), research on PBL with primary school students is limited, and there is even less research reported on the effect of

PBL on the development or application of critical and creative thinking skills in primary school students (see the discussion of the literature on the subject in Chapter 4). PBL is also promoted as increasing student motivation and engagement (e.g., Dochy et al., 2003; Hmelo-Silver, 2004; Specht, 2015), but again, the research evidence with primary school students is extremely limited. The first two research questions to be investigated in Study Three are therefore:

- Research Question 3a: 'What is the impact of PBL on student critical and creative thinking skill application?'
- Research Question 4a: 'What is the impact of PBL on student motivation and engagement?'

The hypotheses associated with these questions are:

- Hypothesis 3a(i): PBL results in both an increased *range*, and an increased *level* of critical and creative thinking skills compared to a control group.
- Hypothesis 3a(ii): Positive gains in the critical and creative thinking skills will persist over time for the PBL students.
- Hypothesis 4a(i): PBL will result in an improvement in motivation and engagement compared to a control group.
- Hypothesis 4a(ii): Positive gains in motivation and engagement will not be retained beyond the PBL intervention.

One of the defining characteristics of PBL that is believed to influence its effectiveness in engaging students, is the nature of the problem presented to students (e.g., Hung, 2016; Jonassen & Hung, 2008; Rossi et al., 2019). PBL teaching is structured around a central problem which provides the stimulus for the student investigation. The problems that are generally recommended to stimulate student learning are real-world problems (Hmelo-Silver, 2004). Two types of real-world problems are promoted for use in primary schools in South Australia to provide the context for teaching with PBL – hypothetical problems, and industry partner problems (DfE, 2022; Department of Education, 2022; Rossi et al., 2019). Of the two types of problems, I have most commonly seen hypothetical problems used in primary school classrooms.

Hypothetical problems are problems that are crafted for the purpose of practising problem-based learning skills. They may be either sourced from a commercial STEM problem resource or developed by the teacher themselves. Teacher-developed problems may be connected to recent events in the news, or to problems in the school environment, e.g., 'We produce a lot of waste in our

school. What environmentally friendly changes can we make to manage our waste more effectively?' Problems sourced from commercial STEM problem resources are carefully structured to cover specific areas of the curriculum, and often include resources to provide background knowledge for the teacher, e.g., Macmillan STEM Investigations (Lioncrest, 2023). Both these types of hypothetical problems, when used in a PBL environment, are presented to the students by the classroom teacher.

The second type of problem, industry partner (IP) problems, are genuine, unsolved problems that are (ideally) currently being worked on by the industry partner. An example of this type of problem is one provided by the Holdfast Bay Council[24] in 2018: 'We have been instructed by the South Australian Government that we must decrease heat islands, increase tree canopy in our local council area by x%, and at the same time increase infill. How can we achieve all these things by 2030?'. Although the IP problem may not be accompanied by background material, the industry partner themselves is ideally available to provide additional information for the students and will respond to student queries during the PBL process. The IP problem is presented to the students by the industry partner instead of the teacher. This introduction of the problem by the person who owns the problem (described as subject presence by W. Hung; 2016) is one of the characteristics of problems which is believed to make a difference.

Both the hypothetical problem and the IP problem described above can be classified as 'real-world' problems. They are similar in that they are current problems, presenting genuine issues which require a solution.

IP problems tend to be more complex than hypothetical problems, and more removed from the experience of both the student and the teacher. This is because the hypothetical problem is created with the classroom in mind and is therefore developed on purpose to control the level of complexity. An IP problem, in contrast, reflects the complexity and unbounded nature of real problems in the real world, providing both benefits and challenges when introduced into the classroom. Using the problems described above as an example, the hypothetical problem sounds relatively simple. It would require participants to develop an understanding of how waste generated by the school is managed by the council, what different waste streams there are, how waste is sorted and stored prior to collection by the council, etc. In contrast, the IP problem requires a broader and greater level of understanding about subjects which the average teacher has no knowledge of – such as: what are heat islands and infill, and how do you classify tree canopy? The

[24] The school attended by the students who tackled this problem was located in the local council area designated as 'Holdfast Bay Council'.

level of complexity of this problem, for both the teachers and the students, would be significantly greater. This greater level of complexity is a characteristic of the IP problems I have observed.

The second difference is in the presentation of the problem. The hypothetical problem is presented by the classroom teacher as a problem that the class will investigate and solve together. The industry partner problem is presented by an industry partner—a 'stranger in the room'—which any teacher will affirm brings a level of excitement and interest from the students which is hard to match. It is difficult to make the presentation of a problem by the familiar classroom teacher as exciting an event. The IP problem, like the hypothetical problem, is presented as a problem that the class will investigate and solve together, but the effect of a stranger in the room who says to the students 'I have a problem – can you help me?' should not be underestimated. It is exactly this human connection which may affect the level of motivation and engagement with the students during PBL with an industry partner (Hung, 2016).

Many promoters of the use of IP problems cite increased student motivation and engagement (M&E) as a reason for the success of PBL (Department of Education Employment and Workplace Relations, 2010; Lonsdale, 2011; Rossi et al., 2019; Torii, 2018). Increases in motivation and engagement are also said to result in improved levels of skills being used (Rossi et al., 2019; Watters et al., 2016). As discussed in Chapter 5 however, actual research demonstrating increases in motivation and engagement, particularly with primary school students, is lacking. If the involvement of industry partners can be shown to help motivate and engage students, resulting in increased development of critical and creative thinking skills, then their use should be encouraged in schools.

Despite the claims made (by Rossi et al., 2019 and others) regarding the successful effects of using industry partners with PBL in schools, there are two issues which must be noted: first, there is no real evidence from research confirming that these positive effects can be attributed to the use of industry partners. Second, although using an IP problem is promoted as an effective way to encourage students to apply critical and creative thinking skills (Lonsdale, 2011; Rossi et al., 2019; Torii, 2018), engaging an industry partner to provide a problem for students to investigate is complex for many teachers (Lonsdale & Anderson, 2012; The Australian Industry Group, 2017; Wilson, 2021). If using a hypothetical problem with PBL, from a commercial resource, achieves the same effect on student critical and creative thinking skills as an industry partner problem, using a hypothetical problem would be the simpler, and therefore the preferred option for many teachers. This feature of PBL (the problem context) is therefore worth further investigation, particularly as there is a significant gap in the research literature (as identified in Chapter 4) pertaining to problem types, the use of industry partners and their problem, and the intersection of these with primary

school students and motivation and engagement. The two additional questions for investigation in this chapter and their associated hypotheses are therefore:

- Research Question 3b: 'What is the impact of using an industry partner problem in STEM PBL on critical and creative thinking skill application?'
- Hypothesis 3b(i): An industry partner problem results in both an increased *range*, and an increased *level* of critical and creative thinking skills compared to a hypothetical problem.
- Hypothesis 3b(ii): Positive gains in the critical and creative thinking skills will persist over time for both the industry partner PBL students and the theory PBL students.
- Research Question 4b: 'What is the impact of using an industry partner problem in STEM PBL on student motivation and engagement?'
- Hypothesis 4b(i): An industry partner problem will result in a greater improvement in motivation and engagement compared to a theory problem.
- Hypothesis 4b(ii): Positive gains in motivation and engagement will not be retained beyond the PBL intervention.

9.2. Method

Research Design

Study three comprised a classroom-based study of the impact of industry partner engagement in PBL, using students and teachers from 4 schools over a period of approximately 16 weeks. The study focussed on two treatment conditions – a hypothetical PBL problem, identified here as the theory PBL problem, and an industry partner problem, identified here as the IP PBL problem, relative to a control group. The participants for Study Three were not randomly allocated to the three treatment conditions, therefore this research is considered a non-randomised, between groups (with control group), quasi-experimental repeat measures (pre- and post-test) design.

The research for Study Three incorporated five steps, as shown in Figure 18. In Step One the teachers were trained in the Problem-Based Learning process to be used during the intervention. This ensured all teachers were using a similar teaching process for the intervention. Step Two was the Time Point One (TP1) pre-intervention data collection which occurred during the first week of the term in which the intervention occurred. It was conducted by the researcher and involved the

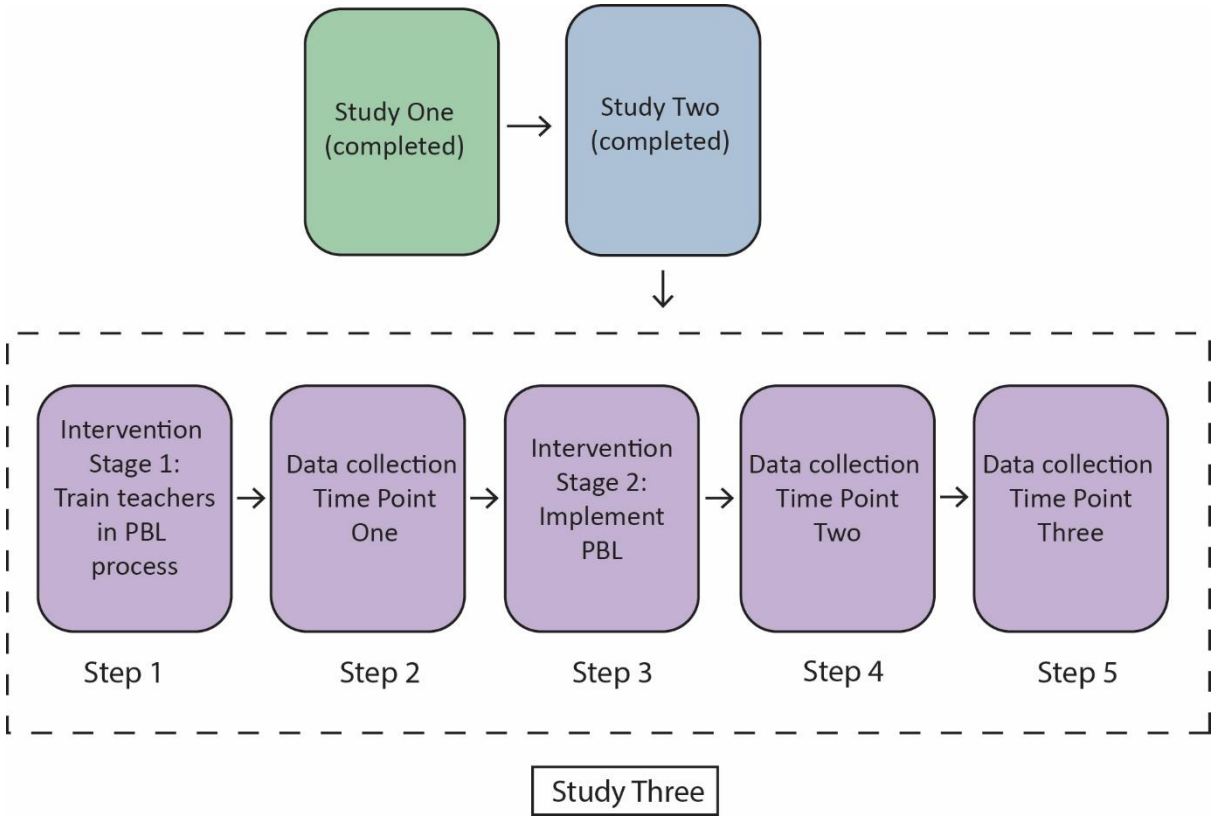


Figure 18

Research design: Study Three

students completing the Critical and Creative Thinking Self-Report Measure and the Motivation and Engagement Scale in the classroom. Step Three was the implementation by the classroom teacher of the PBL intervention in the classroom which continued for a full 10-week term. This step included the use of 'critique panels' to provide external feedback to students on ideas in development. Step Four, post-intervention data collection at Time Point Two (TP2), was similar to the TP1 data collection, but carried out during the final week of the term, following the students' presentation of their problem's solutions. Step Five, follow-up data collection at Time Point Three (TP3), was identical to the TP1 data collection and was carried out approximately six weeks after the TP2 data collection.

Participating teachers were briefed prior to the commencement of the school year, to ensure they had time to plan for both the PBL intervention and the appropriate curriculum lessons during the research term. The classroom teachers taking part in the research nominated the term in

which they would run the PBL unit with their students – either Term 2 or Term 3 of a standard four-term school year. They were aware that the problem (either the theory problem or the IP problem) would have a STEM focus and relate most closely to the biology strand of the Australian Curriculum and had the potential to incorporate money and measurement (a subset of maths) and geography (specifically climate). The teachers were encouraged to schedule the teaching of the appropriate curriculum lessons (biology, measurement and money, and geography) during the same term as the PBL intervention was being run. The knowledge component of the curriculum which would relate most to the problem to be solved, would therefore be fresh in the students' minds during the PBL lessons. The teachers each received one and a half days' training in PBL from the researcher in the second half of the term prior to the intervention being carried out in their classrooms.

Participants

Unlike Studies One and Two, the participating classroom teachers in Study Three were not chosen for background knowledge in critical and creative thinking. The participating teachers were specifically chosen to have little or no experience in Inquiry learning and Problem-Based Learning. All the teachers in Study Three were recruited in South Australia through Catholic Education SA (CESA), and there was no overlap between the participants of Study Three and the previous studies. the STEM Learning and Curriculum Consultant in CESA recommended potential schools that had indicated an interest in PBL. The principals of the schools were contacted, and a short-list of interested schools was identified. Fourteen teachers and the students in their classes from four schools in metropolitan Adelaide, South Australia were identified as potential suitable participants. The schools were located in a variety of suburbs – from coastal white-collar suburbs south of the city and in the hills to the more blue-collar area of the Northern suburbs. The schools also covered a range of ICSEA values[25], ranging from a value in the high 900's to over 1100[26] (ACARA, 2023a). The short-listed teachers were screened through use of an online survey (Appendix J-1) to ensure their suitability.

The participating classroom teachers were responsible for teaching students in Years 5 and 6, with ages ranging from 8 years 2 months to 12 years 5 months. None of the schools or teachers

[25] ICSEA (Index of Community Socio-Educational Advantage) is a scale of socio-educational advantage that is calculated for the majority of schools in Australia. 'ICSEA values are calculated on a scale which has a median of 1000 and a standard deviation of 100. ICSEA values typically range from approximately 500 (representing schools with extremely disadvantaged student backgrounds) to about 1300 (representing schools with extremely advantaged student backgrounds).' Australian Curriculum Assessment and Reporting Authority (ACARA). (2023a). *Guide to understanding the Index of Community Socio-educational Advantage (ICSEA)*. https://myschool.edu.au/media/1900/guide-to-understanding-icsea-values-2023.pdf

[26] exact values are not provided to maintain school anonymity.

included in the research identified as using Inquiry teaching or Problem-Based Learning, none intended to use these styles of teaching prior to the research commencing, and none were registered as International Baccalaureate Schools[27]. Fourteen Year 5 and 6 classes were chosen to take part in the research. In school A four of the six classes chose not to take part in the PBL intervention. These classes were therefore ideal for the control group as they were the same year group as the intervention students and attending the same school. Although it would have been ideal to also have control classes in the other participating schools, none were available. Convenience sampling determined which of the three treatment conditions was applied to each school as shown in Table 43. This ensured a reasonable balance of participants in the treatment conditions and enabled the industry partner's time schedule to be accommodated.

The student participants (*N* = 257) were being taught by the recruited teachers, and who were working at an English language level of Year 4 or above. The distribution of students by treatment group (control, theory PBL or IP PBL), sex and year level can be seen in Table 43. The ideal group size of 65 participants (calculated at a statistical significance (alpha) of .05, effect size of .50, with power at .80: from Lipsey, 1990, as cited in Creswell & Guetterman, 2018, p. 614) can be considered to have been met.

Prior to research commencing, approval was obtained from the University of South Australia's Human Research Ethics Committee; Catholic Education South Australia; and the principals of all schools involved in the research. Information about the research was provided to each school principal, teachers and the families of student participants, and active informed consent was received from all participants. Information and consent forms are included in Appendix J-5. Personal data about each student (date of birth, sex, English language level, diagnosed learning complexities) was provided by each school to enable screening of students with low English language levels for exclusion, and determine whether any of these factors had an influence on the results of the analysis.

[27] IB schools in South Australia use Inquiry teaching.

Table 43

Participants in Study Three

Variable	Categories	Control Group	Theory PBL Group	IP PBL Group
	n =	64	106	87
	School	A	A, B, C	D
Sex	Female	33	67	40
	Male	31	39	47
Learning Complexities	ASD	0	3	5
	ADD/ADHD	0	1	0
	APD	0	1	1
	Dyslexia	0	2	3
	EALD	43	38	5
	Other	0	9	4
English Instructional Level	Below Year 4	0	1	0
	Year 4	0	4	0
	Year 5 and above	64	101	87
Year Group	Year 5	29	59	52
	Year 6	35	47	35

Note. ASD = Autism Spectrum Disorder

ADD/ADHD = Attention Deficit Disorder/Attention Deficit Hyperactivity Disorder

APD = Auditory Processing Disorder

EALD = English as an Additional Language or Dialect

Other = unspecified learning issues

Measures

Three measures were used in this study – two student measures and a teacher post-study questionnaire. The two student measures were combined as a single assessment task (attached in

Appendix J-2). The first was the Critical and Creative Thinking Self-Report Measure, the development of which is described in the previous two chapters. This self-report measure consists of five sets of six statements (30 statements in total), each set of statements relating to a specific critical or creative thinking skill. The statements were presented verbally to the students who indicated with a 3-point categorical scale response by marking numbered boxes with a tick for '*yes I do this'*, a cross for '*no I don't do this'*, or a circle for '*I am unsure if I do this'*.

The second measure was the Motivation and Engagement Scale–Junior School (MES-JS; Martin, 1999-2022). This scale consists of 44 statements measuring 11 first order factors which contribute to the higher order factors of positive motivation and engagement, and negative motivation and engagement e.g., for positive motivation: 'If I try hard, I believe I can do my schoolwork well', and for positive engagement: 'If I can't understand my schoolwork, I keep trying until I do'. This self-report scale was also presented to the students in a verbal format, with responses indicated in a 5-point Likert-style scale of '*disagree strongly'*, '*disagree*', '*neither agree nor disagree'*, '*agree*' or '*agree strongly'*. The statements which comprise the Motivation and Engagement Scale can be found in Appendix J, Table J-3.

The teacher post-study questionnaire was to provide supplementary information – qualitative in nature, but not comprising a qualitative study – to identify if any events had happened during the term that might have impacted on the research outcomes. This questionnaire is attached in Appendix J-4.

Procedure

The procedure included the intervention and the data collection. The intervention consisted of two stages – the first stage was to train the teachers in the form of PBL to be used with the students. The second stage was for the teachers to run the PBL intervention with their students. The data collection occurred at three times: before the intervention (pre-intervention), immediately following the intervention (post-intervention), and approximately eight weeks after the intervention had finished (follow-up).

Teacher Training. Teachers taking part in the research committed to attending nine hours of teacher training in the term before the PBL unit was to be run. The training was designed to provide background information about Problem-Based Learning, STEM and the Critical and Creative Thinking skills, but most importantly was designed to ensure all the teachers were provided with a clear structure to follow when providing the intervention in their classrooms.

The roles of the teacher and researcher were clearly defined – the researcher was to train the teachers in PBL, to collect the data from the students at three time points over approximately 16 weeks (pre-intervention (TP1), post-intervention (TP2), and follow-up (TP3)) and was to support the teachers during the term by meeting with them as requested by the teachers, and offering advice as called-for. The researcher was also to be available to sit on critique panels as needed. The teacher's role was to facilitate the data collection process, follow the PBL process as taught by the researcher, and to use the problem prompt provided by the researcher (either the hypothetical problem or the industry partner problem). The training was provided individually for each of the four schools, carried out over one or two days (their choice), and covered the following:

- What is STEM, how it fits in the Australian Curriculum, and how to assess it.
- The major features of Problem-Based Learning and essential differences to Project-Based Learning.
- Critical and Creative thinking in the Australian Curriculum and current assessment methods.
- Identifying a problem to investigate – commercial resources, hypothetical problems and industry partner problems.
- Running critiquing and feedback sessions
- The Engineering Design Process.
- Programming a 10-week unit.
- The role of the researcher throughout the project and the data collection process.
- Initial mind mapping of the student problem and planning for the 10-week unit.

Teachers were provided with additional PBL resources (hardcopies and a Dropbox link with articles, videos and weblinks) to supplement the training and advice received from the researcher. A selection of the resources is included in Appendix J-6.

Problem-Based Learning (PBL) and Intervention Training. Teachers chosen for the research were identified in the screening questionnaire as having no background in PBL. It was hoped that this would make it more likely they would follow the PBL process taught to them. This was important as there are so many variations in how PBL is applied in classrooms, and it was hoped that this would reduce the variables between the classroom applications of the intervention. Within the structure though, it was also recognised that the teachers would inevitably personalise the process to some

extent. The teachers' own personalities, the variety of students in their classes, and even the school environment could all have an influence on how the program would be run in each class.

The PBL program, as taught to the teachers, was based on the PEPBL criteria (Barrows, 1986) and structured around the Engineering Design Process, as shown in Figure 19. The 10-week program devised for the teachers to follow during the Term, as shown in Table 44, was based on the researcher's previous experience with numerous classes, and provided guidance on the number of lessons per week that would likely be required for each step to be successfully completed. It was emphasised to the teachers that these lessons were problem lessons, not curriculum lessons, which were to be planned and taught separately. By keeping the lessons separate, students were assured of 'problem-solving' time in the PBL lessons, and a focus on the curriculum requirements for their year level in the curriculum lessons (biology, measurement and money, and geography).

Each step of the PBL process would require the students to use critical and creative thinking skills. For example, once they had been presented with the problem, the students' first steps would be to define the problem and then to identify the issues involved. This requires computational thinking skills, as they break down the problem into its components and begin to recognise the complexity of the problem. This is followed by the research phase, where each student gathers information about one or two of the components of the problem. This uses their information processing skills as they find, organise and share information with other students. Planning solutions requires further collaboration, creative idea generation and the assessment of options. Presenting their ideas to a critique panel (typically the researcher, a staff member such as a Deputy Principal, and the industry partner) is an opportunity for students to justify their proposed solution, and to engage in the iterative design process when they are guided to extend or refine their ideas for their final presentation. The skill of 'Connecting Ideas' is used throughout the PBL process, as students begin to recognise how knowledge they have from one sphere of their life (e.g., from activities done outside school, or from another class) can provide insight into the problem they are investigating in another sphere.

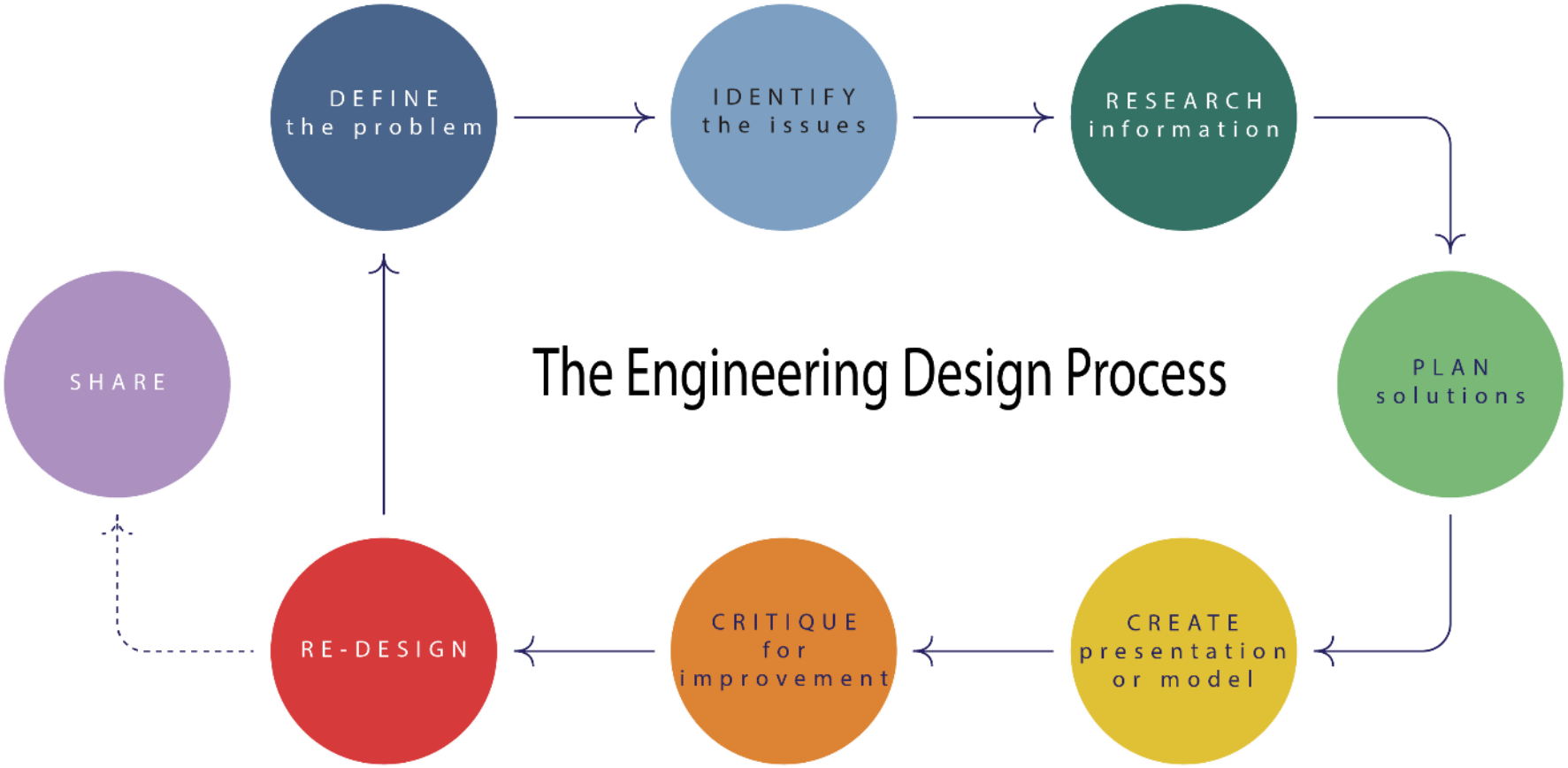


Figure 19

The Engineering Design Process

Table 44

Sample lesson outline for 10 week STEM PBL unit

Week	Engineering Design Process stage	Activity	Notes	Lessons/week
1	Define Identify	• Introduce the problem • Mind mapping (whole class) • Explaining the PBL Engineering Design process to the students	• Industry partner presents to students (where used) • Students define the question to be investigated • Engineering Design diagram used to illustrate the process	1 - 2
2	Research Plan	Students work in research group – topic 1 (3-5 students/group)	• Teacher to determine research groups based on results of mind mapping. Students initially placed in group of personal interest where possible • Instigate method of information sharing between research groups (e.g. Google Docs, MS Teams, etc. All information to be shared) • Commence research	1 - 2
3		Students work in research group – topic 1	• Research	1 - 2
4		Students work in research group – topic 2	• Students move to second research topic (rotation of groups determined by mind mapping results)	1 - 2
5		Students work in research group – topic 2		1 - 2

Table 44 (continued)

Week	Engineering Design Process stage	Activity	Notes	Lessons/week
6	Create	Students form solution groups (preferably 3-5 solution groups, including members of each topic group)	• Students each choose preferred solution from those identified and re-group into solution groups. • Produce prototype, presentation etc. incorporating research carried out on all research topics	2-3
7	Critique	Preliminary presentation of solutions (5 mins max/ group)	• Present to industry partner, member of Leadership team, etc. • Feedback given to students on further information required or specific issues with the proposed solution	1
8	Redesign	Redesign and finalise model and presentation	• Students re-design/ incorporate required changes, complete prototype or presentation	1 - 2
9				1 - 2
10	Share	Final presentations	• Presentations to industry partner/ panel/ other students etc.	1

Note. *Each teacher must create their own Unit Plan which will fit with their school and teaching plan for the term. This sample Unit Plan incorporates the essential STEM PBL components, but does not specify the curriculum lessons (e.g., science, technologies, maths) that will be scheduled by the teacher to support the student work through the STEM PBL unit. These are planned to relate to the topic of the PBL investigation, but are taught through the regular planned curriculum lessons.

It was emphasised to teachers that their role in the learning process during the PBL lessons was to support the students to use these skills, rather than to focus on curriculum outcomes or the solution to the problem.

Although the teachers were provided with a problem for their students to use in the intervention, the training also included information on how to find and effectively use an industry partner, and how to develop hypothetical problems. Teachers were also instructed in running a critique panel and in the focus of feedback to be provided to the students at the critique and final presentations.

The Problem. At the conclusion of their training, the teachers were presented with the problem they were to investigate with their students. To ensure a reasonable balance of participants in the three treatment conditions and to accommodate the industry partner's schedule, the IP problem was given to school D while all three of the other schools were given the hypothetical problem. Only school A had identified classes suitable for use as a control group, as these classes were willing to participate but not to deliver PBL.

The two problems (hypothetical and industry partner) were chosen as they were representative of some of the sorts of problems currently presented to students in primary schools in South Australia. They also covered similar areas of the curriculum, so they were providing students with a similar opportunity for research.

The characteristics of problems recommended in the literature to maintain student interest and motivation (e.g., Blumenfeld et al., 1991, p. 375; Hung, 2016; Jonassen & Hung, 2008) include the following:

- They are authentic, unsolved, current, real-life problems.
- They have time and location proximity.
- They are complex and multifaceted.

Additionally, to be suitable for use for STEM learning in an Australian classroom, problems must be relevant to multiple areas of the Australian Curriculum.

The industry partner problem chosen for this study was:

> Australia currently produces more protein than we need to feed our population, so we export protein to other countries. In 2050 there will be another 2 billion people

> on Earth - how will we feed them and ensure that the solution is tasty, healthy, affordable and good for the environment?

Selected slides from the slide deck which accompanied the presentation have been provided in Appendix J-7. The problem is aligned to the recommended characteristics in the following ways:

- It was a current research problem being investigated by Dr Damien Belobrajdic, a Senior Research Scientist within CSIRO Health and Biosecurity Business Unit.
- The problem was complex and multifaceted, covering multiple curriculum areas and potential research directions.
- It had potential links to the biology curriculum for Years 5 and 6 (Year 5: Living things have structural features and adaptations that help them to survive in their environment (ACSSU043), and Year 6: The growth and survival of living things are affected by physical conditions of their environment (ACSSU094; Australian Curriculum Assessment and Reporting Authority (ACARA), n.d.-c)

There are many commercial resources available which pose hypothetical problems suitable for primary school students to investigate. The hypothetical problem presented to the students in this study was:

> The school is to host a farmers' market, and they are required to produce fresh vegetables fast to sell at the market. A greenhouse will be needed to help the plants grow faster.

This hypothetical problem was taken from a commercially available STEM resource, consisting of engaging problems or challenges linked to outcomes in the Australian Curriculum. The resource was chosen as representative of the many 'STEM problem' resources available commercially. The problem chosen, 'Grow "Fast Food" to sell at Market' from the McMillan STEM Investigations kit for Upper Primary (Forwood, 2018: included in Appendix J-6), was chosen as its topic and the curriculum links were similar to the industry partner problem.

PBL Intervention. The second stage of the intervention was the implementation of the PBL unit, with the identified problem (either hypothetical or IP problem). This is Step Three of the research design illustrated in Figure 18.

School A (control and theory groups) opted to run the intervention in Term 2, 2022, while the other three schools all opted to run it in Term 3, 2022. The data from the control group classes was collected on the same occasions as the data was collected from the theory group classes in

School A. The hypothetical problem was undertaken by two classes from School A, two classes from school B and two classes from school C. Four classes from School D worked on the IP Problem.

In all four schools, the teachers chose to run the Problem-Based Learning lessons with all the students combined as a single group, rather than in their separate classes. This had been recommended in the training, as a larger group provides more opportunities for the students to collaborate and encounter different ideas from their peers. The data collection sessions were run in all the schools with two classes combined as this was more efficient for the schools.

Variations in the way the interventions were run in the different schools are summarised in Table 45. These observations were made from anecdotal evidence provided by the teachers and students in the schools, as the researcher was not present in the schools during the lessons in which the interventions took place.

Data Collection Lesson Time Point One. The first data collection lesson was carried out early in week one of the term in which the intervention was to take place, prior to the students being introduced to the PBL 'problem'. The lesson followed a similar format to that which was delivered and refined in Study Two (described in Chapter 8); however, in Study Three, two measures were presented to the students by the researcher – first, the Critical and Creative Thinking Self-report Measure, followed by the Motivation and Engagement Scale.

The lesson followed the same format as that described for Studies One and Two and commenced with the researcher introducing herself and discussing research in general, and the purpose of collecting this data. This was followed by a short Lego building activity in which pieces of a model were missing. The Lego activity was presented as a challenge to see which of seven different models would be completed first. Working in groups of two or three, the students were required to use strategies to solve the problem of an incomplete model and how to 'do the best job they could' to build the model using alternative pieces. Students were timed, and the completion times were recorded on the board. A brief discussion followed about the patterns that could be seen in the data. The Lego activity was then used as a prompt to discuss topics in the assessment tasks, e.g., what sort of questions one might ask at different times, whether knowledge and experience from an earlier experience was helpful in solving their current problem with the Lego, etc. This activity provided opportunities to discuss the skills used and introduce some of the language used in the assessment task, but also 'broke the ice' between the students and the researcher, who was previously unknown to the students. The Critical and Creative Thinking Self-Report Measure was then presented to

Table 45

Observations of variations in interventions in different schools

School	Involvement of researcher	Comments
A	Initial critique Final critique	The hypothetical prompt was closely followed, with students agreeing on a design for a greenhouse. Students were reluctant to research more than one area of the problem, so the recommended structure of the research was not followed. One solution was therefore presented, with groups of students presenting the component they had researched.
B	No additional involvement	The hypothetical prompt was followed explicitly, with all students working in small groups, developing designs for a greenhouse. Students were expected to produce a model to demonstrate their solution.
C	Initial critique Final critique	The hypothetical prompt was used as motivation but incorporated into the school's philosophy of supporting the local community. Students researched more than one area of the problem and formed 'solution groups' to develop their final solution as recommended. The current unproductive vegetable garden formed the focus of many, but not all of the students' solutions.
D	Initial critique	The students engaged with the industry partner problem and followed the recommended process. Some students researched more than one area of the problem (others only researched one) and formed 'solution groups' to develop their final solution as recommended. Final critique attended by the industry partner

the students with the researcher reading each statement twice – once for the students to hear the whole statement and a second time to confirm what had been said. The students then indicated their response in the numbered boxes on their answer sheet. As in Studies One and Two, students were asked to report on their use of critical and creative thinking skills throughout the year. Throughout the data collection lesson at Time Point One, the students were encouraged to think about and respond to the questions with relation to their work at school in general.

A digital presentation was used to re-enforce the instructions provided by the researcher, and multiple prompts were given for students to ask questions throughout the lesson. 'Brain breaks' in the form of stretches and physical games were inserted into the lesson to ensure the students had

opportunities to refresh themselves and remain stimulated (Ostroff, 2014; Weslake & Christian, 2015).

The Motivation and Engagement Scale followed immediately on completion of the Critical and Creative Thinking Self-Report Measure. As the statements were less complex (more intuitive) each statement was read only once. The students responded by colouring or ticking the appropriate circle corresponding to their chosen response.

Data Collection Lesson Time Point Two. The post-intervention data collection at Time Point Two was carried out in the final week of the term in which the intervention took place, after the students had made their final presentation. On this occasion the Lego task was not used prior to testing. Instead, students were asked explicit questions about work they had been doing during the term, about skills they used when finding and organising information, occasions when they might have made connections to information from different places, etc. The intervention classes were encouraged to think about these questions in relation to the Problem-Based Learning they had recently worked on, while the control group were asked to respond generally about their classwork particularly over the term they had just completed.

Data Collection Lesson Time Point Three. The follow-up data collection at Time Point Three was carried out in week four or week five of the term following the one in which the intervention took place. The students had requested that the Lego task be included, but on this occasion the Lego task was administered following the testing. As in the post-intervention data collection at Time Point Two, prior to the test commencing students were asked explicit questions about work they had been doing during the term, to remind them of skills they used in their daily work.

Throughout the data collection lesson at Time Point Three, the students were encouraged to think about and respond to the questions with relation to their work at school particularly over the term they had completed in which the interventions took place, as well as work they were now doing at school.

Teachers were emailed a link to the digital post-study questionnaire and were followed up to maximise the number of responses.

Control Group. The teachers identified for the control group classes were instructed to continue 'teaching as normal' with their students. The four control group classes had no interaction with the researcher other than during the data collection lessons.

Method for Statistical Analysis

The data collected during Study Three was obtained through the Critical and Creative Self-Report Measure, and the Motivation and Engagement Scale (MES-JS). In addition, non-qualitative, informal feedback was collected through the teacher post-study questionnaire, and through anecdotal feedback from the teachers on the differences in the application of the intervention at different schools. This section summarises the approach to analysis of the data generated at Time Point 1 (TP1), Time Point 2 (TP2) and Time Point 3 (TP3). The results are presented in section 9.3 and their implications are discussed in section 9.4.

All analysis of the quantitative data was carried out using IBM SPSS Statistics (Version 28). Preliminary assumption testing was conducted to check for normality (full results are included in Appendix K, Tables K-1 to K-11). A substantial number of the results of the analysis were not normally distributed. This was observed across all treatment groups and time points, and across all sub-elements. As non-parametric data was therefore indicated, all analysis methods were required to account for this. Levels of skewness and kurtosis are demonstrated in the following examples: e.g., for the theory group at TP3 in Asking Questions Foundation level skewness was -5.428 (*SE* = .255) and kurtosis was 31.656 (*SE* = .506). Motivation and engagement scores were also not normally distributed for all treatment groups and time points, and across all higher order factors, e.g., for the industry partner group at TP2 in Total Positive Motivation the skewness was -1.345 (*SE* = .297) and kurtosis was 2.047 (*SE* = .586). As this data was therefore considered non-parametric, the choice of data analysis techniques reflected this.

The Motivation and Engagement Scale data was examined for internal consistency (Laerd Statistics, 2015a; Pallant, 2013). This included comparisons for all students at each time point, for the four schools at each time point, and for EALD and non-EALD students for each time point, to ensure that the origins of any issues were fully identified. A consistently low Cronbach's alpha was indicated for Disengagement statement three ('I'm not involved in things like class activities and class discussion at school') at TP1 (pre-intervention) for School B, where Cronbach's alpha increased from 0.35 to 0.64 on removal of the item (Disengagement statement 3 – the level reported in previous research is 0.72; Martin, 1999-2022). Comparisons for all students combined at each time point (results reported in Appendix K, Table K-12), for the four schools analysed separately at each time point (results reported Appendix K, Tables K-13 to K-15), and for EALD and non-EALD students at each time point (results reported Appendix K, Table K-16 to K-18) indicated that removal of the item effected a general improvement in Cronbach's alpha across all areas of comparison and did not

disadvantage any group. It was therefore determined to remove this statement from the analysis which followed.

To test the hypothesised difference in change in levels of skills between the control group and PBL group (Hypotheses 3a(i), 3a(ii), 4a(i) and 4a(ii)), the two intervention treatment groups (the theory PBL group and the IP PBL group) were analysed together as a PBL group and compared with the control group. This process was followed with all forms of analysis where comparisons were being made first between the control and PBL groups, and then between the theory PBL and IP PBL groups.

Critical and Creative Thinking. Two forms of analysis were carried out on the data to test Hypothesis 3a(i) ('PBL results in both an increased *range*, and an increased *level* of critical and creative thinking skills compared to a control group.'), Hypothesis 3a(ii) ('Positive gains in the critical and creative thinking skills will persist over time for the PBL students.'), Hypothesis 3b(i) ('An industry partner problem results in both an increased *range*, and an increased *level* of critical and creative thinking skills compared to a hypothetical problem.'), and Hypothesis 3b(ii) ('Positive gains in the critical and creative thinking skills will persist over time for both the industry partner PBL students and the theory PBL students.').

The initial analysis was carried out to determine if there were differences between the control and PBL groups, in the levels of skills being used by the students prior to the commencement of the intervention. This was to determine whether the starting point of the skill level was the same for all the students, or if some groups were already using some of the skills more than other groups of students. As the focus of the hypotheses was on the change that would be observable in the students' level of skill use after the intervention, it was useful to know whether the starting point for the levels of skills was the same for all the students at Time Point One. A Mann-Whitney U test – a rank-based non-parametric alternative to the independent-samples t-test – enables differences between independent groups to be identified, and was therefore used to analyse the data.

The second form of analysis was carried out with a one-way repeated measures analysis of variance (ANOVA) equivalent, suitable for use with non-parametric data. The nature of this study – longitudinal over three time points, with a dichotomous dependent variable – made analysis with Cochrane's *Q* test (Cochran, 1950) most appropriate. This analysis was carried out to determine the level and statistical significance of any change in the levels of skills between the control group and the PBL groups that occurred across the three time points and between the theory PBL and the IP PBL groups. As the indicator of success for the Critical and Creative Thinking Self-Report Measure

was that the student 'showed use of their skills', it was determined that this was only effectively demonstrated by the unequivocable response of '*Yes, I do this'*. A response of '*Unsure*' did not provide enough certainty that the skill was now being used by the student. The three-point categorical scale was therefore collapsed to a dichotomous variable with the two categories '*No*' and '*Unsure*' combined forming a two-point categorical scale with the responses '*Yes, I do this'* and '*No, I don't do this*'. This enabled a clear and unambiguous indication of students who showed the use of skills. Sample size was examined and found adequate to use the χ^2-distribution approximation (Tate & Brown, 1970). Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons (Laerd Statistics, 2017a).

Motivation and Engagement. Motivation and Engagement Scale data generated using the approach developed by Martin (1999-2022), can be analysed at two levels – using the four higher order factors of positive motivation, negative motivation, positive engagement and negative engagement, or by analysing the 11 first order factors of self-belief, valuing, learning focus, planning, task management, persistence, anxiety, uncertain control, failure avoidance, self-sabotage and disengagement (Liem & Martin, 2012). As the emphasis of Research Questions 4a and 4b (RQ 4a: 'What is the impact of PBL on student motivation and engagement?'; RQ 4b: 'What is the impact of using an industry partner problem in STEM PBL on student motivation and engagement?') was on the more general concept of motivation and engagement rather than the more fine grained elements, the individual factors were combined and analysed at the higher order factor level as described by Martin (1999-2022).

To explore Hypotheses 4a(i) and 4a(ii) and determine if there were differences in the levels of motivation and engagement between the control group and the two combined PBL groups, pairwise comparisons were performed using a Friedman Test with a Bonferroni correction for multiple comparisons. The Friedman test is the non-parametric alternative to a one-way repeated measures ANOVA, suitable for use with an ordinal dependent variable (Laerd Statistics, 2015b).

A Friedman test analysis was then carried out to compare the results of the theory PBL group with the IP PBL group in response to Hypotheses 4b(i) and 4b(ii).

9.3. Results of Statistical Analysis

Study 3 sought to provide robust statistical evidence to support the following hypotheses:

- Hypothesis 3a(i): PBL results in both an increased *range*, and an increased *level* of critical and creative thinking skills compared to a control group.

- Hypothesis 3a(ii): Positive gains in the critical and creative thinking skills will persist over time for the PBL students.
- Hypothesis H 3b(i): An industry partner problem results in both an increased *range*, and an increased *level* of critical and creative thinking skills compared to a hypothetical problem.
- Hypothesis H 3b(ii): Positive gains in the critical and creative thinking skills will persist over time for both the industry partner PBL students and the theory PBL students.
- Hypothesis 4a(i): PBL will result in an improvement in motivation and engagement compared to a control group.
- Hypothesis 4a(ii): Positive gains in motivation and engagement will not be retained beyond the PBL intervention.
- Hypothesis 4b(i): An industry partner problem will result in a greater improvement in motivation and engagement compared to a theory problem.
- Hypothesis 4b(ii): Positive gains in motivation and engagement will not be retained beyond the PBL intervention.

This section summarises the outcomes of statistical analysis of the data generated at TP1, TP2 and TP3. The implications of the results are discussed further in section 9.4 below.

Critical and Creative Thinking Statistical Results

The levels of students' skill use before intervention were compared to determine whether the two groups of students starting points were similar at Time Point 1. The comparison was between the control and PBL groups and was made for all the year level statements for all sub-elements. The mean rank of use-of-skills scores was not statistically significantly different between the groups for the majority of the categories of skills[28]. This indicates that statistically, based on the mean scores, both groups of students were starting from the same level of use of almost all of the skills at all year levels. The exceptions were for Asking Questions at the Foundation level (AQF) and Evaluating a Plan at the Year 2 level (EP2). For AQF the skill-use scores for the control group (mean rank = 130.21) were statistically significantly higher than for the PBL group (mean rank 119.87), $U = 5164$, $z = -.965$, $p = .049$ and for EP2 the skill-use scores for the PBL group (mean rank 127.32) were statistically significantly higher than for the control group (mean rank 106.46), $U = 6574.5$, $z = 2.658$, $p = .008$. Neither of these skills at the levels described demonstrated statistically significant levels of change

[28] The results of the Mann-Whitney U test indicated that distributions of the skill-use scores were not similar, as assessed by visual inspection of a box plot, therefore mean ranks were compared rather than medians.

after the intervention, so the difference in the levels of starting points is not considered to have had any effect on the analysis or discussion which follows.

To test the hypothesised difference in change in levels of C&CrT skills between the control group and PBL group (Hypotheses H3a(i)and 3a(ii)), the two intervention treatment groups (the theory PBL group and the IP PBL group) were analysed together as a PBL group and the results of analysis with Cochrane's Q was compared with the control group. The full set of Cochrane's Q test results can be found in Appendix K Tables K-19 to K-22.

The two comparisons of change in use of skills that are of particular interest to test the hypotheses are the change from before the intervention commenced (TP1) to immediately following the end of the intervention at Time Point 2 (H3a(i) and H3b(i)), and the change from TP2 to half-way through the following term at Time Point 3 (H3a(ii) and H3b(ii)). The first change, TP1 – TP2, signals the immediate effect of the PBL intervention on the students' application of skills. The second change, TP2 – TP3, indicates whether the students continued to apply the new skills or levels of skills to their regular lessons, i.e., did changes that were the result of the activities being undertaken during the intervention continue into the work students were doing in the following term.

A summary of the significant results of the analysis for the control and PBL group results is shown in Table 46. A summary of the significant results of the analysis for the theory PBL and the IP PBL group results is shown in Table 47. The tables include directional arrows to indicate whether the statistically significant change was an increase in skill use or a decrease in skill use.

Control Versus PBL Group. Change from TP1 to TP2: H3a(i). As shown in Table 46, the control group indicated significant positive change in two skills over the period of the 10-week term – 'Asking Questions' at a Year 2 level (AQ2) and 'Connecting Information' at a Year 2 level (CI2). In contrast, the PBL group indicated positive change in four skills over the same period – one of the skills the same as the control group, AQ2. The other three skills were at the Foundation and year 4 levels. This demonstrates that there was a greater variety of critical and creative thinking skills used by the PBL group, at a variety of 'levels' at the end of a 10-week PBL intervention, hence Hypothesis 3a(i) is supported.

Control Versus PBL Group: Change from TP2 to TP3: H3a(ii). It was expected that students would continue to use the skills into the following term. With the control group the levels of all but one of the skills remained constant at TP3 i.e., the level of skill reported at TP2 had not statistically significantly changed at TP3. The use of CI6 increased to almost the same level it had been reported at prior to the intervention term. The same result was observed with the PBL group. The PBL group

did indicated a reduction in the use of Creating New Ideas at a year 2 level (CR2) at TP3 however this was the only skill that returned to a pre-intervention level. Hypothesis 3a(ii), that the skills would persist at the new levels, is therefore supported.

Theory PBL Group Versus Industry Partner PBL group: Change from TP1 to TP2: H3b(i). As shown in Table 47 the theory PBL group indicated significant positive change in two skills over the period of the 10-week term – 'Asking Questions' at a Year 2 level (AQ2) and 'Finding and Organising Information' at a Year 4 level (FO4). In contrast, the IP PBL group indicated positive change in five skill levels over the same period, including the skills of 'Connecting Information', 'Creating New Ideas' and 'Finding and Organising Information', at levels Foundation, Year 2 and Year 8. This demonstrates that there was use of a greater variety of critical and creative thinking skills by the IP group, at a variety of 'levels', hence Hypothesis 3b(i) is supported.

Theory PBL group versus Industry Partner PBL group: Change from TP2 to TP3: H3b(ii). Once again, consistent level of use of the skills expected after the conclusion of the PBL intervention was observed. With the Theory PBL group the levels of all but one of the skills remained unchanged from TP2 to TP3 – FO4 dropping back to a lower level of use than recorded at TP1. The industry partner PBL group indicated an increase of CI6, which had fallen in all the groups at the end of the intervention term. This was a strongly significant increase *(p* = .001). All other skills remained at the same level of use as they had been at the end of the intervention term, hence Hypothesis 3b(ii) is supported.

Table 46

Analysis of critical and creative thinking comparing control and PBL groups

Treatment Group	*n*	Sub-element[#]	Year level statement	% 'Yes' response			Pairwise Comparisons and direction of change *p* (Adj. sig.)*	
				TP1 Pre-intervention	TP2 Post-intervention	TP3 Follow-up	TP1-TP2	TP2-TP3
Control	64	AQ	2	46.9	62.5	59.4	.035 ↑	
		CI	2	64.1	75.0	70.3	.019 ↑	
			6	85.9	46.9	81.3	<.001 ↓	<.001 ↑
PBL group	193	CI	2	65.3	71.5	66.3	.009 ↑	
			6	78.2	54.9	71.0	<.001 ↓	<.001 ↑
		CR	F	63.2	69.9	65.3	.045 ↑	
			2	77.7	75.6	66.8		.017 ↓
		FO	F	50.8	61.1	54.4	<.001 ↑	
			4	58.0	62.2	54.9	.028 ↑	

Note. Blank cells indicate no statistically significant change was registered. [#] AQ = Asking Questions, CI = Connecting Information, CR = Creating New Ideas, FO = Finding and Organising Information. * Significance values have been adjusted by the Bonferroni correction for multiple tests.

Table 47

Analysis of critical and creative thinking comparing theory and industry partner groups

Treatment Group	*n*	Sub-element[#]	Year level statement	% 'Yes' response			Pairwise Comparisons and direction of change *p (Adj. sig.) **	
				TP1 Pre-intervention	TP2 Post-intervention	TP3 Follow-up	TP1-TP2	TP2-TP3
Theory PBL group	106	AQ	2	46.2	58.5	57.5	.019 ↑	
		CI	6	84.0	58.5	64.2	<.001↓	
		FO	4	63.2	69.8	54.7	.056 ↑	.056 ↓
IP PBL group	87	CI	2	55.2	69.0	65.5	<.001↑	
			6	71.3	50.6	79.3	.014 ↓	.001↑
			8	49.4	52.9	58.6	.029 ↑	
		CR	F	49.4	63.2	57.5	.011 ↑	
		FO	F	34.5	63.2	57.5	<.001↑	
			2	34.5	46.0	54.0	.022 ↑	

Note. Blank cells indicate no statistically significant change was registered. [#] AQ = Asking Questions, CI = Connecting Information, CR = Creating New Ideas, FO = Finding and Organising Information. * Significance values have been adjusted by the Bonferroni correction for multiple tests.

Motivation and Engagement Statistical Results

The two comparisons of change of level of motivation and engagement that are of particular interest to test the hypotheses are the change from before the intervention commenced (TP1) to immediately following the end of the intervention at Time Point 2 (H4a(i) and H4b(i)), and the change from TP1 to half-way through the following term at Time Point 3 (H4a(ii) and H4b(ii)). The first change, TP1 – TP2, signals the immediate effect of the PBL intervention on the students' levels of motivation and engagement. The second change, TP1 – TP3 provides information on whether the levels of motivation and engagement returned to pre-intervention levels, as hypothesised.

Higher Order Factors. To test the hypothesised difference in change in levels of motivation and engagement between the control group and Problem-Based Learning group (Hypotheses 4a(i) and 4a(ii)), the Friedman test was conducted on the results of the higher order factors obtained from two groups: the two intervention groups together (PBL group), and the control group. The full set of Friedman test results can be found in Appendix K Table K-23 and K-24. A summary of the significant results of the analysis are shown in Table 48. The results of the analysis of higher order factors for the theory PBL group and the IP PBL group are shown in Table 49.

Control Versus PBL Group. Change from TP1 to TP2: H4a(i). An improvement in motivation and engagement was expected to be demonstrated by both increases in positive motivation and engagement and falls in negative motivation and engagement. In line with hypothesis 4a(i), improvement was expected to be demonstrated over the period of the intervention due to the positive influence of the use of PBL. Table 48 shows that the PBL group indicated the only significant change immediately following the intervention – a fall in negative motivation. One area of negative motivation registered a statistically significant reduction, but no areas of engagement registered any change. Hypothesis 4a(i), that 'PBL will result in an improvement in motivation and engagement compared to a control group' was therefore supported.

Control Versus PBL Group. Change from TP1 to TP3: H4a(ii). Following the end of the PBL intervention, students were expected to react negatively to the change back to 'teaching as normal', with decreases in positive motivation and engagement, and increases in negative motivation and engagement. The PBL students indicated statistically significant falls in all four of the higher order factors for motivation and engagement from TP1 to TP3. The fall in negative motivation was particularly significant *(p* = .002), and this is discussed further below. The increase in positive engagement registered by the control group ran counter to the results for the PBL group, who registered a fall in this factor for the same period. Further investigation showed that this increase in

Table 48

Analysis of motivation and engagement higher order factors comparing control and PBL groups

Higher Order factor	Treatment Group	$\chi^2(2)$	Median			Pairwise Comparisons and direction of change *p (Adj. sig.)**	
			TP1 Pre-intervention	TP2 Post-intervention	TP3 Follow-up	TP1-TP2	TP1-TP3
Positive Motivation	Control		4.4167	4.5000	4.3750		
	PBL	12.823	4.4167	4.4167	4.2500		.018 ↓
Positive Engagement	Control	7.988	3.6667	4.0000	3.8750		.037 ↑
	PBL	9.166	3.7500	3.7917	3.6667		.029 ↓
Negative Motivation	Control		3.2500	3.1667	3.0417		
	PBL	13.624	2.7500	2.6667	2.4167	.022↓	.002 ↓
Negative Engagement	Control		1.8571	1.5714	1.5714		
	PBL	9.131	1.7143	1.6429	1.5714		.022 ↓

Note. Blank cells indicate no statistically significant result was obtained. Control group $n = 64$, PBL group $n = 193$. *Significance values given where $p < .050$. These significance values have been adjusted by the Bonferroni correction for multiple tests.

Table 49

Analysis of motivation and engagement higher order factors comparing theory and IP PBL groups

Total Higher Order factors	Treatment Group	$\chi^2(2)$	Median			Pairwise Comparisons and direction of change *p* (Adj. sig.)*	
			TP1 Pre-intervention	TP2 Post-intervention	TP3 Follow-up	TP1 –TP2	TP1 - TP3
Positive Motivation	Theory PBL	8.255	4.3333	4.4167	4.2500		
	IP PBL		4.4167	4.4167	4.2083		
Positive Engagement	Theory PBL	7.350	3.8333	3.5833	3.6667		.029 ↓
	IP PBL		3.7083	3.8333	3.6667		
Negative Motivation	Theory PBL	9.595	2.9167	2.7917	2.7083	.044↓	.021↓
	IP PBL		2.5000	3.1667	3.0417		

Note. Blank cells indicate no statistically significant result was obtained. Theory PBL group *n* = 106, IP PBL group *n* = 87. *Significance values given where *p* < .050. These significance values have been adjusted by the Bonferroni correction for multiple tests.

positive engagement for the control group was not reflected in the results of the two PBL classes who attended the same school as the control group. The classroom teachers were unable to suggest any reasons for this increase in positive engagement. Hypothesis 4a(ii) stated that positive gains in motivation and engagement would not persist beyond the PBL intervention. Hypothesis 4a(ii) was therefore not supported.

Theory PBL Group Versus IP PBL Group. Change from TP1 to TP2: H4b(i). Comparing the results between the theory PBL and the IP PBL groups in Table 49, the theory PBL group indicated the only statistically significant change during the period of the intervention. Although it was expected to see increases in positive motivation and engagement and decreases in negative motivation and engagement, the only change reported was a reduction in negative motivation. Hypothesis 4b(i), that 'An industry partner problem will result in a greater improvement in motivation and engagement compared to a theory problem' was therefore not supported.

Theory PBL Group Versus IP PBL Group. Change from TP1 to TP3: H4b(ii). A comparison between the theory PBL group and the IP PBL group, indicates the only significant changes from TP1 to TP3 have occurred in the theory PBL group. Decreases occurred in both positive engagement and in negative motivation, reducing the levels of both factors to below the levels recorded at TP1. As a decrease in negative motivation is considered to be an 'improvement' in motivation, Hypothesis 4b(ii), that positive gains in motivation and engagement will not persist beyond the PBL intervention, was therefore not supported.

9.4. Discussion

The hypotheses proposed in response to Research Question 3 and 4, and results of the analyses are attached in Table 50 and 51.

New Insights into Critical and Creative Thinking Skills: H 3a(i) and H 3b(i)

The increased use of critical and creative thinking skills in the PBL group supports the claims in the literature that Problem-Based Learning has a positive effect on student skills. Furthermore, the even greater increases in the IP PBL group are conclusive evidence that there is a significant benefit to using an industry partner and their problem in STEM PBL. The results of the study support the informal hypothesis that stimulated the PhD project. My observation that students benefitted from the presence of an industry partner was accurate, and this study has provided evidence in

Table 50

Research Question 3 and sub-questions 3a and 3b with associated hypotheses and results

			Research Questions and Hypothesis	Result
RQ 3	How do the characteristics of STEM PBL affect the use of critical and creative thinking skills during learning?			
	RQ 3a	‘What is the impact of PBL on student critical and creative thinking skill application?’		
		H 3a(i)	PBL results in both an increased *range*, and an increased *level* of critical and creative thinking skills compared to a control group.	Supported
		H 3a(ii)	Positive gains in the critical and creative thinking skills will persist over time for the PBL students.	Supported
	RQ 3b	‘What is the impact of using an industry partner problem in STEM PBL on critical and creative thinking skill application?’		
		H 3b(i)	An industry partner problem results in both an increased *range*, and an increased *level* of critical and creative thinking skills compared to a hypothetical problem.	Supported
		H 3b(ii)	Positive gains in the critical and creative thinking skills will persist over time for both the industry partner PBL students and the theory PBL students.	Supported

Table 51

Research Question 4 and sub-question 4a and 4b, with associated hypotheses and results

				Result
		Research Questions and Hypothesis		
RQ 4	How are student motivation and engagement measured in primary school classrooms during learning activities?			
	RQ 4a	'What is the impact of PBL on student motivation and engagement?'		
		H 4a(i)	PBL will result in an improvement in motivation and engagement compared to a control group.	Partially supported
		H 4a(ii)	Positive gains in motivation and engagement will not persist beyond the PBL intervention.	Supported
	RQ 4b	'What is the impact of using an industry partner problem in STEM PBL on student motivation and engagement?'		
		H 4b(i)	An industry partner problem will result in a greater improvement in motivation and engagement compared to a theory problem.	Not supported
		H 4b(ii)	Positive gains in motivation and engagement will not persist beyond the PBL intervention.	Not supported

support of that observation. Additionally, the study has also provided evidence that the use of PBL as a pedagogical approach enables students to apply critical and creative thinking skills during their learning.

Hypotheses 3a(i) and 3b(i) both focus on the *range* of skills and the *level* of the skills being used, as indicators of success. The range of skills, i.e., the type of different skills that are being used will be examined first, followed by an examination of the levels of the skills that are being used by students.

Type of Skill – Effect of Problem-Based Learning. As all the students – including the control group – were attending school and engaged in learning activities over the period of the research, it should be expected that there would be an increase in the use of at least some of the skills. The use of PBL in the intervention, however, has enabled students to use more of the skills than the students in the control group. The structure of the PBL lessons provides opportunities for the students to be stretched and challenged at various stages – from the problem definition stage right through to the presentation of their final solution. By challenging the students within the carefully structured Engineering Design framework of a STEM PBL Unit, they were supported to use their skills beyond the use of skills reported by students being taught through less 'active' lessons. PBL has long been touted as enabling the development of skills (e.g., Hmelo-Silver, 2004; Shepherd, 1999; Smith et al., 2022), but as we have seen, this claim has proved difficult to demonstrate (see Chapter 3). At the commencement of this research the decision was made to focus on the application of skills rather than their development. There were two main reasons for this decision: 1) the measure was only identifying the skills currently used by students. Students were not asked if they had *ever* used the skills described. It could not be determined therefore whether the students had already developed (but were choosing not to use) skills at more complex levels than they were currently using. 2) it was assumed that students would only report the use of skills that were explicitly stimulated by the type of learning they were engaged in. This research has clearly demonstrated the increased *use* of critical and creative thinking skills by primary school students engaged in STEM PBL. Further, however, the research indicates the *development* of the skills, as the students continued to report their use of the skills, even after they had returned to 'teaching as normal'. If, as a result of being taught with a PBL approach, students both develop and apply more skills than they would otherwise, this approach should be promoted for use in schools more.

The range of skills used by PBL students was double the number used by the students in the control group. As well as 'Asking Questions' and 'Connecting Information', students in the PBL group

used ‘Creating New Ideas’ and ‘Finding and Organising Information’ at significantly greater levels. ‘Evaluating a Plan or Idea’ was the only skill not used – including by the control group. The structure of the PBL program run in the classes did not incorporate a feedback-loop for the students to reflect on their practice — something also apparently missing from the classes engaged in by the control group. This lack of opportunity to use those skills is clearly reflected in the self-report scores from the students.

Type of Skill – Effect of an Industry Partner. Furthermore, students who had a genuine industry partner problem to investigate demonstrated an increase in the use of a greater variety of skills than the students engaged in a hypothetical problem. The theory group reported the use of only two skills – ‘Asking Questions’ and ‘Finding and Organising Information’. Both skills are used extensively in PBL. The additional skills used by the IP PBL students were ‘Connecting Information’ and ‘Creating New Ideas’.

‘Connecting Information’ is the ability to connect ideas beyond the confines of the lesson. This is an extremely valuable skill, as it enables greater understanding of complex issues, and the development of more imaginative solutions. It is also a difficult concept for some students to accept. The inability to see connections beyond the lesson is often described as ‘siloing’ – the inability or refusal to acknowledge that maths is done anywhere outside a maths lesson, or that science has any influence on their thinking outside the science lesson, and that nothing they do at school has any relevance at home, or vice versa. The skill of ‘Connecting Information’ is not intentionally taught to the students, nor does it receive a particular focus in the PBL process, but by encouraging the teachers to teach related curriculum lessons during the same term as the PBL, the hope is that students will recognise the relevance of these lessons to the problem they are investigating. It is up to the students to make the connections, and the increase in the use of these skills indicates that they have done exactly that. By presenting students with problems that go beyond the walls of the school – particularly problems that are brought to them by a complete stranger – it appears that the use of this skill of being able to connect ideas from other lessons, times and places can be stimulated.

Although the teachers were aware that Study Three was focussing on the use of critical and creative thinking skills by students, it was of interest that no teachers commented on the use of skills by students in their feedback to the researcher through the post-study questionnaire, or during informal discussions with the researcher. This lack of focus on the skills may illustrate a lack of awareness of the use of skills, accounting for the poor level of agreement between the teachers and students in Study Two (Chapter 8).

Level of Skills. It was notable that despite the students working at a Year 5 and 6 level, it is not generally in those year level skills that increases were observed. Most of the skills the students reported an increase in were the Foundation and year 2 level skills. The levels of skills used by the students further supports the earlier discussion (Chapter 7 discussion, Study One) about the way different skills develop and are used, and the mismatch between the current critical and creative thinking continuum and the stages students are at in their schooling. Some skills seem to remain relevant and are used at all levels depending on the requirements of the task at hand, e.g., with 'Asking Questions', while for other skills some of the less complex levels may be subsumed as later, more complex levels of the skill are mastered, e.g., for 'Evaluating a Plan or Idea'. The descriptions of the skills in the current C&CrT continuum, and in the measure developed for this study, require deeper examination. The year level labels applied to the descriptions of the skills in the continuum do not appear to align with the year groups of the students. Further research is required to determine the most accurate descriptions of the skills, and to represent the way they are actually used by students during learning.

The Long-Term Effect of PBL on Critical and Creative Thinking: H 3a(ii) and H 3b(ii).

That Hypotheses 3a(ii) and 3b(ii) were supported is evidence that the positive effects of learning through PBL, i.e. the increased use of skills, does not end with the withdrawal of this pedagogical approach to teaching and learning. This is a very positive result, as it indicates that once students have used a new skill or a skill at a higher level, they are more likely to continue to use it at that level. This is also taken as evidence that PBL is enabling students to *develop* these new skills. Testing the students further in a longer-term longitudinal study would provide greater insights into how effective this approach is with primary school students in the long term.

New Insights into Motivation and Engagement: H 4a(i), and H 4b(i)

The literature makes claims of increases in motivation and engagement for students taking part in PBL (e.g., Belland et al., 2013; Dochy et al., 2003; Stepien & Gallagher, 1993). Further, anecdotal evidence from teachers also supports this view, e.g. one teacher in School C reporting "Engagement was terrific". It was expected therefore that increases would be seen in both positive motivation and engagement *during* the term the PBL intervention occurred. It was also expected that decreases would be observed in negative motivation and engagement. This was not the general trend observed. Only one higher order factor registered a significant change during the PBL intervention – negative motivation.

A reduction in negative motivation is encouraging, as it is characterised by students exhibiting counterproductive behaviours such as self-sabotage and failure avoidance, and suffering from anxiety. To try to understand the possible effect of these results, the data was further analysed, at the level of the first order factors.

The motivation and engagement survey (MES-JS) utilised to collect the motivation and engagement responses from the students, is capable of being analysed at the more granular level of 11 first order factors – three each for positive motivation and engagement and negative motivation, and two for negative engagement. The results of this analysis are shown in Table 52 for a comparison of the Control and PBL results, and Table 53 for a comparison of the theory PBL group and IP PBL group results.

The reduction in negative motivation was fully accounted for by a fall in the first order factor of 'uncertain control' indicating that these students felt they had more control over their learning during the PBL. As 'uncertain control' is negatively correlated with achievement, mastery motivation, competence evaluation, IQ test scores and teachers' ratings of competence (Martin, 2001), this reduction can be seen as a real benefit of using PBL. Although this was the only change in the motivation and engagement scores covering the period of the intervention (TP1 – TP2), further reductions were observed at TP3.

The Long-Term Effect of PBL on Motivation and Engagement: H 4a(ii) and H 4b(ii).

At the level of higher order factors (Table 48), positive motivation and positive engagement both fell at TP3, but only for the PBL students. This fall is taken to indicate that the PBL students were negatively affected by the PBL coming to an end. It is possible that, being unfamiliar with PBL, while engaged with the PBL process the students found it more challenging than engaging. When it ended however and they returned to 'teaching as normal', they found themselves less engaged. For students who are usually taught with a more traditional, less 'active', less student-driven pedagogical approach, Problem-Based Learning can be challenging. For some students, the opportunity to use their own initiative about choosing what they believe is important in their investigation may motivate them. For others, this level of free choice may de-incentivise them. Comments from the teachers in the post-study questionnaire reflected both attitudes; the comment from a teacher from School A describing some students as "struggling", while others "did well finding their own problems and identifying solutions". Examining the first order factors (Table 52) this decrease in positive engagement is accounted for by a strongly significant ($p < .001$) fall in 'task management'. This

Table 52

Analysis of motivation and engagement first order factors comparing control and PBL groups

Higher Order Factors	First Order Factors	Treatment Groups	χ2(2)	Median			Pairwise Comparisons and direction of change *p (Adj. sig.)**	
				TP1 Pre-intervention	TP2 Post-intervention	TP3 Follow-up	TP1-TP2	TP1-TP3
Positive Engagement	Total Task Management	Control	4.881	80.00	85.00	85.00		
		PBL	17.206	80.00	80.00	75.00		< .001↓
	Total Planning	Control	14.098	65.00	80.00	80.00	.035↑	.035↑
		PBL	2.034	70.00	70.00	70.00		
Negative Motivation	Total Uncertain Control	Control	.148	60.00	50.00	55.00		
		PBL	17.716	50.00	47.50	45.00	.012↓	< .001↓
	Total Failure Avoidance	Control	1.014	65.00	60.00	65.00		
		PBL	8.626	50.00	45.00	42.50		.024↓
Negative Engagement	Total Self Sabotage	Control	2.456	40.00	35.00	35.00		
		PBL	23.780	35.00	35.00	30.00		< .001↓

Note. Control group *n* = 64, PBL group *n* = 193. *Significance values given where *p* < .05. These significance values have been adjusted by the Bonferroni correction for multiple tests.

Table 53

Analysis of motivation and engagement first order factors comparing theory and industry partner (IP) PBL groups

Higher Order Factors	First Order Factors	Treatment Groups	χ2(2)	Median			Pairwise Comparisons and direction of change *p (Adj. sig.)**	
				TP1 Pre-intervention	TP2 Post-intervention	TP3 Follow-up	TP1-TP2	TP1-TP3
Positive Engagement	Total Task Management	Theory PBL	8.972	80.00	80.00	75.00		.024↓
		IP PBL	8.385	80.00	80.00	80.00		.026[a]↓
Negative Motivation	Total Uncertain Control	Theory PBL	13.164	50.00	50.00	50.00	.005[a] ↓	.022[a] ↓
		IP PBL	9.030	45.00	45.00	40.00		.024 ↓
Negative Engagement	Total Self Sabotage	Theory PBL	11.898	40.00	35.00	30.00		.010↓
		IP PBL	12.395	35.00	35.00	30.00		.008↓

Note. Theory PBL group *n* = 106, IP PBL group *n* = 87. *Significance values given where *p* < .05. These significance values have been adjusted by the Bonferroni correction for multiple tests. [a] This is recorded as a fall in score although it is not reflected in the median.

appears to support the argument that students enjoy having some autonomy over their learning. Task management ('When I do homework, I get organized so I can do it well') is described by Martin (1999-2022) as being an outward reflection of self-regulation. Further information from the students is required to understand the reason for this change, however the alteration from a student-led style of learning to a more structured style of lessons may have caused the students to feel a reduction in control over their learning, which may have manifested in this reduced engagement.

Negative motivation and negative engagement also indicated significant change, all showing a decrease for the PBL group at TP3. This is a very positive result for PBL, demonstrating the longer-term positive effects of this pedagogical approach. The students engaged in PBL showed reductions in 'uncertain control' ('I don't think I have much control over how well I do in my schoolwork'), 'failure avoidance' ('The main reason I try at school is because I don't want people to think bad things about me) and 'self-sabotage' ('Sometimes I don't try hard at school so I can have a reason if I don't do well') The results demonstrated by the PBL students were further supported by similar results from the theory PBL and IP PBL groups, shown in Table 53.

The relationship between PBL and motivation and engagement is more complex than has been possible to explore in this study. Further research which could provide more detailed information on this relationship could:

- a) measure motivation and engagement at more carefully timed intervals, e.g., immediately following the industry partner's presentation to the students, and immediately preceding or following the industry partner's critique. An alternative would be to use an experience sampling method (ESM) which prompts participants to provide real-time or near-real-time data about their experiences (Hektner et al., 2007);
- b) compare the results of Study Three with a follow-up study on the same participants during and after a second intervention with PBL. It is possible that the very different teaching approach demonstrated by PBL caused the students to be uncomfortable to the point where they did not respond very positively to the motivation and engagement scale. Being presented with a PBL intervention a second time would possibly be less intimidating for the students, and their familiarity with the process might produce a greater recognition of changes in their motivation and engagement; and

- c) interview the students to capture the nuances of their responses to the questions in the MES-JS through qualitative data.

Measuring students' levels of motivation and engagement was expected to provide a clear indication of whether Problem-Based Learning, with its active challenges, provided greater levels of motivation and engagement. Although the results were not as straightforward as expected from the literature, the analysis of motivation and engagement at the level of first order factors certainly provides support for the statements that PBL has a positive effect on motivation and engagement.

It is also worth considering whether motivation and engagement are the most relevant characteristics to be measuring in primary school students. Are excitement and interest, for example, effectively captured by measures of motivation and engagement? The comment from a teacher in School D perhaps provides examples of other important characteristics which further studies might examine, e.g., "independence and a sense of wonder".

The Industry Partner, and Motivation and Engagement. Hypotheses 4b(i), which suggested that there would be more improvement in motivation and engagement with the use of an IP problem than a hypothetical problem, was not supported. As previously stated, the relationship between PBL, and motivation and engagement, is complex. That extends to the component of PBL that is the 'Problem'. Further investigation should be made to understand the relationship between the effect of different types of problem and motivation and engagement. The industry partner problem that was presented to the students in this study was significantly more complex than the hypothetical problem. Although the two problems set for the students (the IP and hypothetical problem) were intentionally chosen for their similarity in subject content (Biological sciences – food, growing food (plants and animals); Mathematics – costs of production; Geography – effects of climate on living things), they were necessarily different. The hypothetical problem was devised to ensure that sufficient information and direction were provided to steer the students' investigation in a particular direction. The industry partner problem was open-ended, incorporating potential topics the students – and teachers – had little understanding or previous experience with. With the added complexity of the industry partner problem, it is perhaps more surprising that the students in the IP PBL group maintained the same level of motivation and engagement during the intervention, as the control group who were not being challenged by unfamiliar methods, demands, and a 'stranger in the room'. Perhaps a further benefit of the IP problem (or the industry partner) is that it provides a sort of 'buffering' effect. Despite the increased challenge with the IP problem, the connection with the industry partner provides such an increased level of motivation/ engagement/ excitement/ interest that the negative challenges of the complexities are over-ridden. Clearly further

investigation is required to determine whether the benefit comes from the use of a genuine, complex, unsolved problem, whether it is due to the presence of 'the stranger in the room', or a combination of both.

The Benefits of PBL and Industry Partners

It has been shown in this study that PBL is responsible for a measurable and strongly significant positive change in the levels of motivation and engagement reported by students. While also responsible for increases in the application of critical and creative thinking skills during learning, the use of an industry partner to provide a genuine, complex, unsolved problem has been shown to increase the application of skills even further.

9.5. Limitations, Constraints and Recommendations.

One variable that may have affected the results of the data was the presence of different teachers in each of the classrooms, presenting the PBL sessions. This could have been avoided if the researcher had provided the lessons for all the students. Although this would have ensured that the teaching was more uniform across the classes, it was neither possible nor desirable for the researcher to teach all the students. It was not possible as the COVID-19 pandemic and time limitations for both the researcher and the four schools made it impossible to schedule. It was not desirable, as a stranger in the classroom can affect students in unforeseen ways which could have an influence on the research. Firstly, each school has particular ways of operating which an outsider does not immediately pick up on. These can be subtle, e.g., in specific terminology or phrases used by staff, or more obvious, such as specific class routines. Secondly, students can feel reluctant to engage openly with their peers and the teacher before they have built a trusting relationship with them. In primary schools this is often the privilege held by the classroom teacher, and in my experience a stranger in the room can inhibit this. It was therefore preferable to train the teachers in the PBL process to reduce the potential variables introduced by the interventions being run by different teachers in different schools. Variations in their management of the intervention were anticipated and were reported by the teachers. Further, for this research to be applicable to teaching in classrooms, having the teacher manage the intervention provides a more genuine comparison than use of a research would.

The schools all reported that the PBL process had been followed as closely as possible. Some variations such as changes in the timeline to accommodate whole school events and excursions were noted. School A (theory Group) reported that "students only swapped research areas once as they enjoyed their chosen area of research and claimed responsibility and connected with it well",

and one teacher said, “I may have prompted them too much but that was a select group of students who struggled.” School D (IP PBL group) reported that they “did not run a second research group before creating our resolution groups. This was due to time and other school commitments.” It is not known how much influence these changes to the recommended schedule may have had on the results of the students’ motivation and engagement, or on their application of critical and creative thinking skills. It is important to recognise however, that teaching occurs in classrooms, with all the changes and variables that are inevitable in real-life schools. Although the research process cannot be controlled fully, the results reflect teaching in the real world and the results that might be expected when PBL is applied in real schools with real students, in different environments.

Three areas of research which would add greatly to our understanding of the students’ experience and the development of the C&CrT skills, but were beyond the scope of this research are:

1) including some qualitative research to provide a richer understanding of the students’ experience;
2) identifying in more detail when and how the students use different levels of the skills; and
3) identifying specifically what students’ find motivating and/or engaging in the PBL and industry partner experiences.

The data collection was carried out in this study at the ‘standard’ time points for a repeated-measures experimental design of pre-intervention, post-intervention and a follow-up (Creswell & Guetterman, 2018). It is possible however that the timing of the data collection for the students’ responses of motivation and engagement may have impacted on the responses they gave. For example, if the MES-JS had been completed by the students at closer proximity to the industry partners input, e.g., immediately following the industry partner’s presentation to the students, and immediately preceding or following the industry partner's critique, the levels of motivation and engagement described by the students may have been vastly different. Further research in which the shorter version of the MES-JS is applied at more regular intervals, may shed more light on the complex relationship between motivation and engagement, and the industry partner’s effect on the students.

It should be noted that the numbers of participants varied greatly when the total data set was split into groups of interest, and at particular time points when significant numbers of students were absent. The ideal group size of 65 participants (calculated at an alpha of .05, effect size of .50, with power at .80) (Creswell & Guetterman, 2018) was not always met when the groups were

compared. This was considered during the analysis as it may have affected the reliability of some of the results, and certainly affects the ability to generalize the results to wider populations.

Two variables which have not been addressed in this research are the possible effects of socio-educational advantage and cultural expectations on the motivation and engagement of students. Reports for example of 'lower academic self-concepts and intrinsic and instrumental values' (Wang et al., 2020, p 435) in East Asian students compared with Western students point to possible cultural effects on motivation and engagement. These differences are particularly observed in the areas of long-term and short-term career goals, and the value placed on family expectations versus personal expectations or preferences (Wang et al., 2020). Although beyond the scope of this research, these factors should be identified as possible influences on the responses, motivation and engagement of students. The ICSEA values (Index of Community Socio-Educational Advantage) of the four schools ranged from the high 900s (more educationally disadvantaged than 64% of schools in Australia), to over 1100 (more educationally disadvantaged than 12% of schools in Australia). Education level of parents and career expectations in families would also have the potential to affect the aspirations and motivation and engagement of students in the classroom.

Language background other than English ranged across the schools from 45% to 1%. The range of language backgrounds is an indicator for a range of cultural backgrounds, which brings with it different expectations of schooling, lessons, and the value of education. It would be extremely interesting to see if there was a difference in the way in which students from different cultural backgrounds responded to the challenge of an industry partner problem. Ideally the treatment groups would have included a cross-section of students from across ICSEA and culturally diverse groups, but the nature of the research (schools and classes) and difficulty obtaining research consent due to COVID-19 made this impossible. Further, data relating to personal details such as immigration status and cultural backgrounds was not available to the researcher, as these issues were beyond the scope of the research questions.

9.6. Conclusions from Study Three

Problem-Based Learning (PBL) has a long history of use in teaching and is currently promoted and used extensively in SA public schools for engaging students in multifaceted, often STEM related, problems. Previous research on the effectiveness of Problem-Based Learning has been inconclusive, due both to the huge variety of methods used in the teaching, and to the variations in the focus of the assessment used to indicate success or otherwise.

Although unable to provide answers on whether PBL and industry partners influence the development of student skills, this research has unequivocally demonstrated the successful use of PBL in increasing the use of critical and creative thinking skills in primary school students during learning. It also supports the previously untested assertions that using an industry partner with their genuine, messy, complicated, unsolved problem helps students use their critical and creative thinking skills more than using a commercial hypothetical problem.

Although the roles of motivation and engagement in the development of the skills require further investigation to be fully understood, there does appear to be a positive, complex relationship between the use of industry partners and student motivation and engagement.

Furthermore, this study has highlighted questions about the effectiveness and accuracy of the Critical and Creative Thinking learning continuum (V8.4) in the Australian Curriculum. The labelling of statements as being age- or school-year appropriate may need to be re-considered. Certainly, the descriptions of the skills at different ages or stages requires detailed study if a truly representative learning progression or continuum of skills is to be produced.

Chapter 10.

Conclusions from Part B: Investigating Problem-Based Learning and Skill Development in the Classroom

10.1. Introduction

The original overarching question asked in this thesis was: 'What is the impact of an industry partner in STEM Problem-Based Learning on the development of critical and creative thinking skills in primary school students?', later amended to 'What is the impact of an industry partner in STEM Problem-Based Learning on the *application* of critical and creative thinking skills in primary school students?' The *application* of the skills during learning was successfully measured, however the *development* of the skills was also positively indicated by the results, and a positive correlation between PBL and critical and creative skill application was shown. Further, the influence of PBL on specific factors of motivation and engagement of students was revealed.

The three empirical studies of which Part B is comprised, approached the overarching question together – first by determining the essential critical and creative thinking skills to be examined, second by assuring the manner in which they were measured was the most suitable, and finally by examining the effect STEM Problem-Based Learning had on the use of critical and creative thinking skills being used in the classroom.

The overarching question sought to generate empirical evidence of the effectiveness of PBL and industry partners in a primary school classroom. Since the research for this thesis began, calls have been made to increase critical and creative thinking skills in students of all ages (World Economic Forum, 2020), and the focus of educational policy and research in Australia and across the world has increasingly turned to assessing students' problem-solving skills to measure their critical and creative abilities (Berry et al., 2022; Organisation for Economic Co-operation and Development (OECD), 2023; Scoular et al., 2020). The measure developed in this PhD project approaches the measurement of student skills from a new angle, focusing on the skills students are using *during* their learning, rather than during an artificial event such as an examination, or through assessing an 'artifact' such as a presentation or model, for evidence of ability.

10.2. Contributions

These three studies provide three contributions to primary education research. First, a novel assessment task has been developed for use with school students from Year 5 (approximate age 10 years), to measure the critical and creative skills they are using during their learning. Although developed specifically to measure the skills described in the Australian Curriculum's Critical and Creative Thinking learning continuum, the measure is also valid for use worldwide, as the skills assessed are of global relevance. Second, the research brought to light the significant discrepancies between the teachers' view and the students' view, of the skills students are using during their learning. This finding has implications for how teachers assess and report on their students' learning and raises new questions about the extent to which both teachers and students understand critical and creative thinking skills. In the final study the Critical and Creative Thinking Self-Report Measure was applied pre-, during-, and post- a STEM Problem-Based Learning intervention. Students reported significant changes in their use of skills, changes which correlated with both the use of Problem-Based Learning, and the use of an industry partner's STEM problem. They also reported improvements in their motivation and engagement. The third major contribution is therefore empirical evidence that STEM Problem-Based Learning, and the use of industry partners, increases students' use of critical and creative thinking skills, and their motivation and engagement in learning.

10.3. Addressing Gaps in the Literature

The literature review included as a component of the conceptual research in Part A identified several gaps in research relevant to the thesis topic.

1) Primary school students. A lack of research relating to primary school students was identified for the areas under investigation in the research in this thesis. This included research on:

a) primary students use of critical and creative thinking skills during learning;
b) suitable measures for assessing the critical and creative thinking skills in the Australian Critical and Creative Thinking learning continuum with primary school students;
c) the effect of PBL on skill use in the primary classroom; the effect of an industry partner in primary school STEM PBL;
d) the effect of an industry partner on motivation and engagement in primary school STEM PBL;
e) the effect of different types of problem in PBL taught in primary school classrooms.

Research outcomes relating to tertiary and secondary students cannot be reliably applied to younger students who are developmentally different from the adults and adolescents on whom the majority of research has been carried out (e.g., Hmelo-Silver, 2004; Mergendoller et al., 2006). This PhD project addressed this gap by carrying out research on primary school students, enabling conclusion to be drawn specifically on primary aged students, their motivation and engagement, and their use of skills.

2) The developmental progression of individual critical and creative thinking skills. A significant amount of research has been carried out to identify which students are more likely to develop creative skills, on identifying creativity when it has occurred, and on identifying the critical thinking skills applied in artificial testing situations. What is missing however, is research identifying the stages that the creative and critical thinking skills proceed through, as they are used by students during learning. This PhD project commenced by analysing the progression of some skills described in the C&CrT continuum and identifying elements which need to be strengthen and skill progressions which require restructuring. The adaptation of the SOLO model (Biggs & Collis, 1982) to some critical and creative thinking skills signals an initial step in the research which is required to understand both how the skills develop, and how they are used in different circumstances. By understanding how the skills develop, we are better placed to find effective ways to enable students to develop their skills.

3) An effective measure of critical and creative skills during learning activities. This complex task was achieved despite the lack of consensus in the research literature relating to the concepts of critical thinking and creative thinking. The lack of a suitable measure was addressed by the production of a novel assessment task, the Critical and Creative Thinking Self-Report Measure, specifically for this purpose. This task provides a unique and robust measure of the use of critical and creative thinking skills during learning, for use at all levels of schooling, including primary.

4) The impact of STEM Problem-Based Learning on the use of critical and creative thinking skills. Despite decades of research which has been reported on PBL, its effect on the use of critical and creative thinking skills has been under-investigated, or under-reported. This may have been due to the lack of reliable measures for these skills. This thesis has demonstrated that increases in skills have been measured in primary school students due to the application of a PBL approach to teaching and learning. If these results can be replicated and the findings are found to apply to a more general population, this will help support the case for STEM PBL to be implemented more widely in primary schools. Further research should be carried out to determine whether the results can be extrapolated to a wider range of students.

5) The effects of different forms of 'problem' in Problem-Based Learning on the use of critical and creative thinking skills. The comparisons of the effects of different types of 'problems' (i.e., hypothetical problems versus genuine industry partner problems), is under-researched generally, and the effect of problem type on the critical and creative thinking skills is previously unreported. This thesis provides some insight into this area, by comparing a hypothetical problem with an industry partner problem and measuring the use of critical and creative thinking skills to demonstrate the positive impact of industry partner involvement. The success demonstrated with the use of an industry partner problem has significant implications for resourcing education and for teaching. These implications include, e.g.: at a school level – the need to consider school-wide timetabling to enable 'problem lessons'; the need to provide teachers with support finding and using an industry partner. At a tertiary level – the need to incorporate training in the use of PBL and IPs in teaching.

6) The impact of PBL on motivation and engagement. Little research has been reported on the relationships between Problem-Based Learning and motivation and engagement, particularly evidence of which components of PBL affect students' motivation and learning. With primary school students, information on this is particularly sparse. The results of Study Three indicated that greater gains were made with reductions in negative motivation and engagement, than the expected increases in positive motivation and engagement. The present study provides a foundation on which to build. Further research, particularly measuring students' motivation and engagement at more regular and perhaps more significant points during the intervention, may present more insight.

7) The effect of the use of industry partner problems on student motivation and engagement. This gap was investigated in Study Three, and provided evidence that industry partner problems have a similar positive effect to PBL, on motivation and engagement. However, the relationship between the use of an industry partner and student motivation and engagement appears to be more complex than expected and may require more detailed data collection to provide meaningful results.

8) Comparing student and teacher assessments of student critical and creative thinking skills. The results of Study Two indicated the level of agreement between the teacher's and the student's assessment of the student's use of skills was very low. This may indicate low levels of accuracy of teachers, or students, or both. It has implications however for the accuracy of teacher reporting on cognitive skills, and on how teachers are trained to recognise and teach these skills. If indeed cognitive skills cannot be accurately assessed by teachers through means other than a self-report measure, this may have further implications for how we assess and report on students.

9) The effect of multimedia theories on assessment tasks. This area has received no research focus previously. Although investigated in Study Two by implementing the measure using two alternative presentation formats, no discernible difference was found between them. This may have been due to the materials being low in element interactivity, or due to the nature of an assessment task which requires students to recall previous activities and report on them, versus a learning activity requiring students to understand and apply new learning. Although inconclusive, this research provides insight on a new application of existing theories.

10.4. Limitations of the Studies and Recommendations

The studies were carried out during the years when the Covid-19 pandemic caused restrictions on access to students in schools. This impacted on the number and type of schools the researcher was able to collect data from. As all the schools in the final data collection were metropolitan schools in Adelaide, South Australia, the extrapolation of the findings to a wider population cannot be assured.

Not all industry partners have the skills to engage students effectively. Dr Belobrajdic was an ideal industry partner to use as he has prior experience working with primary school students engaged in PBL. Any industry partner entering the classroom for the first time must be carefully primed to ensure they present their material effectively for a young audience, without losing the complexity of their genuine problem. Although the value of industry partners has been demonstrated by this research, the effect of the complexity of industry partners problems and 'the stranger in the room' have not yet been fully explored.

On the basis of the conceptual work described in Part A, the skills chosen for inclusion in the Critical and Creative Thinking Self-Report Measure were those specified in the Critical and Creative Thinking learning continuum in the Australian Curriculum. Researchers in other countries may need to adapt the statements in the Critical and Creative Thinking Self-Report Measure to reflect the emphasis of their own national curriculum. The choice of skills being assessed in the Critical and Creative Thinking Self-Report Measure should not therefore be a limitation on the use of the measure in applications outside Australia.

Chapter 11.

Can you Prove it?

On September 19th 2019, as I described how the presence of an industry partner seemed to affect the behaviour of my PBL students, a little voice in my head said “Can you prove it?”

It is now January 2024. As I write, a new research article about the measurement of critical and creative thinking skills used in STEM learning ('Creativity, Critical Thinking, Communication, and Collaboration: Assessment, Certification, and Promotion of 21st Century Skills for the Future of Work and Education': Thornhill-Miller et al., 2023) has just arrived on my desk. The latest webinar from the OECD is titled ‘Unleashing potential: Rethinking assessment for 21st century learning’ (Piacentini, 2023). Research on critical and creative thinking skills, their teaching and measurement, are in focus worldwide.

The overarching question asked in this thesis was: ‘‘What is the impact of an industry partner in STEM Problem-Based Learning on the development of critical and creative thinking skills in primary school students?’ Investigated in this thesis through a combination of conceptual research and empirical research, the answer to this question is threefold.

First, STEM Problem-Based Learning helps students demonstrate the application of a wider range of critical and creative thinking skills during learning. Second, using an industry partner and their problem in STEM PBL helps students demonstrate the application of an even wider range of critical and creative thinking skills during learning than using a hypothetical problem. Third, the increase in the application of skills in these two conditions suggests, but is not evidence of, critical and creative thinking skills being *developed* by STEM PBL and the use of an industry partner.

The development of a measure to assess the use of critical and creative thinking skills enabled the student use of skills to be measured during learning. It also brought to light the need to provide more accurate descriptions of the skills and how they develop than are currently provided in the Critical and Creative Thinking learning continuum in the Australian Curriculum.

The positive effect of Problem-Based Learning on increasing the application of critical and creative thinking skills in student learning highlights the need for a more comprehensive understanding of how these skills develop and are applied by students during their learning. It also

brings into question whether a continuum is the most accurate way to illustrate the use and development of critical and creative thinking skills. Furthermore, if this positive effect of PBL on the application of student skills is supported by further research, the use of Problem-Based Learning in primary schools should be increased, and schools and teachers should be supported to utilise it in their teaching.

The effect of an industry partner on increasing students' skill use during STEM PBL was successfully demonstrated. The relationship between PBL and motivation and engagement was confirmed, but the relationships between industry partners and the motivation and engagement of the students was more complex to unravel. The results of this research highlighted the need for further work to understand the relationship between PBL and motivation and engagement, particularly where industry partners are used. With increased use of industry partners also comes the need to provide additional support for teachers, to enable them to access industry partners and effectively use them in the classroom.

Finally, the investigation found that an industry partner problem is measurably more effective than a hypothetical problem, in helping students apply their critical and creative thinking skills. This highlights the need for further research to enable a more comprehensive understanding of the characteristics of problems – particularly hypothetical and industry partner problems – and how they affect the skills students use.

In addition to the findings relating to the effects of PBL and industry partners on critical and creative thinking skills, were the findings on the effects of PBL on student motivation and engagement. Although the relationship between PBL, industry partners and motivation and engagement has not been fully explored by this research, the results indicate that significant improvements are achieved in motivation and engagement by teaching with Problem-Based Learning.

The research described in this thesis was a result of a desire to be able to prove whether industry partners influence the way students develop and use critical and creative thinking skills. Although initially conceived as a purely empirical study which would rely heavily on the researcher's *critical* thinking skills, the investigation ultimately relied heavily on *creative* thinking to adapt the studies in response to findings in the literature so they could be successfully implemented. Critical and creative thinking are essential to the development of a strong workforce to support the Australian economy and they are essential for students to be able to navigate successfully through the education system. If implementing Problem-Based Learning in schools, and using industry

partners to provide relevant, genuine, complex problems for students to investigate has a positive effect on the use of critical and creative thinking skills by students, what are we waiting for?

The proverb quoted at the beginning of this thesis, 'Give a man a fish and you feed him for a day. Teach him how to fish and you feed him for his lifetime', describes the long-term benefits of teaching skills and understanding, rather than succumbing to what is often the easier action of responding to the immediate problem — providing the answer or dealing with the symptom rather than the cause.

This adage perfectly illustrates the importance of focussing on ways to ensure our students are developing skills which will be useful for their lifetime, rather than teaching only subject knowledge — facts and information which can become quickly outdated and can be accessed relatively easily with current technologies. The research described in this thesis demonstrates that critical and creative thinking skills can be fostered more readily by choosing to engage students through a particular pedagogical approach. The use of STEM Problem-Based Learning, particularly when used in conjunction with an industry partner, can make a positive, measurable difference in the skills being used by students in the classroom. This has the potential to improve the preparation of students for careers in STEM, and for rapid transformations in the future of life and work.

References

Ab Kadir, M. A. (2018). An inquiry into critical thinking in the Australian curriculum: examining its conceptual understandings and their implications on developing critical thinking as a "general capability" on teachers' practice and knowledge. *Asia Pacific Journal of Education*, *38*(4), 533-549.

Acar, S., & Runco, M. A. (2012). Creative Abilities: Divergent thinking. In *Handbook of Organizational Creativity* (pp. 115-139). Academic Press.

Adeniji, S. M., Baker, P., & Schmude, M. (2022). Structure of the Observed Learning Outcomes (SOLO) model: A mixed-method systematic review of research in mathematics education. *EURASIA Journal of Mathematics, Science and Technology Education*, *18*(6).

Albanese, M. A., & Mitchell, S. (1993). Problem-based learning: A review of literature on its outcomes and implementation issues. *Academic medicine*, *68*, 52-81.

Aldridge, J. M., Fraser, B. J., & Huang, T.-C. I. (1999). Investigating classroom environments in Taiwan and Australia with multiple research methods. *The Journal of Educational Research*, *93*(1), 48-62.

Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of personality and social psychology*, *43*(5), 997-1013.

Amabile, T. M. (1983). The social psychology of creativity: A componential conceptualization. *Journal of personality and social psychology*, *45*(2), 357-376.

Amabile, T. M. (1996). *Creativity in context*. Westview Press.

Amabile, T. M. (2012). Componential theory of creativity. *Harvard Business School*, *12*(96), 1-10.

Amabile, T. M., & Gryskiewicz, N. D. (1989). The creative environment scales: Work environment inventory. *Creativity Research Journal*, *2*(4), 231-253.

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of educational psychology*, *84*(3), 261 - 271.

Ananiadou, K., & Claro, M. (2009). *21st Century Skills and Competences for New Millennium Learners in OECDCountries* (No. 41). OECD Publishing. http://dx.doi.org/10.1787/218525261154

Anderson, A. R., Christenson, S. L., Sinclair, M. F., & Lehr, C. A. (2004). Check & Connect: The importance of relationships for promoting engagement with school. *Journal of school psychology*, *42*(2), 95-113.

Anderson, E. (1998). *Motivational and cognitive influences on conceptual knowledge: The combination of science observation and interesting texts* [Unpublished doctoral dissertation]. University of Maryland, College Park.

Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A Taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives*. Allyn & Bacon.

Anderson, N. R., & West, M. A. (1998). Measuring climate for work group innovation: development and validation of the team climate inventory. *Journal of Organizational Behavior*, *19*(3), 235-258.

Andrich, D. (1985). An elaboration of Guttman scaling with Rasch models for measurement. *Sociological methodology*, *15*, 33-80.

Appleton, J. J. (2012). Systems consultation: Developing the assessment-to-intervention link with the Student Engagement Instrument. In S. L. Christensen, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 725-741). Springer.

Appleton, J. J., Christenson, S. L., & Furlong, M. J. (2008). Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in the Schools*, *45*(5), 369-386.

Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the Student Engagement Instrument. *Journal of school psychology*, *44*(5), 427-445.

Arslan, P. Y. (2012). A Review of Multimedia Learning Principles: Split-Attention, Modality, and Redundancy Effects. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, *8*(1), 114-122.

Australian Curriculum Assessment and Reporting Authority. (2019). *Critical and Creative Thinking*. Retrieved 9/5/2019 from https://www.australiancurriculum.edu.au/f-10-curriculum/general-capabilities/critical-and-creative-thinking/

Australian Curriculum Assessment and Reporting Authority (ACARA). (2012). *The Shape of the Australian Curriculum Version 4.0*. ACARA. https://acaraweb.blob.core.windows.net/resources/The_Shape_of_the_Australian_Curriculum_v4.pdf

Australian Curriculum Assessment and Reporting Authority (ACARA). (2013a). *General capabilities*. https://docs.acara.edu.au/resources/General_Capabilities_2011.pdf

Australian Curriculum Assessment and Reporting Authority (ACARA). (2013b). *General Capabilities in the Australian Curriculum*. Australian Curriculum Assessment and Reporting Authority. https://www.australiancurriculum.edu.au/f-10-curriculum/general-capabilities/

Australian Curriculum Assessment and Reporting Authority (ACARA). (2017). *Structure*. https://www.australiancurriculum.edu.au/F-10-curriculum/structure/

Australian Curriculum Assessment and Reporting Authority (ACARA). (2018). *Collaborative problem solving online assessment final report*. ACARA. https://www.acara.edu.au/docs/default-

source/curriculum/201802-acara-collaborative-problem-solving-online-assessment-final-report.pdf?sfvrsn=2

Australian Curriculum Assessment and Reporting Authority (ACARA). (2022). *General Capabilities Creative and Critical Thinking v9*. ACARA. https://v9.australiancurriculum.edu.au/f-10-curriculum/general-capabilities/critical-and-creative-thinking?element=0&sub-element=0

Australian Curriculum Assessment and Reporting Authority (ACARA). (2023a). *Guide to understanding the Index of Community Socio-educational Advantage (ICSEA)*. https://myschool.edu.au/media/1900/guide-to-understanding-icsea-values-2023.pdf

Australian Curriculum Assessment and Reporting Authority (ACARA). (2023b). *What's changed in the new Australian Curriculum?* ACARA. Retrieved 14 June 2023 from https://v9.australiancurriculum.edu.au/resources/stories/curriculum-changes#accordion-41087ee733-item-594b90cc58

Australian Curriculum Assessment and Reporting Authority (ACARA). (n.d.-a). *Australian Curriculum Critical and Creative Thinking*. https://www.australiancurriculum.edu.au/f-10-curriculum/general-capabilities/critical-and-creative-thinking/

Australian Curriculum Assessment and Reporting Authority (ACARA). (n.d.-b). General Capabilities Critical and Creative Thinking learning continuum. In *Australian Curriculum General Capabilities Version 8.4*: ACARA.

Australian Curriculum Assessment and Reporting Authority (ACARA). (n.d.-c). *Science (Version 8.4)*. https://www.australiancurriculum.edu.au/f-10-curriculum/science/

Australian Government Department of Education. (2022). *Australian Curriculum*. https://www.education.gov.au/australian-curriculum

Australian Psychological Society. (2022). *Position Statement. Evidence-based practice and practice-based evidence in psychology*. https://psychology.org.au/getmedia/3478fa00-0a90-43ff-8d90-99a42ea53981/22aps-ps-psysch-p1.pdftarget=

Baas, M., Koch, S., Nijstad, B. A., & De Dreu, C. K. (2015). Conceiving creativity: The nature and consequences of laypeople's beliefs about the realization of creativity. *Psychology of Aesthetics, Creativity, and the Arts*, *9*(3), 340-354.

Baer, J., & Kaufman, J. C. (2005). Bridging generality and specificity: The amusement park theoretical (APT) model of creativity. *Roeper Review*, *27*(3), 158-163.

Baker, M., Rudd, R., & Pomeroy, C. (2001). Relationships between critical and creative thinking. *Journal of Southern Agricultural Education Research*, *51*(1), 173-188.

Bandura, A. (1981). Self-referent thought: A developmental analysis of self-efficacy. *Social cognitive development: Frontiers and possible futures*, *200*(1), 239.

Barbot, B., Besançon, M., & Lubart, T. (2016). The generality-specificity of creativity: Exploring the structure of creative potential with EPoC. *Learning and Individual Differences*, *52*, 178-187.

Barbot, B., Besançon, M., & Lubart, T. I. (2011). Assessing creativity in the classroom. *The Open Education Journal*, *4*, 58-66.

Barbot, B., Hass, R. W., & Reiter-Palmon, R. (2019). Creativity assessment in psychological research:(Re) setting the standards. *Psychology of Aesthetics, Creativity, and the Arts*, *13*(2), 233.

Barron, F. X. (1969). *Creative person and creative process*. Holt, Rinehart & Winston.

Barrows, H. (2000). Forward. In D. H. Evensen & C. E. Hmelo (Eds.), *Problem-based learning: A research perspective on learning interactions*. Routledge.

Barrows, H. S. (1986). A taxonomy of problem-based learning methods. *Medical Education*, *20*(6), 481-486.

Barrows, H. S. (1996). Problem-based learning in medicine and beyond: A brief overview. *New Directions for Teaching and Learning*, *1996*(68), 3-12.

Bates, C., & Nettelbeck, T. (2001). Primary school teachers' judgements of reading achievement. *Educational Psychology*, *21*(2), 177-187.

Batey, M. (2007). *A psychometric investigation of everyday creativity* University of London, University College London (United Kingdom)].

Baumeister, R. F. (2016). Toward a general theory of motivation: Problems, challenges, opportunities, and the big picture. *Motivation and emotion*, *40*, 1-10.

Beghetto, R. A., & Karwowski, M. (2017). Toward untangling creative self-beliefs. In *The Creative Self* (pp. 3-22). Elsevier.

Belland, B. R., Kim, C., & Hannafin, M. J. (2013). A framework for designing scaffolds that improve motivation and cognition. *Educational psychologist*, *48*(4), 243-270.

Berkson, L. (1993). Problem-based learning: have the expectations been met? *Academic medicine*, *68*(October Suppliment), S79 - S88.

Berry, A., Buntting, C., Corrigan, D., Gunstone, R., & Jones, A. (2022). *Education in the 21st century: STEM, creativity and critical thinking*. Springer Nature.

Besemer, S. P., & O'Quin, K. (1999). Confirming the three-factor creative product analysis matrix model in an American sample. *Creativity Research Journal*, *12*(4), 287-296.

Biggs, J. (1999). *Teaching for quality learning at university: what the student does*. SRHE and Open University Press.

Biggs, J., & Collis, K. (1982). *Evaluating the quality of learning: the SOLO taxonomy (structure of the observed learning outcome)*. Academic Press.

Biggs, J., & Collis, K. (1989). Towards a model of school-based curriculum development and assessment using the SOLO taxonomy. *Australian journal of education*, *33*(2), 151-163.

Biggs, J., & Tang, C. (2011). *Teaching for quality learning at university*. McGraw-Hill International.

Bloom, B. S., Engelhardt, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: Cognitive domain*. McKay.

Blumenfeld, P. C., Kempler, T. M., & Krajcik, J. S. (2006). Motivation and cognitive engagement in learning environments. In K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 475-488). Cambridge University Press.

Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational psychologist*, *26*(3-4), 369-398.

Bollen, K. A. (1989). *Structural equations with latent variables*. John Wiley & Sons.

Brown, S., & Race, P. (2013). Using effective assessment to promote learning. In L. Hunt & D. Chalmers (Eds.), *University Teaching in Focus: A learning-centred approach*. London: Routledge.

Burris, S. H. (2005). *Effect of problem-based learning on critical thinking ability and content knowledge of secondary agriculture students* University of Missouri--Columbia].

Burris, S. H., & Garton, B. L. (2007). Effect of Instructional Strategy on Critical Thinking and Content Knowledge: Using Problem-Based Learning in the Secondary Classroom. *Journal of Agricultural Education*, *48*(1), 106-116.

Byrd, R. E. (1986). *Creativity and risk-taking*. Pfeiffer International Publishers.

Care, E., Anderson, K., & Kim, H. (2016). Visualizing the breadth of skills movement across education systems. *Brookings Institution*.

Care, E., Kim, H., Vista, A., & Anderson, K. (2018). Education System Alignment for 21st Century Skills: Focus on Assessment. *Center for Universal Education at The Brookings Institution*.

Carrell, M. R., & Dittrich, J. E. (1978). Equity theory: The recent literature, methodological considerations, and new directions. Academy of Management Review, 3(2), 202-210.

Carson, S. H., Peterson, J. B., & Higgins, D. M. (2005). Reliability, validity, and factor structure of the creative achievement questionnaire. *Creativity Research Journal*, *17*(1), 37-50.

Carter, C. P., Reschly, A. L., Lovelace, M. D., Allpeton, J. J., & Thompson, D. (2012). Measuring Student Engagement Among Elementary Students: Pilot of the Student Engagement Instrument-Elementary Version. *School Psychology Quarterly*, *27*, 61-73.

Cerpa, N., Chandler, P., & Sweller, J. (1996). Some conditions under which integrated computer-based training software can facilitate learning. *Journal of Educational Computing Research*, *15*(4), 345-367.

Chang, C.-C., Tseng, K.-H., & Lou, S.-J. (2012). A comparative analysis of the consistency and difference among teacher-assessment, student self-assessment and peer-assessment in a Web-based portfolio assessment environment for high school students. *Computers & Education*, *58*(1), 303-320. https://doi.org/10.1016/j.compedu.2011.08.005

Chapman, J. W., & Tunmer, W. E. (1995). Development of young children's reading self-concepts: An examination of emerging subcomponents and their relationship with reading achievement. *Journal of educational psychology*, *87*(1), 154-167.

Chapman, J. W., & Tunmer, W. E. (1997). A longitudinal study of beginning reading achievement and reading self-concept. *British Journal of Educational Psychology*, *67*(3), 279-291.

Christenson, S. L., Reschly, A. L., & Wylie, C. (Eds.). (2012a). *Handbook of research on student engagement*. Springer.

Christenson, S. L., Reschly, A. L., & Wylie, C. (2012b). Preface. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. v-viii). Springer.

Clapham, M. M. (2011). Testing/measurement/assessment. In M. A. Runco & S. Pritzker (Eds.), *Encyclopedia of creativity*.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*(3), 149-210.

Cochran, W. G. (1950). The comparison of percentages in matched samples. *Biometrika*, *37*(3/4), 256-266.

Cockburn, C. (2005). *Implementing curriculum in South Australia: the SACSA Framework*. http://www.curriculum.edu.au/leader/sacsa_framework,11031.html?issueID=9789

Colangelo, N., Kerr, B., Hallowell, K., Huesman, R., & Gaeth, J. (1992). The Iowa Inventiveness Inventory: Toward a measure of mechanical inventiveness. *Creativity Research Journal*, *5*(2), 157-163.

Coleman, L. (2022). *Fostering critical and creative thinking in the lower secondary science classroom: A qualitative analysis of exemplary teachers' practice* University of South Australia].

Collis, K., & Biggs, J. (1986). Using the SOLO taxonomy. In *Set: Research Information for Teachers*: NZCER/ACER.

Comparing Class and Grade Level across Countries. (n.d.). http://seandelaney.com/wp-content/uploads/2018/07/Comparing-Class-and-Grade-Levels-across-Countries.pdf

Condliffe, B. (2017). Project-Based Learning: A Literature Review. Working Paper [Working Paper]. *MDRC*, 1-84. https://eric.ed.gov/?id=ED578933

Conradty, C., & Bogner, F. X. (2022). Measuring Students' School Motivation. *Education Sciences*, *12*(6), 378.

Costa, A. L., & Kallick, B. (2004). *Discovering & exploring habits of mind*. Hawker Brownlow Education.

Costa Jr, P. T., & McCrae, R. R. (1992). Four ways five factors are basic. *Personality and Individual Differences*, *13*(6), 653-665.

Costa Jr, P. T., & McCrae, R. R. (2008). The revised neo personality inventory (neo-pi-r). *The SAGE handbook of personality theory and assessment*, *2*(2), 179-198.

Coulaud, H. (2013). Prototype for a computerized version of EPoC. In: University Paris Descartes Paris.

Cozby, P. C., & Bates, S. (2012). *Methods in Behavioral Research* (11 ed.). McGraw-Hill.

Creswell, J. W., & Guetterman, T. C. (2018). *Educational Research: Planning, Conducting and Evaluating Quantitative and Qualitative Research* (6 ed.). Pearson Education Inc.

Cropley, A. J. (2000). Defining and measuring creativity: Are creativity tests worth using? *Roeper Review*, *23*(2), 72-79.

Cropley, A. J. (2006). In praise of convergent thinking. *Creativity Research Journal*, *18*(3), 391-404.

Cropley, D. H. (2015a). Creativity in engineering. In G. E. Corazza & S. Agnoli (Eds.), *Multidisciplinary Contributions to the Science of Creative Thinking* (pp. 155-173). Springer.

Cropley, D. H. (2015b). *Creativity in engineering: Novel solutions to complex problems*. Academic Press.

Cropley, D. H. (2019a). Creativity, innovation and the future of engineering. World Engineers Convention (2019: Melbourne, Australia),

Cropley, D. H. (2019b). *Homo Problematis Solvendis - Problem-solving Man: A History of Human Creativity*. Springer Nature.

Cropley, D. H. (2020). *Femina Problematis Solvendis—Problem solving Woman: A History of the Creativity of Women*. Springer Nature.

Cropley, D. H. (2023). Making an Impact: Fit-for-Purpose Creativity Assessment. In F. K. Reisman (Ed.), *Unpacking Creativity: Culture, Innovation, and Motivation in Global Contexts* (pp. 154-174). KIE Publications.

Cropley, D. H. (in press). Deploying Creativity for Good: How engineers solve worthy problems. In R. J. Sternberg & S. Karami (Eds.), *Transformational Creativity: Learning for a better future*. Palgrave MacMillan.

Cropley, D. H., & Cropley, A. J. (2008). Elements of a universal aesthetic of creativity. *Psychology of Aesthetics, Creativity, and the Arts*, *2*(3), 155-161.

Cropley, D. H., & Cropley, A. J. (2015). *The psychology of innovation in organizations*. Cambridge University Press.

Cropley, D. H., & Kaufman, J. C. (2012). Measuring functional creativity: Non-expert raters and the Creative Solution Diagnosis Scale. *The Journal of Creative Behavior*, *46*(2), 119-137.

Cropley, D. H., Medeiros, K. E., & Damadzic, A. (2023). The intersection of human and artificial creativity. In *Creative provocations: Speculations on the future of creativity, technology & learning* (pp. 19-34). Springer.

Csikszentmihalyi, M. (1990). Flow: The psychology of optimal experience. *1990*.

Csikszentmihalyi, M. (1996). *Creativity: Flow and the psychology of discovery and invention*. Harper Collins.

Cuevas, J., & Dawson, B. L. (2018). A test of two alternative cognitive processing models: Learning styles and dual coding. *Theory and Research in Education*, *16*(1), 40-64.

Cumming, G., & Finch, S. (2005). Inference by eye: confidence intervals and how to read pictures of data. *American Psychologist*, *60*(2), 170-180.

Curtis, D., & Denton, R. (2003). *The authentic performance based assessment of problem solving*. NCVER.

Davies, M. (2019). New Directions in the Teaching of Critical Thinking. *Change: The Magazine of Higher Learning*, *51*(5), 18-27. https://doi.org/10.1080/00091383.2019.1652063

Davies, M., & Barnett, R. (2015). *The Palgrave Handbook of Critical Thinking in Higher Education*. Springer.

Davis, G. A., & Rimm, S. (1982). Group Inventory for Finding Interests (GIFFI) I and II: Instruments for identifying creative potential in the junior and senior high school. *The Journal of Creative Behavior*, *16*, 50-57.

Davis, T. R., & Luthans, F. (1980). A social learning approach to organizational behavior. Academy of Management Review, 5(2), 281-290.

de Bono, E. (2009). CoRT 1 Breadth Tools. Retrieved 10 October 2011, from www.deBonoForSchools.com

Deakin Crick, R. (2012). Deep engagement as a complex system: Identity, learning power and authentic enquiry. In S. L. Christensen, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 675-694). Springer.

Deal, J. E., Halverson Jr, C. F., Martin, R. P., Victor, J., & Baker, S. (2007). The Inventory of Children's Individual Differences: Development and validation of a short version. *Journal of Personality Assessment*, *89*(2), 162-166.

DeLuca, C., Chapman-Chin, A., & Klinger, D. A. (2019). Toward a teacher professional learning continuum in assessment for learning. *Educational Assessment*, *24*(4), 267-285.

Department for Education (DfE) South Australia. (2017). *STEM Learning Companion Document Strategy Overview 2017 to 2020*.

Department for Education (DfE) South Australia. (2022). *STEM school-industry partnerships*. https://www.education.sa.gov.au/parents-and-families/curriculum-and-learning/stem-learning/stem-school-industry-partnerships

Department for Education (DfE) South Australia. (n.d.-a). Inspire numeracy guidebook - actions for secondary teachers. In Department for Education (Ed.).

Department for Education (DfE) South Australia. (n.d.-b). STEM Learning. https://edi.sa.edu.au/library/document-library/educational-resources/stem-learning/problem-and-inquiry-based-learning.pdf

Department for Education (DfE) South Australia. (n.d.-c). STEM School Industries Pathways Brochure. In *https://www.education.sa.gov.au/parents-and-families/curriculum-and-learning/stem-learning/stem-school-industry-partnerships*.

Department for Education and Child Development (DECD). (2013). *Guidelines for the Implementation of the Australian Curriculum in DECD schools: Reception-Year 10*. Department for Education and Child Development,

Department for Education and Child Development (DECD). (2016). *STEM Learning Strategy for DECD Preschool to Year 12*. Department for Education and Child Development.

Department of Education. (2022). *Links to 21st century learning*. Australian Government. https://www.education.gov.au/australian-curriculum/national-stem-education-resources-toolkit/i-want-know-about-stem-education/what-works-best-when-teaching-stem/links-21st-century-learning

Department of Education Employment and Workplace Relations. (2010). *Unfolding opportunities: a baseline study of school-business relationships in Australia. Final Report*. P. KPA. https://docs.education.gov.au/system/files/doc/other/partnerships_for_schools_businesses_and_communities_appendix_1_and_2_final_study_report.pdf

DeVellis, R. F. (2012). *Scale development: Theory and applications.* Sage Publications.

Dewey, J. (1910). *How We Think*. D.C. Heath & Co.

Diakidoy, I.-A. N., Stylianou, P., Karefillidou, C., & Papageorgiou, P. (2005). The relationship between listening and reading comprehension of different types of text at increasing grade levels. *Reading psychology*, *26*(1), 55-80.

Diedrich, J., Jauk, E., Silvia, P. J., Gredlein, J. M., Neubauer, A. C., & Benedek, M. (2018). Assessment of real-life creativity: The Inventory of Creative Activities and Achievements (ICAA). *Psychology of Aesthetics, Creativity, and the Arts*, *12*(3), 304.

Dingli, S. (2008). Thinking outside the box: Edward de Bono's lateral thinking. In *The Routledge companion to creativity* (pp. 338-350). Routledge.

Dochy, F., Segers, M., Van den Bossche, P., & Gijbels, D. (2003). Effects of problem-based learning: A meta-analysis. *Learning and instruction*, *13*(5), 533-568.

Dollinger, S. J., Urban, K. K., & James, T. A. (2004). Creativity and openness: Further validation of two creative product measures. *Creativity Research Journal*, *16*(1), 35-47.

Domino, G. (1994). Assessment of creativity with the ACL: An empirical comparison of four scales. *Creativity Research Journal*, *7*(1), 21-33.

Donnelly, K., & Wiltshire, K. (2014). *Review of the Australian Curriculum: final report*. Retrieved from https://apo.org.au/sites/default/files/resource-files/2014-10/apo-nid41699.pdf

Doolittle, J. H. (1990). *Creative Reasoning Test*. Midwest Publications/ Critical Thinking Press.

Dornbusch, S. M., & Steinberg, L. (1990). *Measures of school engagement*. Temple University Philadelphia.

Downing, S. M. (2006). Twelve steps for effective test development. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 3-25). Lawrence Erlbaum Associates. http://www.danangtimes.vn/Portals/0/Docs/314154729-0805852654.pdf#page=18

Drake, K. N., & Long, D. (2009). Rebecca's in the dark: A comparative study of problem-based learning and direct instruction/experiential learning in two 4th-grade classrooms. *Journal of Elementary Science Education*, *21*(1), 1-16.

Duke, N., Halvorsen, A.-L., Strachan, S. L., Kim, J., & Konstantopoulos, S. (2017). Putting PBL to the Test: The Impact of Project-based Learning on Second-grade Students' Social Studies and Literacy Learning and Motivation. https://www.pblworks.org/sites/default/files/2019-02/Duke%20Halvorsen%20Strachan%20Kim%20Konstantopoulos%20November%202018.pdf

Dunn, O. J. (1964). Multiple comparisons using rank sums. *Technometrics*, *6*(3), 241-252.

Dunning, D. (2011). The Dunning–Kruger effect: On being ignorant of one's own ignorance. *Advances in experimental social psychology*, *44*, 247-296.

Education Council. (2015). *National STEM school education strategy: A comprehensive plan for science, technology, engineering and mathematics education in Australia*. Education Council. http://www.educationcouncil.edu.au/site/DefaultSite/filesystem/documents/National%20STEM%20School%20Education%20Strategy.pdf

Education Services Australia (ESA). (2019). *Learning Progressions and Online Formative Assessment National Initiative - Discovery Phase. Final Report - Attachment 3, Critical and Creative Thinking: A Report on Existing and Future Work*. https://www.ofai.edu.au/media/diplwvzt/a-03_attachment-3_critical-and-creative-thinking.pdf

Ellerton, P. (2017). *On critical thinking and collaborative inquiry* (Education Future Frontiers, Issue. https://pdfs.semanticscholar.org/917a/be01e83d41c6df5099a1c987c66f0f38bc3b.pdf

Ellerton, P. (n.d.). The Critical Thinking Matrix. In: University of Queensland Critical Thinking Project.

Ennis, R., & Weir, E. (1985). *The Ennis-Weir Critical Thinking Essay Test*. Midwest Publications.

Ennis, R. H. (1993). Critical Thinking Assessment. *Theory Into Practice*, *32*(3), 179-186.

Ennis, R. H., Gardiner, W. L., Guzzetta, J., Morrow, R., Paulus, D., & Ringel, L. (1964). *The Cornell Conditional-reasoning Test, Form x*. Cornell Critical Thinking Project.

Ennis, R. H., Gardiner, W. L., Morrow, R., Paulus, D., & Ringel, L. (1964). *The Cornell Class-reasoning Test, Form X*. University of Illinois.

Ennis, R. H., Millman, J., & Tomko, T. N. (1985). *Cornell critical thinking tests level X & level Z: Manual*. Midwest Publications Pacific Grove, CA.

Facione, N. C., & Facione, P. (1994). *The California critical thinking skills test: Form A and B test manual*. The California Academic Press.

Facione, N. C., Facione, P. A., & Giancarlo, C. A. (1994). The Disposition Toward Critical Thinking as a Measure of Competent Clinical Judgment: The Development of the California Critical Thinking Disposition Inventory. *Nursing Education*, 345-350.

Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. research findings and recommendations.* https://search-proquest-com.access.library.unisa.edu.au/docview/63064314?accountid=14649

Facione, P. A. (1991). Using the California Critical Thinking Skills Test in Research, Evaluation, and Assessment.

Facione, P. A. (2020). Critical thinking: What it is and why it counts. *Insight assessment*, 1-31. https://www.insightassessment.com/wp-content/uploads/ia/pdf/whatwhy.pdf

Fajari, L. E. W., Sarwanto, & Chumdari. (2020). Influence of PBL-based multimedia and learning motivation on students critical-thinking skills in elementary schools. *Journal of Physics: Conference Series*, *1511*(1).

Feist, G. J. (2010). The function of personality in creativity: The nature and nurture of the creative personality. In J. C. Kaufman & R. J. Sternberg (Eds.), *The Cambridge handbook of creativity* (pp. 113-130). Cambridge University Press.

Field, A. (2018). *Discovering Statistics using IBM SPSS Statistics* (5th ed.). Sage.

Finn, J. D., & Zimmer, K. S. (2012). Student engagement: What is it? Why does it matter? In S. L. Christensen, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 97-131). Springer.

Fortier, M. S., Vallerand, R. J., & Guay, F. (1995). Academic motivation and school performance: Toward a structural model. *Contemporary Educational Psychology*, *20*(3), 257-274.

Forwood, C. (2018). *STEM Investigations for Upper Primary*. Macmillan Education.

Foster, N., & Schleicher, A. (2022). Assessing Creative Skills. *Creative Education*, *13*, 1-29. https://doi.org/10.4236/ce.2022.131001

Fraser, B. J. (1979). A Test of Enquiry Skills. In (pp. 77-94).

Fredricks, J. A., Blumenfeld, P. C., Friedel, J. M., & Paris, A. H. (2005). School Engagement. In K. A. Moore & L. H. Lippman (Eds.), *What Do Children Need to Flourish?* Springer.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59-109.

Fredricks, J. A., & McColskey, W. (2012). The measurement of student engagement: A comparative analysis of various methods and student self-report instruments. In S. L. Christensen, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 763-782). Springer.

Fry, E. (1977). Fry's readability graph: Clarifications, validity, and extension to level 17. *Journal of reading*, *21*(3), 242-252.

Fulmer, S. M., & Frijters, J. C. (2009). A review of self-report and alternative approaches in the measurement of student motivation. *Educational Psychology Review*, *21*, 219-246.

Furnham, A., Batey, M., Anand, K., & Manfield, J. (2008). Personality, hypomania, intelligence and creativity. *Personality and Individual Differences*, *44*(5), 1060-1069.

Furrer, C., & Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of educational psychology*, *95*(1), 148.

Gaardboe, S. (2019). *STEM Learning Evaluation Grant Report Brighton Primary School*. Department for Education,.

Gaardboe, S. (2020). Researching school-industry STEM partnerships. *Teacher*. https://www.teachermagazine.com/au_en/articles/researching-school-industry-stem-partnerships

Gagné, F., & St Père, F. (2001). When IQ is controlled, does motivation still predict achievement? *Intelligence*, *30*(1), 71-100.

Gallagher, S. A., & Gallagher, J. J. (2013). Using problem-based learning to explore unseen academic potential. *Interdisciplinary Journal of Problem-Based Learning*, *7*(1), 111-131.

Gambrell, L. B., Palmer, B. M., Codling, R. M., & Mazzoni, S. A. (1996). *Motivation to Read Profile (MRP)*. APA PsycTests. https://doi.org/10.1037/t43709-000

Gamse, B. C., Martinez, A., & Bozzi, L. (2017). Calling STEM experts: how can experts contribute to students' increased STEM engagement? *International Journal of Science Education, Part B*, *7*(1), 31-59.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. Basic Books.

Gardner, H. (2009). *5 Minds for the Future*. McGraw-Hill.

Gernes, A. (2021). *Impact of Brain Breaks on Student Engagement, Motivation, and Achievement in a Third-Grade Classroom* (Publication Number 6) [Education Masters, Winona State University]. https://openriver.winona.edu/educationmasterspapers/6

Giancarlo, C. A., Blohm, S. W., & Urdan, T. (2004). Assessing secondary students' disposition toward critical thinking: Development of the California Measure of Mental Motivation. *Educational and Psychological Measurement*, *64*(2), 347-364.

Gibbons, J. D., & Chakraborti, S. (2011). *Nonparametric statistical inference* (5th ed.). Chapman & Hall/CRC Press.

Gilbert, R. (2018). General capabilities in the Australian Curriculum: Promise, problems and prospects. In A. Reid & D. Price (Eds.), *The Australian Curriculum: Promises, problems and possibilities*. Australian Curriculum Studies Association.

Gilbert, R. (2019). General capabilities in the Australian curriculum: Promise, problems and prospects. *Curriculum Perspectives*, *39*(2), 169-177.

Gilson, L. L., & Goldberg, C. B. (2015). Editors' comment: so, what is a conceptual paper? *Group and Organization Management*, *40*(2), 127-130.

Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and instruction*, *15*(4), 313-331.

Gonski, D., Arcus, T., Boston, K., Gould, V., Johnson, W., O'Brien, L., Perry, L-A., Roberts, M. . (2018). *Through Growth to Achievement: Report of the Review to Achieve Educational Excellence in Australian Schools*. https://www.education.gov.au/quality-schools-package/resources/through-growth-achievement-report-review-achieve-educational-excellence-australian-schools

Gottfried, A. E. (1986). *Children's Academic Intrinsic Motivation Inventory:(CAIMI)*. Psychological Assessment Resources.

Gottfried, A. E. (2019). Academic intrinsic motivation: Theory, assessment, and longitudinal research. In *Advances in motivation science* (Vol. 6, pp. 71-109). Elsevier.

Gottfried, A. E., & Gottfried, A. W. (1996). A longitudinal study of academic intrinsic motivation in intellectually gifted children: Childhood through early adolescence. *Gifted Child Quarterly*, *40*(4), 179-183.

Gough, H. G. (1979). A creative personality scale for the Adjective Check List. *Journal of personality and social psychology*, *37*(8), 1398-1405.

Gough, H. G. (1987). *The California Psychological Inventory, administrators guide*. Consulting Psychologists Press.

Gough, H. G. (1992). Assessment of creative potential in psychology and the development of a creative temperament scale for the CPI. In J. C. Rosen & P. McReynolds (Eds.), *Advances in psychological assessment* (Vol. 8, pp. 225-257). Springer Science+Business Media.

Gough, H. G., & Heilbrun, A. B. (1983). *The Adjective Check List Manual* (2nd ed.). Consulting Psychologists Press.

Grade Equivalents. (n.d.). https://www.asmilan.org/admissions/grade-equivalents

Gralewski, J., & Karwowski, M. (2019). Are teachers' ratings of students' creativity related to students' divergent thinking? A meta-analysis. *Thinking Skills and Creativity*, *33*. https://doi.org/10.1016/j.tsc.2019.100583

Green, J., Martin, A. J., & Marsh, H. W. (2007). Motivation and engagement in English, mathematics and science high school subjects: Towards an understanding of multidimensional domain specificity. *Learning and Individual Differences*, *17*(3), 269-279.

Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, *P* values, confidence intervals, and power: a guide to misinterpretations. *European journal of epidemiology*, *31*(4), 337-350.

Guilford, J. P. (1959). Three faces of intellect. *American Psychologist*, *14*(8), 469-479.

Guilford, J. P. (1967). *The nature of human intelligence*. McGraw-Hill.

Guilford, J. P. (1976). *Creativity tests for children*. Sheridan Psychological Services.

Guilford, J. P., & Hoepfner, R. (1966). *Structure-of-intellect Factors and Their Tests, 1966: Studies of Aptitudes of High-level Personnel*. University of Southern California.

Guthrie, J. T., & Wigfield, A. (2000). Engagement and Motivation in Reading. In *Handbook of Reading Research* (Vol. 3).

Haner, U.-E. (2005). Spaces for creativity and innovation in two established organizations. *Creativity and Innovation Management*, *14*, 288-298.

Hasinger, E. (2023). *Cornell Critical Thinking Test Guide*. tests.com. https://www.tests.com/Cornell-Critical-Thinking-Testing

Hatcher, D. (2015). Critical Thinking Instruction: A Realistic Evaluation The Dream vs. Reality. *Inquiry: Critical Thinking Across the Disciplines*, *30*.

Heilmann, G., & Korte, W. B. (2010). *The Role of Creativity and Innovation in School Curricula in the EU27: A content analysis of curricula documents*. European Commission Joint Research Centre. ftp://ftp.jrc.es/pub/EURdoc/JRC61106_TN.pdf

Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (2007). *Experience sampling method: Measuring the quality of everyday life*. Sage.

Henk, W. A., & Melnick, S. A. (1995). The Reader Self-Perception Scale (RSPS): A new tool for measuring how children feel about themselves as readers. *The reading teacher*, *48*(6), 470-482.

Hennessey, B. A. (1994). The consensual assessment technique: An examination of the relationship between ratings of product and process creativity. *Creativity Research Journal*, *7*(2), 193-208.

Hennessey, B. A. (2010). The creativity-motivation connection. In J. C. Kaufman & R. J. Sternberg (Eds.), *The Cambridge handbook of creativity* (pp. 342-365). Cambridge University Press.

Henriksen, D., Mishra, P., & Mehta, R. (2015). Novel, effective, whole: Toward a NEW framework for evaluations of creative products. *Journal of Technology and Teacher Education*, *23*(3), 455-478.

Heppner, P. P., & Petersen, C. H. (1982). The development and implications of a personal problem-solving inventory. *Journal of Counseling Psychology*, *29*(1), 66.

Hertzka, A. F., & Guilford, J. P. (1955). *Logical reasoning*. Sheridan Supply Co.

Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review*, *16*(3), 235-266.

Hocevar, D. (1979). The Development of the Creative Behavior Inventory (CBI).

Hocevar, D. (1981). Measurement of creativity: Review and critique. *Journal of Personality Assessment*, *45*(5), 450-464.

Hocevar, D., & Bachelor, P. (1989). A taxonomy and critique of measurements used in the study of creativity. *Handbook of creativity*, 53-75.

Hofkens, T. L., & Pianta, R. C. (2022). Teacher–Student Relationships, Engagement in School, and Student Outcomes. In A. L. Reschly & S. L. Christenson (Eds.), *Handbook of research on student engagement* (pp. 431-449). Springer.

Horowitz, R., & Samuels, S. J. (1985). Reading and listening to expository text. *Journal of Reading Behavior*, *17*(3), 185-198.

Hung, P.-H., Hwang, G.-J., Lee, Y.-H., Wu, T.-H., Vogel, B., Milrad, M., & Johansson, E. (2014). A problem-based ubiquitous learning approach to improving the questioning abilities of elementary school students. *Journal of Educational Technology & Society*, *17*(4), 316-334.

Hung, W. (2011). Theory to reality: A few issues in implementing problem-based learning. *Educational Technology Research and Development*, *59*(4), 529-552.

Hung, W. (2016). All PBL starts here: The problem. *Interdisciplinary Journal of Problem-Based Learning*, *10*(2), 10. https://doi.org/org/10.7771/1541-5015.1604

Hung, W., Jonassen, D. H., & Liu, R. (2008). Problem-Based Learning. In *Handbook of research on educational communications and technology* (Vol. 3, pp. 485-506).

Hung, W., Mehl, K., & Holen, J. B. (2013). The relationships between problem design and learning process in problem-based learning environments: Two cases. *The Asia-Pacific Education Researcher*, *22*, 635-645.

Hung, Y.-j., Samuelson, B. L., & Chen, S.-c. (2016). Relationships between peer-and self-assessment and teacher assessment of young EFL learners' oral presentations. In M. Nikolov (Ed.), *Assessing young learners of English: Global and local perspectives* (pp. 317-338). Springer International.

Insight Assessment. (2023a). *California Critical Thinking Disposition Inventory (CCTDI)*. California Academic Press. https://www.insightassessment.com/product/cctdi

Insight Assessment. (2023b). *California Critical Thinking Skills Test (CCTST)*. California Academic Press. https://www.insightassessment.com/product/cctst

Insight Assessment. (2023c). *Educate Insight Grades 3-5*. California Academic Press. https://www.insightassessment.com/product/educate-insight-grades-3-5

Insight Assessment. (2023d). *Educate Insight Grades 6-8*. California Academic Press. https://www.insightassessment.com/product/educate-insight-grades-6-8

Isaksen, S. G., Lauer, K. J., & Ekvall, G. (1999). Situational outlook questionnaire: A measure of the climate for creativity and change. *Psychological Reports*, *85*(2), 665-674.

Ivcevic, Z. (2007). Artistic and everyday creativity: An act-frequency approach. *The Journal of Creative Behavior*, *41*(4), 271-290.

Ivcevic, Z. (2009). Creativity map: Toward the next generation of theories of creativity. *Psychology of Aesthetics, Creativity, and the Arts*, *3*(1), 17 -21.

Ivcevic, Z., & Mayer, J. D. (2009). Mapping dimensions of creativity in the life-space. *Creativity Research Journal*, *21*(2-3), 152-165.

Jaakkola, E. (2020). Designing conceptual articles: four approaches. *AMS review*, *10*(1-2), 18-26.

Jensen, K. J. (2015). *A Meta-analysis of the effects of problem-and project-based learning on academic achievement in grades 6-12 populations* Seattle Pacific University].

Jerzembek, G., & Murphy, S. (2013). A narrative review of problem-based learning with school-aged children: implementation and outcomes. *Educational Review*, *65*(2), 206-218.

Jiang, D., & Sweller, J. (2022). The Transient Information Principle in Multimedia Learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed.). Cambridge University Press. https://doi.org/10.1017/9781108894333

Jimoyiannis, A. (2011). Using SOLO taxonomy to explore students' mental models of the programming variable and the assignment statement. *Themes in Science and Technology Education*, *4*(2), 53-74.

Johnson, D. L. (1979). *The Creativity Checklist*. Stoelting.

Jolly, A. (2016). *STEM by Design: Strategies and Activities for Grades 4-8*. Routledge.

Jonassen, D. H. (1997). Instructional design models for well-structured and III-structured problem-solving learning outcomes. Educational Technology Research and Development, 45(1), 65-94.

Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educational Technology Research and Development*, *48*(4), 63-85.

Jonassen, D. H., & Hung, W. (2008). All problems are not equal: Implications for Problem-Based Learning. *Interdisciplinary Journal of Problem-Based Learning*, *2*(2), 6 - 28.

Kalmar, D. A., & Sternberg, R. J. (1988). Theory knitting: An integrative approach to theory development. *Philosophical psychology*, *1*(2), 153-170.

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *13*(4), 351-371.

Kalyuga, S., & Sweller, J. (2014). The Redundancy Principle in Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 247 - 262). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369

Kalyuga, S., & Sweller, J. (2021). The Redundancy Principle in Multimedia Learning. In L. Fiorella & R. E. Mayer (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 212-220). Cambridge University Press. https://doi.org/DOI: 10.1017/9781108894333.021

Kaplan, L. D. (1991). Teaching intellectual autonomy: The failure of the critical thinking movement. *Educational Theory*, *41*, 361-370.

Kapoor, H., Reiter-Palmon, R., & Kaufman, J. C. (2021). Norming the muses: Establishing the psychometric properties of the Kaufman Domains of Creativity Scale. *Journal of Psychoeducational Assessment*, *39*(6), 680-693.

Karwowski, M., Han, M.-H., & Beghetto, R. A. (2019). Toward dynamizing the measurement of creative confidence beliefs. *Psychology of Aesthetics, Creativity, and the Arts*, *13*(2), 193-202.

Karwowski, M., Lebuda, I., & Wiśniewska, E. (2018). Measuring creative self-efficacy and creative personal identity. *The International Journal of Creativity & Problem Solving*.

Kaufman, J.C. (2009). Creativity 101 (2nd ed.). Springer Publishing Company.

Kaufman, J. C. (2012). Counting the muses: Development of the Kaufman Domains of Creativity Scale (K-DOCS). *Psychology of Aesthetics, Creativity, and the Arts*, *6*(4), 298-308.

Kaufman, J. C. (2019). Self-assessments of creativity: Not ideal, but better than you think. *Psychology of Aesthetics, Creativity, and the Arts*, *13*(2), 187.

Kaufman, J. C., & Baer, J. (Eds.). (2005). *Creativity across domains: Faces of the muse*. Lawrence Erlbaum Associates Publishers.

Kaufman, J. C., & Beghetto, R. A. (2009). Beyond big and little: The four c model of creativity. *Review of General Psychology*, *13*(1), 1-12.

Kaufman, J. C., Cole, J. C., & Baer, J. (2009). The construct of creativity: Structural model for self-reported creativity ratings. *The Journal of Creative Behavior*, *43*(2), 119-134.

Kaufman, J. C., Glăveanu, V., & Baer, J. (Eds.). (2017). *The Cambridge Handbook of Creativity Across Domains*. Cambridge University Press.

Kaufman, J. C., Lee, J., Baer, J., & Lee, S. (2007). Captions, consistency, creativity, and the consensual assessment technique: New evidence of reliability. *Thinking Skills and Creativity*, *2*(2), 96-106.

Kaufman, J. C., & Reiter-Palmon, R. (2017). *The creative response evaluation at work test manual*.

Kaufman, J. C., Waterstreet, M. A., Ailabouni, H. S., Whitcomb, H. J., Roe, A. K., & Riggs, M. (2010). Personality and self-perceptions of creativity across domains. *Imagination, Cognition and Personality*, *29*(3), 193-209.

Kelley, T. R., & Knowles, J. G. (2016). A conceptual framework for integrated STEM education. *International Journal of STEM Education*, *3*, 1-11.

Kim, C., Park, S. W., Cozart, J., & Lee, H. (2015). From motivation to engagement: The role of effort regulation of virtual high school students in mathematics courses. *Journal of Educational Technology & Society*, *18*(4), 261-272.

Kimberlin Education. (2018). *IBL, PBL and PJBL, WHAT'S THE DIFFERENCE?* https://kimberlineducation.com/ibl-pbl-and-pjbl-whats-the-difference/

Kirschenbaum, R. J. (1989). *Understanding the creative activity of students: Including an instruction manual for the creative behavior inventory*. Creative Learning Press.

Kirschner, P. A., Park, B., Malone, S., & Jarodzka, H. (2017). *Toward a cognitive theory of multimedia assessment (CTMMA)* (M. J. Spector, B. B. Lockee, & M. D. Childress, Eds.). Springer Cham.

Kirton, M. (1989). *Adaptors and innovators: Styles of creativity and problem solving*. Routledge.

Klem, A. M., & Connell, J. P. (2004). Relationships matter: Linking teacher support to student engagement and achievement. *Journal of school health*, *74*, 262-273.

Koski, A., Puustinen, M., Säntti, J., & Salminen, J. (2023). Exploring the research-based approach of academic classroom teachers in Finland. *Teachers and Teaching*, 1-17.

Kosovich, J. J., Hulleman, C. S., & Barron, K. E. (2017). Measuring motivation in educational settings: A Case for pragmatic measurement. In K. A. Renninger & S. E. Hidi (Eds.), *The Cambridge handbook on motivation and learning* (pp. 39-60). Routledge.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory Into Practice*, *41*(4), 212-218.

Kuhn, D. (1999). A Developmental Model of Critical Thinking. *Educational Researcher*, *28*, 16-25+46.

Kumar, V., Kemmler, D., & Holman, E. R. (1997). The Creativity Styles Questionnaire-Revised. *Creativity Research Journal*, *10*(1), 51-58.

Laerd Statistics. (2015a). Chronbach's alpha using SPSS Statistics. *Statistical tutorials and software guides*. https://statistics.laerd.com/

Laerd Statistics. (2015b). Friedman test using SPSS Statistics. *Statistical tutorials and software guides*. https://statistics.laerd.com/

Laerd Statistics. (2017a). Cochrane's Q test using SPSS Statistics. *Statistical tutorials and software guides*. https://statistics.laerd.com/

Laerd Statistics. (2017b). *Statistical tutorials and software guides*. Retrieved 27 March 2023 from https://statistics.laerd.com/

Lai, E. R. (2011). Critical thinking: A literature review. *Pearson's Research Reports*, *6*, 1-49.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159-174.

Lapuz, A. M., & Fulgencio, M. N. (2020). Improving the critical thinking skills of secondary school students using problem-based learning. *Lapuz, AME, & Fulgencio, MN (2020). Improving the Critical Thinking Skills of Secondary School Students using Problem-Based Learning. International Journal of Academic Multidisciplinary Research,(4)*, *1*, 1-7.

Larmer, J. (2014). Project-Based Learning vs. Problem-Based Learning vs. X-BL. *Project-Based Learning*. https://www.edutopia.org/blog/pbl-vs-pbl-vs-xbl-john-larmer

Lassig, C. J. (2013). Approaches to creativity: How adolescents engage in the creative process. *Thinking Skills and Creativity*, *10*, 3-12.

Lau, S., & Cheung, P. C. (2010). Creativity assessment: Comparability of the electronic and paper-and-pencil versions of the Wallach–Kogan Creativity Tests. *Thinking Skills and Creativity*, *5*(3), 101-107.

Leedy, P. D., & Ormrod, J. E. (2015). *Practical Research Planning and Design* (11th ed.). Pearson.

Lei, H., Cui, Y., & Zhou, W. (2018). Relationships between student engagement and academic achievement: A meta-analysis. *Social Behavior and Personality: an international journal*, *46*(3), 517-528.

Lepper, M. R., Sethi, S., Dialdin, D., & Drake, M. (1997). Intrinsic and extrinsic motivation: A developmental perspective. *Developmental psychopathology: Perspectives on adjustment, risk, and disorder*, 23-50.

Lester, D., Hvezda, J., Sullivan, S., & Plourde, R. (1983). Maslow's hierarchy of needs and psychological health. *The Journal of General Psychology*, *109*(1), 83-85.

Leucht, M., Tiffin-Richards, S., Vock, M., Pant, H. A., & Koeller, O. (2012). English teachers' diagnostic skills in judging their students' competencies on the basis of the Common European Framework of Reference. *Zeitschrift für Entwicklungspsychologie und pädagogische Psychologie*, *44*(4), 163-177.

Liem, G. A. D., & Martin, A. J. (2012). The Motivation and Engagement Scale: Theoretical framework, psychometric properties, and applied yields. *Australian Psychologist*, *47*(1), 3-13.

Lindner, M. (2022). Principles for Educational Assessment with Multimedia. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning*. Cambridge University Press. https://doi.org/10.1017/9781108894333

Lioncrest. (2023). *Macmillan STEM Investigations*. https://lioncrest.com.au/our-range/stem-science/macmillan-stem-investigations-detail

Locke, E. A., & Latham, G. P. (2004). What should we do about motivation theory? Six recommendations for the twenty-first century. *Academy of Management Review*, *29*(3), 388-403.

Lodico, M. G., Spaulding, D. T., & Voegtle, K. H. (2010). *Methods in educational research: From theory to practice*. John Wiley & Sons.

Long, H., Kerr, B. A., Emler, T. E., & Birdnow, M. (2022). A Critical Review of Assessments of Creativity in Education. *Review of Research in Education*, *46*(1), 288-323. https://doi.org/10.3102/0091732X221084326

Lonsdale, M. (2011). *School-Community Partnerships in Australian Schools*. Australian Council for Educational Research (ACER). https://research.acer.edu.au/policy_analysis_misc/7/

Lonsdale, M., & Anderson, M. (2012). Preparing 21st century learners: the case for school-community collaborations. https://research.acer.edu.au/policy_analysis_misc/16/

Low, R., & Sweller, J. (2005). The Modality Principle in Multimedia Learning. In R. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 147-158). Cambridge University Press. https://doi.org/DOI: 10.1017/CBO9780511816819.010

Lubart, T., Barbot, B., & Besançon, M. (2019). Creative potential: assessment issues and the EPoC Battery/Potencial creativo: temas de evaluación y batería EPoC. *Studies in Psychology*, *40*(3), 540-562.

Lubart, T., Besançon, M., & Barbot, B. (2011). *Evaluation du Potentiel Creatif (EPoC)*. Hogrefe.

Lubart, T., Zenasni, F., & Barbot, B. (2013). Creative potential and its measurement. *International Journal for Talent Development and Creativity*, *1*(2), 41-51.

Lucas, B., & Smith, C. (2018). The capable country: Cultivating capabilities in Australian education, Mitchell Institute policy report No. 03/2018. https://apo.org.au/sites/default/files/resource-files/2018-10/apo-nid198271.pdf

Machts, N., Kaiser, J., Schmidt, F. T., & Moeller, J. (2016). Accuracy of teachers' judgments of students' cognitive abilities: A meta-analysis. *Educational Research Review*, *19*, 85-103.

Maestrales, S., Marias Dezendorf, R., Tang, X., Salmela-Aro, K., Bartz, K., Juuti, K., . . . Schneider, B. (2022). US and Finnish high school science engagement during the COVID-19 pandemic. *International Journal of Psychology*, *57*(1), 73-86.

Mahatmya, D., Lohman, B. J., Matjasko, J. L., & Farb, A. F. (2012). Engagement across developmental periods. In S. L. Christensen, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 45-63). Springer.

Main, P. (n.d.). *Dual Coding: A Teacher's Guide*. https://www.structural-learning.com/post/dual-coding-a-teachers-guide

Malinen, O.-P., Väisänen, P., & Savolainen, H. (2012). Teacher education in Finland: a review of a national effort for preparing teachers for the future. *Curriculum Journal*, *23*(4), 567-584.

Malloy, J. A., Marinak, B. A., Gambrell, L. B., & Mazzoni, S. A. (2013). Assessing motivation to read: The motivation to read profile–revised. *The reading teacher*, *67*(4), 273-282.

Malmberg, J., Järvelä, S., & Kirschner, P. A. (2014). Elementary school students' strategic learning: does task-type matter? *Metacognition and learning*, *9*(2), 113-136.

Marinak, B. A., Malloy, J. B., Gambrell, L. B., & Mazzoni, S. A. (2015). Me and my reading profile: A tool for assessing early reading motivation. *The reading teacher*, *69*(1), 51-62.

Marks, H. M. (2000). Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American Educational Research Journal*, *37*(1), 153-184.

Marsh, H. W., Craven, R. G., Hinkley, J. W., & Debus, R. L. (2003). Evaluation of the Big-Two-Factor Theory of academic motivation orientations: An evaluation of jingle-jangle fallacies. *Multivariate Behavioral Research*, *38*(2), 189-224.

Martin, A. J. (1999-2022). *Motivation and Engagement Scale (MES) Test User Manual*. Lifelong Achievement Group.

Martin, A. J. (2001). The Student Motivation Scale: A tool for measuring and enhancing motivation. *Australian Journal of Guidance and Counselling*, 1-20.

Martin, A. J. (2007). Examining a multidimensional model of student motivation and engagement using a construct validation approach. *British Journal of Educational Psychology*, *77*(2), 413-440.

Martin, A. J. (2008a). Enhancing student motivation and engagement: The effects of a multidimensional intervention. *Contemporary Educational Psychology*, *33*(2), 239-269.

Martin, A. J. (2008b). How Domain Specific Is Motivation and Engagement across School, Sport, and Music? A Substantive-Methodological Synergy Assessing Young Sportspeople and Musicians. *Contemporary Educational Psychology*, *33*(no. 4), 785-813.

Martin, A. J. (2009). Motivation and engagement across the academic life span: A developmental construct validity study of elementary school, high school, and university/college students. *Educational and Psychological Measurement*, *69*(5), 794-824.

Martin, A. J. (2012). Part II commentary: Motivation and engagement: Conceptual, operational, and empirical clarity. In S. L. Christensen, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 303-311). Springer.

Martin, A. J. (2022). The Role of Academic Engagement in Students' Educational Development: Insights from Load Reduction Instruction and the 4M Academic Engagement Framework. In A. L. Reschly, S. L. Christenson, & A. J. Martin (Eds.), *Handbook of research on student engagement* (pp. 487-509). Springer.

Martin, A. J., Ginns, P., Anderson, M., Gibson, R., & Bishop, M. (2021). Motivation and engagement among Indigenous (Aboriginal Australian) and non-Indigenous students. *Educational Psychology*, *41*(4), 424-445.

Martin, A. J., Ginns, P., & Papworth, B. (2017). Motivation and engagement: Same or different? Does it matter? *Learning and Individual Differences*, *55*, 150-162.

Martin, A. J., Yu, K., Papworth, B., Ginns, P., & Collie, R. J. (2015). Motivation and engagement in the United States, Canada, United Kingdom, Australia, and China: Testing a multi-dimensional framework. *Journal of Psychoeducational Assessment*, *33*(2), 103-114.

Marzano, R. J. (2010). What Are the General Skills of Thinking and Reasoning and How Do You Teach Them? *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, *71*(5), 268-273. https://doi.org/10.1080/00098659809602721

Marzano, R. J., Pickering, D., & Pollock, J. E. (2001). *Classroom instruction that works: Research-based strategies for increasing student achievement*. Association for Supervision and Curriculum Development.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, *50*(4), 370-396.

Mayer, R. E. (1997). Multimedia instruction: are we asking the right question. *Educational psychologist*, *32*, 1-19.

Mayer, R. E. (2005a). Cognitive theory of multimedia learning. In *The Cambridge handbook of multimedia learning* (Vol. 41, pp. 31-48). Cambridge University Press.

Mayer, R. E. (2005b). Principles for Managing Essential Processing in Multimedia Learning : Segmenting, Pretraining, and Modality Principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 169-182). Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.012

Mayer, R. E. (2009). The Promise of Multimedia Learning. In R. E. Mayer (Ed.), *Multimedia Learning* (2 ed., pp. 3-27). Cambridge University Press. https://doi.org/DOI: 10.1017/CBO9780511811678.003

Mayer, R. E. (2022). Cognitive Theory of Multimedia Learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 57-72). Cambridge University Press. https://doi.org/10.1017/9781108894333

McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Research quarterly for exercise and sport*, *60*(1), 48-58.

McCoach, D. B. (2002). A validation study of the school attitude assessment survey. *Measurement and Evaluation in Counseling and Development*, *35*(2), 66-77.

McCrae, R. R., & Costa Jr, P. T. (1997). Personality trait structure as a human universal. *American Psychologist*, *52*(5), 509.

McGaw, B. (2013). Developing 21st century competencies through disciplines of knowledge. an International Symposium on Education and 21st Century Competencies, Muscat, Sultanate of Oman,

Mednick, S. A. (1962). The associative basis of the creative process. *Psychological Review*, *69*, 220-232.

Mednick, S. A., & Mednick, M. T. S. (1967). *Examiner's manual: Remote associates test*. Houghton Mifflin.

Meece, J. L., Wigfield, A., & Eccles, J. S. (1990). Predictors of math anxiety and its influence on young adolescents' course enrollment intentions and performance in mathematics. *Journal of educational psychology*, *82*(1), 60.

Meeker, M. (1985). *Structure of Intellect Learning Abilities Test*. Western Psychological Services.

Mergendoller, J. R., Maxwell, N. L., & Bellisimo, Y. (2006). The effectiveness of problem-based instruction: A comparative study of instructional methods and student characteristics. *Interdisciplinary Journal of Problem-Based Learning*, *1*(2), 49-69.

Merriam-Webster. (n.d.). *Continuum*. Retrieved January 30 from https://www.merriam-webster.com/dictionary/continuum

Michael, W. B., & Colson, K. R. (1979). The development and validation of a life experience inventory for the identification of creative electrical engineers. *Educational and Psychological Measurement*, *39*(2), 463-470.

Midgley, C., Maehr, M., Hruda, L., Anderman, E., Anderman, L., Gheen, M., . . . Nelson, J. (2000). Patterns of adaptive learning scales (PALS) manual. *Ann Arbor, MI: University of Michigan*.

Ministerial Council on Education, E., Training and Youth Affairs (MCEETYA),. (2008). *Melbourne declaration on educational goals for young Australians*. Ministerial Council on Education, Employment, Training and Youth Affairs.

Mislevy, R. J. (2012). Four metaphors we need to understand assessment. In *The Gordon Commission on the Future of Assessment in Education*. http://www.gordoncommission.org.

Mitchell, R. F. (2009). Current trends in classroom research. In M. Long & C. Doughty (Eds.), *The handbook of language teaching* (pp. 675-705). Wiley-Blackwell.

Mitchell, T. R. (1982). Motivation: New directions for theory, research, and practice. *Academy of Management Review*, *7*(1), 80-88.

Montali, J., & Lewandowski, L. (1996). Bimodal reading: Benefits of a talking computer for average and less skilled readers. *Journal of Learning Disabilities*, *29*(3), 271-279.

Moos, R. H. (1986). *Work environment scale manual*. Consulting Psychologists Press.

Moos, R. H., & Insel, P. M. (1986). Work environment scale.

Moreno, R., & Mayer, R. E. (2002). Verbal redundancy in multimedia learning: When reading helps listening. *Journal of educational psychology*, *94*(1), 156.

Moseley, D., Baumfield, V., Elliott, J., Gregson, M., Higgins, S., Miller, J., & Newton, D. P. (2005). *Frameworks for thinking: a handbook for teaching and learning*. Cambridge University Press.

Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of educational psychology*, *87*(2), 319.

Mumford, M. D., Mobley, M. I., Reiter-Palmon, R., Uhlman, C. E., & Doares, L. M. (1991). Process analytic models of creative capacities. *Creativity Research Journal*, *4*(2), 91-122.

Mumford, M. D., Supinski, E. P., Baughman, W. A., Costanza, D. P., & Threlfall, K. V. (1997). Process-based measures of creative problem-solving skills: V. Overall prediction. *Creativity Research Journal*, *10*(1), 73-85.

Murdoch, K. (2015). *The Power of Inquiry*. Seastar Education.

Murdoch, K. (2019). *A Model for Designing a Journey of Inquiry*. https://static1.squarespace.com/static/55c7efeae4b0f5d2463be2d1/t/5dcb82551bdcf03f365b0a6f/1573618265386/A+MODEL+FOR+DESIGNING+A+JOURNEY+OF+INQUIRY.pdf

Murphy, P. K., & Alexander, P. A. (2000). A motivated exploration of motivation terminology. *Contemporary Educational Psychology*, *25*(1), 3-53.

National Research Council Institute of Medicine (NRC). (2004). *Engaging schools: Fostering high school students' motivation to learn*. National Academies Press.

Newell, A., Shaw, J. C., & Simon, H. A. (1962). The processes of creative thinking. In H. E. Gruber, G. Terrell, & M. Wertheimer (Eds.), *Contemporary Approaches to Creative Thinking* (pp. 63-119). Atherton.

Newman, M., Ambrose, K., Corner, T., Evans, J., Morris-Vincent, P., Quinn, S., . . . Vernon, L. (2003, April 2003). *Evaluating Educational Impact: The Approach Followed by the Project on the Effectiveness of Problem Based Learning (PEPBL)* American Educational Research Association Conference, Chicago, USA.

Newmann, F. M. (1992). *Student engagement and achievement in American secondary schools*. Teachers College Press, Columbia University.

Newmann, F. M., Wehlage, G. G., & Lamborn, S. D. (1992). The Significance and Sources of Student Engagement. In F. M. Newmann (Ed.), *Student Engagement and Achievement in American Secondary Schools*. Teachers College Press, Columbia University.

Norman, G., & Schmidt, H. G. (1992). The psychological basis of problem-based learning: A review of the evidence. *Academic medicine*, *67*(9), 557-565.

Norris, S. P., & King, R. (1984). *The Design of a Critical Thinking Test on Appraising Observations. Studies in Critical Thinking, Research Report No. 1*.

O'Neil Jr, H. F., Abedi, J., & Spielberger, C. D. (1994). The measurement and teaching of creativity. In H. F. J. O'Neil & M. Drillings (Eds.), *Motivation: Theory and research.* (pp. 245-263). Lawrence Erlbaum Associates, Inc.

Office of the Chief Scientist. (2013). *Science, Technology, Engineering and Mathematics in the National Interest: A Strategic Approach*. Canberra: Australian Government

Office of the Chief Scientist. (2014). *Science, Technology, Engineering and Mathematics: Australia's Future*. Australian Government, Canberra

Olin, A., & Pörn, M. (2023). Teachers' professional transformation in teacher-researcher collaborative didactic development projects in Sweden and Finland. *Educational Action Research*, *31*(4), 821-838.

Organisation for Economic Co-operation and Development (OECD). (2020). *PISA 2021 Creative Thinking Framework (Third Draft)*. OECD. https://www.oecd.org/pisa/publications/PISA-2021-creative-thinking-framework.pdf

Organisation for Economic Co-operation and Development (OECD). (2023). *Innovating Assessments to Measure and Support Complex Skills* (M. Piacentini & N. Foster, Eds.). OECD Publishing. https://doi.org/10.1787/e5f3e341-en

Osborn, A. F. (1953). *Applied Imagination*. Scribner's.

Ostroff, W. L. (2014). Don't Just Sit There...Pay Attention!". *Educational Leadership*, 70-74.

Paas, F., & Sweller, J. (2022). Implications of Cognitive Load Theory for Multimedia Learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge Hanbook of Multimedia Learning* (3rd ed.). Cambridge University Press. https://doi.org/10.1017/9781108894333

Paivio, A., & Clark, J. M. (2006). Dual coding theory and education. Draft chapter presented at the conference on Pathways to Literacy Achievement for High Poverty Children at The University of Michigan School of Education,

Pallant, J. (2013). *SPSS survival manual*. McGraw-hill education (UK).

Parsons, A. W., Parsons, S. A., Malloy, J. A., Gambrell, L. B., Marinak, B. A., Reutzel, D. R., . . . Fawson, P. C. (2018). Upper elementary students' motivation to read fiction and nonfiction. *The Elementary School Journal*, *118*(3), 505-523.

Patston, T., Kaufman, J., & Cropley, D. H. (2022). *Creative Actions*. Hawker Brownlow.

Patston, T. J., Kaufman, J. C., Cropley, A. J., & Marrone, R. (2021). What is creativity in education? A qualitative study of international curricula. *Journal of Advanced Academics*, *32*(2), 207-230.

Paul, R. (2011). Critical Thinking Movement: 3 Waves. https://www.criticalthinking.org/pages/critical-thinking-movement-3-waves/856

Paul, R., & Elder, L. (2007). Consequential validity: using assessment to drive instruction. https://www.criticalthinking.org/files/White%20PaperAssessmentSept2007.pdf

Pearson. (2023). *Watson Glaser Critical Thinking Appraisal*. Pearson. https://www.talentlens.com/recruitment/assessments/watson-glaser-critical-thinking-appraisal.html

Pease, M. A., & Kuhn, D. (2011). Experimental analysis of the effective components of problem-based learning. *Science Education*, *95*(1), 57-86.

Pegg, J. (2010). *Promoting the acquisition of higher-order skills and understandings in primary and secondary mathematics* Research Conference 2010, https://research.acer.edu.au/cgi/viewcontent.cgi?article=1094&context=research_conference

Pegg, J. (2018). Structure of the Observed Learning Outcome (SOLO) Model. In S. Lerman (Ed.), *Encyclopedia of Mathematics Education*. Springer.

Pegg, J., & Tall, D. (2005). The fundamental cycle of concept construction underlying various theoretical frameworks. *International Reviews on Mathematical Education*, *37*, 468-475.

Pekrun, R. (2020). Commentary: Self-Report Is Indispensable to Assess Students' Learning. *Frontline Learning Research*, *8*(3), 185-193.

Perkins, D. N. (1994). *The Intelligent Eye: Learning to think by looking at art*. Getty Publications.

Permatasari, C. P., Yerizon, Y., Arnawa, I. M., & Musdi, E. (2020). Improving Students' Problem-Solving Ability through Learning Tools Based on Problem Based Learning. *Journal of Physics: Conference Series*, *1554*(1). https://doi.org/https://doi.org/10.1088/1742-6596/1554/1/012017

Perry, J. C. (2008). School engagement among urban youth of color: Criterion pattern effects of vocational exploration and racial identity. *Journal of Career Development*, *34*(4), 397-422.

Phillips, L. M. (1988). *The Design and Development of the Phillips-Patterson Test of Inference Ability in Reading Comprehension*.

Piacentini, M. (2023). *Unleashing potential: Rethinking assessment for 21st century learning* [Webinar]. OECD website, OECD Directorate for Education and Skills. https://oecdedutoday.com/oecd-education-webinars/

Pianta, R. C., Hamre, B. K., & Allen, J. P. (2012). Teacher-student relationships and engagement: Conceptualizing, measuring, and improving the capacity of classroom interactions. In S. L. Christensen, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 365-386). Springer.

Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of educational psychology*, *95*(4), 667-686.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of educational psychology*, *82*(1), 33-40.

Plucker, J. A., & Beghetto, R. A. (2004). Why Creativity Is Domain General, Why It Looks Domain Specific, and Why the Distinction Does Not Matter. In R. J. Sternberg, E. L. Grigorenko, & J. L. Singer (Eds.), *Creativity: From potential to realization* (pp. 153-167). American Psychological Association.

Plucker, J. A., Beghetto, R. A., & Dow, G. T. (2004). Why isn't creativity more important to educational psychologists? Potentials, pitfalls, and future directions in creativity research. *Educational psychologist*, *39*(2), 83-96.

Plucker, J. A., Makel, M. C., & Qian, M. (2019). Assessment of creativity. In A. S. Kaufman & J. C. Sternberg (Eds.), *The Cambridge handbook of creativity* (2nd ed., pp. 44-68).

Plucker, J. A., Qian, M., & Wang, S. (2011). Is originality in the eye of the beholder? Comparison of scoring techniques in the assessment of divergent thinking. *The Journal of Creative Behavior*, *45*(1), 1-22.

Praetorius, A.-K., Berner, V.-D., Zeinz, H., Scheunpflug, A., & Dresel, M. (2013). Judgment confidence and judgment accuracy of teachers in judging self-concepts of students. *The Journal of Educational Research*, *106*(1), 64-76.

Pringle, A., & Sowden, P. T. (2017). The Mode Shifting Index (MSI): A new measure of the creative thinking skill of shifting between associative and analytic thinking. *Thinking Skills and Creativity*, *23*, 17-28.

Puccio, G. J., Treffinger, D. J., & Talbot, R. J. (1995). Exploratory examination of the relationship between creativity styles and creative products. *Creativity Research Journal*, *8*, 152-157.

Puryear, J. S., & Lamb, K. N. (2020). Defining creativity: How far have we come since Plucker, Beghetto, and Dow? *Creativity Research Journal*, *32*(3), 206-214.

Race, P. (2010). *Making Learning Happen: A guide for post-compulsory education* (2nd ed.). Sage Publications Ltd.

Race, P. (2014). *Making learning happen: A guide for post-compulsory education* (3rd ed.). Sage.

Ramalingam, D., Anderson, P., Duckworth, D., Scoular, C., & Heard, J. (2020a). Creative Thinking: Definition and Structure. https://research.acer.edu.au/ar_misc/43

Ramalingam, D., Anderson, P., Duckworth, D., Scoular, C., & Heard, J. (2020b). *Creative Thinking: Skill Development Framework*. Australian Council for Educational Research Ltd. https://research.acer.edu.au/ar_misc/40

Reeve, J., & Jang, H. (2022). Agentic engagement. In *Handbook of research on student engagement* (pp. 95-107). Springer.

Reid, A. (2018). The journey towards the first Australian Curriculum. In A. Reid & D. Price (Eds.), *The Australian Curriculum: Promises, problems and possibilities* (pp. 3-18). ACSA.

Reisman, F., Keiser, L., & Otti, O. (2016). Development, use and implications of diagnostic creativity assessment app, RDCA–Reisman Diagnostic Creativity Assessment. *Creativity Research Journal*, *28*(2), 177-187.

Reiter-Palmon, R., Robinson-Morral, E. J., Kaufman, J. C., & Santo, J. B. (2012). Evaluation of self-perceptions of creativity: Is it a useful criterion? *Creativity Research Journal*, *24*(2-3), 107-114.

Reschly, A. L., Betts, J., & Appleton, J. J. (2014). An examination of the validity of two measures of student engagement. *International Journal of School & Educational Psychology*, *2*(2), 106-114.

Reschly, A. L., & Christenson, S. L. (2006a). Prediction of dropout among students with mild disabilities: A case for the inclusion of student engagement variables. *Remedial and Special Education*, *27*(5), 276-292.

Reschly, A. L., & Christenson, S. L. (2006b). School Completion. In G. G. Bear & K. M. Minke (Eds.), *Children's needs III: Development, prevention, and intervention* (pp. 103–113). National Association of School Psychologists.

Reschly, A. L., & Christenson, S. L. (2012). Jingle, jangle, and conceptual haziness: Evolution and future directions of the engagement construct. In *Handbook of research on student engagement* (pp. 3-19). Springer.

Reschly, A. L., & Christenson, S. L. (2022). Jingle-Jangle Revisited: history and further evolution of the Student Engagement Construct. In A. L. Reschly & S. L. Christenson (Eds.), *Handbook of research on student engagement* (pp. 3-24). Springer.

Rhodes, M. (1961). An analysis of creativity. *The Phi Delta Kappan*, *42*(7), 305-310.

Richards, R., Kinney, D. K., Bennet, M., & Merzel, A. P. C. (1988). Assessing everyday creativity: Characteristics of the Lifetime Creativity Scales and validation with three large samples. *Journal of personality and social psychology*, *54*, 476-485.

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: a systematic review and meta-analysis. *Psychological bulletin*, *138*(2), 353.

Rimm, S., & Davis, G. A. (1980). Five years of international research with GIFT: An instrument for the identification of creativity. *The Journal of Creative Behavior*, *14*(1), 35-46.

Ritchhart, R. (2002). *Intellectual character: What it is, why it matters, and how to get it*. John Wiley & Sons.

Ritchhart, R., Church, M., & Morrison, K. (2011). *Making thinking visible: How to promote engagement, understanding, and independence for all learners*. John Wiley & Sons.

Ross, J. D., & Ross, C. M. (1976). *Ross Test of Higher Cognitive Processes*. Academic Therapy Publications.

Rossi, D., Crisp, T., Andrew, J., & Pearse, S. (2019). *STEM School-Industry Partnerships*. Department for Education. https://edi.sa.edu.au/library/document-library/educational-resources/stem-learning/industry_partnership_guide.pdf

Rost, D. H. (2009). *Intelligenz: Fakten und Mythen*. Beltz Weinheim.

Rotgans, J. I., & Schmidt, H. G. (2011). Situational interest and academic achievement in the active-learning classroom. *Learning and instruction*, *21*(1), 58-67.

Rovers, S. F., Clarebout, G., Savelberg, H. H., de Bruin, A. B., & van Merriënboer, J. J. (2019). Granularity matters: comparing different ways of measuring self-regulated learning. *Metacognition and learning*, *14*, 1-19.

Royer, J. M., Sinatra, G. M., & Schumer, H. (1990). Patterns of individual differences in the development of listening and reading comprehension. *Contemporary Educational Psychology*, *15*(2), 183-196.

Runco, M. A. (1987a). The generality of creative performance in gifted and nongifted children. *Gifted Child Quarterly*, *31*(3), 121-125.

Runco, M. A. (1987b). Interrater agreement on a socially valid measure of students' creativity. *Psychological Reports*, *61*(3), 1009-1010.

Runco, M. A. (1999a). Appendix 2: Tests of Creativity. In M. A. Runco & S. Pritzker (Eds.), *Encyclopedia of creativity* (Vol. 2). Academic Press.

Runco, M. A. (1999b). Divergent Thinking. In M. A. Runco & S. Pritzker (Eds.), *Encyclopedia of creativity* (pp. 577-582). Academic Press.

Runco, M. A., & Acar, S. (2012). Divergent thinking as an indicator of creative potential. *Creativity Research Journal*, *24*(1), 66-75.

Runco, M. A., & Acar, S. (2019). Divergent Thinking. In A. S. Kaufman & R. J. Sternberg (Eds.), *The Cambridge Hanbook of Creativity* (2nd ed., pp. 224-254). Cambridge University Press.

Runco, M. A., & Jaeger, G. J. (2012). The standard definition of creativity. *Creativity Research Journal*, *24*(1), 92-96.

Runco, M. A., Plucker, J. A., & Lim, W. (2001). Development and psychometric integrity of a measure of ideational behavior. *Creativity Research Journal*, *13*(3-4), 393-400.

Runco, M. A., Walczyk, J. J., Acar, S., Cowger, E. L., Simundson, M., & Tripp, S. (2014). The incremental validity of a short form of the Ideational Behavior Scale and usefulness of distractor, contraindicative, and lie scales. *The Journal of Creative Behavior*, *48*(3), 185-197.

Ryan, R. M. (2012). *The Oxford handbook of human motivation*. Oxford University Press.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25*(1), 54-67.

Ryan, R. M., & Deci, E. L. (2022). Self-determination theory. In F. Maggino (Ed.), *Encyclopedia of quality of life and well-being research*. Springer. https://doi.org/10.1007/978-3-319-69909-7_2630-2

Sadoski, M., & Paivio, A. (2013). *Imagery and Text: A Dual Coding Theory of Reading and Writing*. Routledge. https://doi.org/10.4324/9780203801932

Saeed, S., & Zyngier, D. (2012). How motivation influences student engagement: A qualitative case study. *Journal of Education and learning*, *1*(2), 252-267.

Said-Metwaly, S., Van den Noortgate, W., & Kyndt, E. (2017). Approaches to measuring creativity: A systematic literature review. *Creativity. Theories–Research-Applications*, *4*(2), 238-275.

Salmela-Aro, K., Tang, X., & Upadyaya, K. (2022). Study demands-resources model of student engagement and burnout. In *Handbook of research on student engagement* (pp. 77-93). Springer.

Sambell, K., Brown, S., & Race, P. (2019). Assessment to support student learning: eight challenges for 21st century practice. *All Ireland Journal of Teaching and Learning in Higher Education (AISHE-J) Creative Commons Attribution-NonCommercial-ShareAlike*, *11*(2).

Santos Meneses, L. F. (2020). Critical thinking perspectives across contexts and curricula: Dominant, neglected, and complementing dimensions. *Thinking Skills and Creativity*, *35*, 1-14. https://doi.org/10.1016/j.tsc.2019.100610

Savery, J. R. (2006). Overview of problem-based learning: definition and distinctions. Interdisciplinary Journal of Problem-Based Learning,

Savery, J. R. (2015). Overview of problem-based learning: Definitions and distinctions. *Essential readings in problem-based learning: Exploring and extending the legacy of Howard S. Barrows*, *9*(2), 5-15.

Scarino, A. (2019). The Australian Curriculum and its conceptual bases: a critical analysis. *Curriculum Perspectives*, *39*(1), 59-65.

Schaefer, C. E., & Anastasi, A. (1968). A Biographical Inventory for Identifying Creativity in Adolescent Boys. *Journal of Applied Psychology*, *58*, 42-48.

Schraw, G., & Dennison, R. S. (1994). *Metacognitive Awareness Inventory*. American Psychological Association. https://psycnet.apa.org/doiLanding?doi=10.1037%2Ft21885-000

Schraw, G., & Lehman, S. (2001). Situational interest: A review of the literature and directions for future research. *Educational Psychology Review*, *13*, 23-52.

Schumm, M. F., & Bogner, F. X. (2016). Measuring adolescent science motivation. *International Journal Science Education*, *38*, 434–449.

Schunk, D. H., & DiBenedetto, M. K. (2020). Motivation and social cognitive theory. *Contemporary Educational Psychology*, *60*, 10.

Schunk, D. H., & DiBenedetto, M. K. (2022). Self-Efficacy and Engaged Learners. In A. L. Reschly & S. L. Christenson (Eds.), *Handbook of research on student engagement*. Springer.

Schwab, D. P., Olian-Gottlieb, J. D., & Heneman, H. G. (1979). Between-subjects expectancy theory research: A statistical review of studies predicting effort and performance. Psychological bulletin, 86(1), 139.

Scoular, C., & Care, E. (2018). Teaching Twenty-First Century Skills: Implications at System Levels in Australia. In *Assessment and Teaching of 21st Century Skills* (pp. 145-162). Springer.

Scoular, C., Ramalingam, D., Duckworth, D., & Heard, J. (2020). *Assessment of General Capabilities: Skills for the 21st-century Learner. Final Report*. Australian Council for Educational Research Ltd. https://research.acer.edu.au/ar_misc/47

Setyowati, S. (2022). The Effectiveness of The Problem Based Learning on Grade Five Critical Thinking Skills. *ANARGYA: Jurnal Ilmiah Pendidikan Matematika*, *5*(2), 191-196.

Shaffer, E., & Steiger, J. (1971). *Judgment: Deductive logic and assumption recognition*. Instructional Objectives Exchange.

Shakibaie, S. (n.d.). *Child's Average Attention Span By Age: From Toddler To Teens*. Ready Kids Occupational Therapy,. https://readykids.com.au/average-attention-span-by-age/

Sharp, A. M., Oscanyan, F. S., & Lipman, M. (1980). *Philosophy in the classroom*. Temple University Press.

Shaw, R. D. (2014). How Critical is Critical Thinking? *Music Educators Journal*, *101*, 65-70.

Shepherd, N. G. (1999). *The Probe Method: A problem-based learning model's affect on critical thinking skills of fourth and fifth grade social studies students* North Carolina State University].

Shernoff, D. J., Csikszentmihalyi, M., Shneider, B., & Shernoff, E. S. (2003). Student engagement in high school classrooms from the perspective of flow theory. *School Psychology Quarterly*, *18*(2), 158.

Shernoff, D. J., & Schmidt, J. A. (2008). Further evidence of an engagement–achievement paradox among US high school students. *Journal of Youth and Adolescence*, *37*, 564-580.

Shipman, V. (1983). *New Jersey Test of Reasoning Skills Form B*. Totowa Board of Education. https://digitalcommons.montclair.edu/iapc_nj_test_reasoningskills/2/

Siegel, S. M., & Kaemmerer, W. F. (1978). Measuring the perceived support for innovation in organizations. *Journal of Applied Psychology*, *63*(5), 553-562.

Silvia, P. J., Kaufman, J. C., & Pretz, J. E. (2009). Is creativity domain-specific? Latent class models of creative accomplishments and creative self-descriptions. *Psychology of Aesthetics, Creativity, and the Arts*, *3*(3), 139-148.

Skinner, E., Furrer, C., Marchand, G., & Kindermann, T. (2008). Engagement and disaffection in the classroom: Part of a larger motivational dynamic? *Journal of educational psychology*, *100*(4), 765-781.

Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection: Conceptualization and assessment of children's behavioral and emotional participation in academic activities in the classroom. *Educational and Psychological Measurement*, *69*(3), 493-525.

Skinner, E. A., & Raine, K. E. (2022). Unlocking the positive synergy between engagement and motivation. In *Handbook of research on student engagement* (pp. 25-56). Springer.

Smith, J. M., & Schaefer, C. E. (1969). Development of a creativity scale for the Adjective Check List. *Psychological Reports*, *25*(1), 87-92.

Smith, K., Maynard, N., Berry, A., Stephenson, T., Spiteri, T., Corrigan, D., . . . Smith, T. (2022). Principles of Problem-Based Learning (PBL) in STEM Education: Using Expert Wisdom and Research to Frame Educational Practice. *Education Sciences*, *12*(10), 728.

Sombatteera, S., & Kalyuga, S. (2012). When dual sensory mode with limited text presentation enhance learning. *Procedia-Social and Behavioral Sciences*, *69*, 2022-2026.

Specht, D. M. (2015). *Probe Method's Impact on Students' Motivation and Critical Thinking Skills* Walden University]. Minnesota USA.

Spinath, B. (2005). Akkuratheit der Einschätzung von Schülermerkmalen durch Lehrer und das Konstrukt der diagnostischen Kompetenz: Accuracy of teacher judgments on student characteristics and the construct of diagnostic competence. *Zeitschrift für pädagogische Psychologie*, *19*(1/2), 85-95.

Steinberg, L. (1997). *Beyond the classroom*. Simon and Schuster.

Stepien, W., & Gallagher, S. (1993). Problem-based learning: As authentic as it gets. *Educational Leadership*, *50*, 25-25. https://www.ascd.org/el/articles/problem-based-learning-as-authentic-as-it-gets

Sternberg, R. J. (1986). *Critical Thinking: Its Nature, Measurement, and Improvement*.

Sternberg, R. J., & Lubart, T. I. (1996). Investing in creativity. *American Psychologist*, *51*(7), 677-688.

Sticht, T. G., Beck, L. J., Hauke, R. N., Kleiman, G. M., & James, J. H. (1974). *Auding and reading: A developmental model*. HumRRO Press.

Strobel, J., & Van Barneveld, A. (2009). When is PBL more effective? A meta-synthesis of meta-analyses comparing PBL to conventional classrooms. *Interdisciplinary Journal of Problem-Based Learning*, *3*(1), 44 - 58.

Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of educational psychology*, *104*(3), 743-762.

Svensäter, G., & Rohlin, M. (2023). Assessment model blending formative and summative assessments using the SOLO taxonomy. *European Journal of Dental Education*, *27*(1), 149-157.

Sweller, J. (2005a). Implications of Cognitive Load Theory for Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19-30). Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.003

Sweller, J. (2005b). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 159-167).

Tate, M. W., & Brown, S. M. (1970). Note on the Cochran Q test. *Journal of the American Statistical Association*, *65*(329), 155-160.

Taylor, G., Jungert, T., Mageau, G. A., Schattke, K., Dedic, H., Rosenfield, S., & Koestner, R. (2014). A self-determination theory approach to predicting school achievement over time: The unique role of intrinsic motivation. *Contemporary Educational Psychology*, *39*(4), 342-358.

Taylor, I. A. (1975). An emerging view of creative actions. In I. A. Taylor & J. W. Getzels (Eds.), *Perspectives in creativity* (pp. 297-325). Aldine.

The Australian Industry Group. (2017). *Strengthening School-Industry STEM Partnerships: Final Project Report*. The Australian Industry Group. http://cdn.aigroup.com.au/Reports/2017/AiGroup_OCS_STEM_Report_2017.pdf

Thomas, J. W. (2000). *A Review of the Research on Project-Based Learning*. http://www.bie.org/index.php/site/RE/pbl_research/29

Thomson, S. (2021). Australia: PISA Australia - Excellence and Equity? In N. Crato (Ed.), *Improving a Country's Education*. Springer. https://doi.org/10.1007/978-3-030-59031-4

Thomson, S., De Bortoli, L., & Underwood, C. (2017). *PISA 2015: Reporting Australia's results*. ACER.

Thornhill-Miller, B., Camarda, A., Mercier, M., Burkhardt, J.-M., Morisseau, T., Bourgeois-Bougrine, S., . . . Mourey, F. (2023). Creativity, Critical Thinking, Communication, and Collaboration: Assessment, Certification, and Promotion of 21st Century Skills for the Future of Work and Education. *Journal of Intelligence*, *11*(3), 54.

Tierney, P., & Farmer, S. M. (2002). Creative self-efficacy: Its potential antecedents and relationship to creative performance. *Academy of Management Journal*, *45*(6), 1137-1148.

Tishman, S., Perkins, D. N., & Jay, E. (1995). *The thinking classroom: Learning and teaching in a culture of thinking*. Allyn & Bacon.

Torii, K. (2018). *Connecting the worlds of learning and work: Prioritising school-industry partnerships in Australia's education system*. Mitchell Institute. www.mitchellinstitute.org.au

Torrance, E. P. (1966). *Torrance tests of creative thinking: Technical norms manual*. Personnel Press.

Torrance, E. P. (1977). *Creativity in the Classroom; What Research Says to the Teacher*. National Education Association,.

Torrance, E. P. (1993). Understanding creativity: Where to start? *Psychological Inquiry*, *4*(3), 232-234.

Torrance, E. P. (1998). *Torrance Test of Creative Thinking: Norms and technical manual*. Scholastic Testing Services.

Torrance, E. P. (1999). *Torrance Tests of Creative Thinking: Norms and technical manual*. Scholastic Testing Services.

Torrance, E. P. (2008). *The Torrance Tests of Creative Thinking Norms-Technical Manual Figural (Streamlined) Forms A and B*. Scholastic Testing Service.

Torrance, E. P., & Safter, H. T. (1989). The long range predictive validity of the Just Suppose Test. *The Journal of Creative Behavior*, *23*, 219-223.

Treffinger, D. J., Young, G. C., Selby, E. C., & Shepardson, C. (2002). *Assessing Creativity: A Guide for Educators*.

Urban, K. (2005). Assessing creativity: The Test for Creative Thinking - Drawing Production (TCT-DP). *International Education Journal*, *6*(2), 272-280.

Urban, K. K. (n.d.). *Test for Creative Thinking - Drawing Production (TCT-DP) Manual*. Pearson.

Urban, K. K., & Jellen, H. G. (1996). *Test for Creative Thinking - Drawing Production (TCT-DP)*. Swets and Zeitlinger.

Urhahne, D., & Wijnia, L. (2021). A review on the accuracy of teacher judgments. *Educational Research Review*, *32*, 100374. https://doi.org/https://doi.org/10.1016/j.edurev.2020.100374

Verlaan, W., Pearce, D. L., & Zeng, G. (2017). Revisiting Sticht: The changing nature of the relationship between listening comprehension and reading comprehension among upper

elementary and middle school students over the last 50 years. *Literacy Research and Instruction*, *56*(2), 176-197.

Vernon, D. T., & Blake, R. L. (1993). Does problem-based learning work? A meta-analysis of evaluative research. *Academic medicine*, *68*(7), 550-263.

Victorian Curriculum and Assessment Authority (VCAA). (2019). *Guide to Formative Assessment Rubrics*. Melbourne, Victoria: Victorian Curriculum and Assessment Authority Retrieved from https://www.vcaa.vic.edu.au/Documents/viccurric/formative-assessment/GuideFormativeAssessmentRubrics_2019.pdf

Vincent-Lancrin, S., González-Sancho, C., Bouckaert, M., de Luca, F., Fernández-Barrerra, M., Jacotin, G., . . . Vidal, Q. (2019). *Fostering Students' Creativity and Critical Thinking*. OECD Publishing. https://doi.org/doi:https://doi.org/10.1787/62212c37-en

Voelkl, K. E. (1996). Measuring students' identification with school. *Educational and Psychological Measurement*, *56*(5), 760-770.

Walker, A., & Leary, H. (2009). A problem based learning meta analysis: Differences across problem types, implementation types, disciplines, and assessment levels. *Interdisciplinary Journal of Problem-Based Learning*, *3*(1), 12-43.

Wallach, M. A., & Kogan, N. (1965). *Modes of thinking in young children - A study of the creativity-intellegence distinction*. Rinehart and Winston.

Wallas, G. (1926). *The Art of Thought*. Butler & Tanner Ltd.

Walton, H. J., & Matthews, M. (1989). Essentials of problem-based learning. *Medical Education*, *23*(6), 542-558.

Wang, M.-T., Degol, J. L., & Henry, D. A. (2019). An integrative development-in-sociocultural-context model for children's engagement in learning. *American Psychologist*, *74*(9), 1086.

Wang, M.-T., Guo, J., & Degol, J. L. (2020). The role of sociocultural factors in student achievement motivation: A cross-cultural review. Adolescent Research Review, 5(4), 435-450.

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA statement on *p*-Values: Context, Process, and Purpose. *The American Statistician*, *70*(2), 129-133.

Watson, G. (1980). *Watson-Glaser critical thinking appraisal*. Psychological Corporation.

Watson, G., & Glaser, E. M. (1980). *Watson-Glaser critical thinking appraisal: Forms A and B; Manual*. Psychological Corporation.

Watters, J. J., Pillay, H., & Flynn, M. (2016). *Industry-school partnerships: A strategy to enhance education and training opportunities* (064695749X).

Weslake, A., & Christian, B. J. (2015). Brain breaks: Help or hindrance? *TEACH COLLECTION of Christian education*, *1*(1). https://research.avondale.edu.au/teachcollection

Wieland, P. S. (2014). *Learning a technical language: The modality effect in nursing students' retention of new vocabulary* University of Houston-Clear Lake]. Texas.

Wigfield, A., & Guthrie, J. T. (2004). Motivations for Reading Questionnaire. *Journal of educational psychology*.

Wiggins, G. (1990). The case for authentic assessment. *Practical Assessment, Research, and Evaluation*, *2*.

Wilder, S. (2015). Impact of problem-based learning on academic achievement in high school: a systematic review. *Educational Review*, *67*(4), 414-435.

Williams, J. (2013). Classroom research. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 541-554). Routledge.

Wilson, K. (2021). Exploring the challenges and enablers of implementing a STEM project-based learning programme in a diverse junior secondary context. *International journal of science and mathematics education*, *19*(5), 881-897.

Woolley, M. E., Bowen, G. L., & Bowen, N. K. (2004). Cognitive pretesting and the developmental validity of child self-report instruments: Theory and applications. *Research on Social Work Practice*, *14*(3), 191-200.

World Economic Forum. (2016). *The Future of Jobs 2016. Employment, Skills and Workforce Strategy for the Fourth Industrial Revolution*. World Economic Forum.

World Economic Forum. (2018). *The Future of Jobs Report 2018*. World Economic Forum.

World Economic Forum. (2020). *The Future of Jobs Report 2020*. World Economic Forum.

Wyatt-Smith, C. M., & Adie, L. (2018). Assessment: the trilogy of standards, evidence and judgement in Australian education reform. In A. Reid & D. Price (Eds.), *The Australian Curriculum: Promises, problems and possibilities*. Australian Curriculum Studies Association.

Yates, L., & Collins, C. (2010). The absence of knowledge in Australian curriculum reforms. *European Journal of Education*, *45*(1), 89-102.

Zeng, L., Proctor, R. W., & Salvendy, G. (2011). Can traditional divergent thinking tests be trusted in measuring and predicting real-world creativity? *Creativity Research Journal*, *23*(1), 24-37.

Zhang, X., & Bartol, K. M. (2010). Linking empowering leadership and employee creativity: The influence of psychological empowerment, intrinsic motivation, and creative process engagement. *Academy of Management Journal*, *53*(1), 107-128.

Zilbersheid, U. (2013). The historical character of human nature in Freud's theories. *The American journal of psychoanalysis*, *73*, 184-204.

Zumbach, J., Kumpf, D., & Koch, S. (2004). Using multimedia to enhance problem-based learning in elementary school. *Information technology in childhood education annual*, *2004*(1), 25-37.

Appendix A STEM Learning Evaluation Grant Report

STEM Learning Evaluation Grant Report

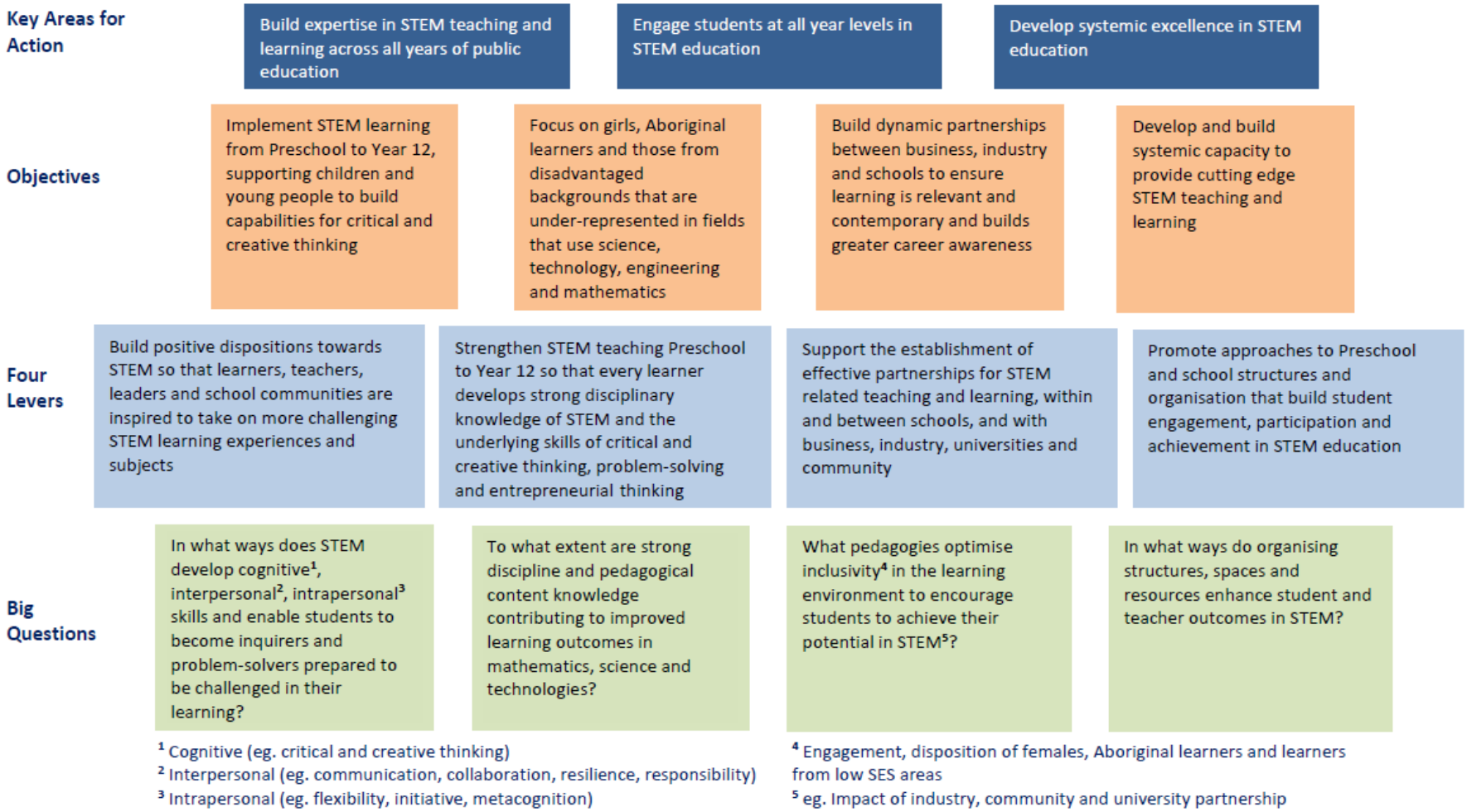
Brighton Primary School

Sue Gaardboe

May 2019

Government of South Australia
Department for Education

STEM Learning Framework: Key Areas for Action, Objectives, Four Levers and Big Questions



Key Messages

At Brighton Primary School we have been teaching STEM to Year 6 students, using partnerships with Industry to provide genuine issues for the students to relate to. The Industry Partner provides the 'problem', which forms the context in which the students develop their skills and make connections across curriculum areas.

An important part of the success in this methodology is connected to the Critical and Creative Thinking skills, as students are required to: critically question the information they are presented with; collaborate with their peers; find and process complex information; develop imaginative solutions to complex problems; consider alternatives; reflect on and justify their choices; transfer knowledge into new contexts; use logic and reasoning; draw conclusions and develop a course of action; evaluate and justify their ideas against given criteria.

Our evaluation is based primarily on the student's perceptions of their own skills (a self-assessment rubric) measured against the Critical and Creative Capabilities in the Australian Curriculum.

Key Findings and Outcomes

- Self-assessment rubrics based on the Australian Curriculum General Capabilities were developed to enable assessment of the 4 areas of Critical and Creative Thinking.
- Overall there was a **positive change** in the students' reported self-perception of their abilities in Critical and Creative Thinking, following the STEM Problem Based Learning lessons.
- The change in the boys' self-perception compared before and after the STEM PBL lessons was **statistically insignificant**.
- The change in the girls' self-perception compared before and after the STEM PBL lessons was **statistically significant.**
- The difference between the girls' self-perception of their abilities in Critical and Creative Thinking, compared to the boys' self-perception of their abilities in Critical and Creative Thinking, **prior** to their exposure to the STEM PBL lessons, was **statistically insignificant.**
- The difference between the girls' self-perception of their abilities in Critical and Creative Thinking, compared to the boys' self-perception of their abilities in Critical and Creative Thinking, **following** their exposure to the STEM PBL lessons, was **statistically significant.**
- Teachers' ability to accurately assess their students on the Critical and Creative Capabilities is questionable.

Contents

STEM Learning Framework
Key Messages
Key Findings and Outcomes *i*
Executive Summary *iii*
Acknowledgement *iii*
Introduction 1
Site Context 2
Practice Problem (or Aspiration) 2
Practice Model 2
Relationship to STEM Strategy 3
Methods 4
The Innovation 4
Data Collection Methods 5
Findings 6
Summary and Conclusion 11
Suggestions for Further Research 12
Appendix 13
Budget 18

Executive Summary

The results we have obtained indicate that this method of teaching STEM, and its potential to positively influence students' skills in Critical and Creative Thinking, deserves further investigation.

Acknowledgement

Professor David Cropley, School of Engineering, UniSA
Associate Professor David Curtis, School of Education, Flinders University
Year 6 students at Brighton Primary School
Courtney Elliot, Year 6 Teacher, Brighton Primary School
Candice Mangan, Year 6 Teacher, Brighton Primary School
Aaron Lindsay, Year 6 Teacher, Brighton Primary School
Christie Evans, STEM Coordinator, Seaview Downs Primary School
Rebecca Marrone, PhD Candidate, UniSA

Introduction

STEM Problem Based Learning (STEM PBL) was introduced by me into Brighton Primary School in Term 3 in 2018, as part of a Department for Education project to develop STEM teaching, in STEM Works Primary Schools. This is a style of Inquiry based teaching we have developed from a Problem Based Learning Model, incorporating an Industry partner to provide a genuine problem for the students to solve. Although most of the teachers involved in our STEM Problem Based Learning Project had used some form of Inquiry teaching with students previously, this project took them deeper into student-led learning than they had previously experienced. The Department for Education used our experience with the teachers to produce a promotional video about using STEM Problem Based Learning. (https://vimeo.com/318145983).

As teachers, we are expected to assess student learning across the range of curriculum subject areas, but we are also expected to be able to assess and report on student skills in the General Capabilities – for example: how our students deal with challenges; how they research questions and find information; how they develop the skills of asking relevant and probing questions; how they use logic and reasoning; and how they transfer knowledge from one experience or curriculum area to another.

STEM PBL Teaching provides the opportunity to engage in genuine problem based learning. The Industry Partner (in this project, the Holdfast Bay Council) provides the problem (How can we increase tree canopy and urban infill, while reducing heat islands in our local council area?), and the teachers identify the curriculum links within the problem (in this case the teachers made links with the science, maths and technologies curricula). They teach the subject knowledge to the students both in focussed curriculum lessons run alongside the STEM lessons (maths, science etc), and by integrating mini-lessons into the STEM lessons as questions related to the investigation are identified by the students.

During our STEM PBL project, the students were not explicitly taught skills in Critical or Creative Thinking. I have observed that, despite their position in the Australian Curriculum General Capabilities, these skills are not consistently explicitly taught in primary schools. A series of rubrics, based on the descriptions of student expectations in the Australian Curriculum General Capabilities, covering the four areas of Critical and Creative Thinking, were used by the students as self-assessment tools.

At the start of the project, the students were given the self-assessment rubrics which, depending on the teacher, were discussed with the students to varying extents. The students were asked to assess the level of skill they thought they regularly demonstrated and, where possible, to give an example of where they had demonstrated use of such a skill in the past. At the conclusion of the project the students were again asked to assess the level of skill they thought they had demonstrated, this time in their STEM PBL lessons, and

to give examples where possible. The teachers were also requested to provide an assessment based on Rubric 1, of each of the students in their class.

Site Context

Brighton Primary School is located near the beach, 15 km south of the City of Adelaide. It is a Category 6 school, catering for students from Reception to Year 7 and incorporating the Centre for Deaf Education. The school currently has 689 students enrolled across 26 classes. The staffing structure includes 4 Leadership positions, and 35 teachers in a staff of 69.

The following statistics apply:

CALD: 75 students

ATSI: 12 students

School Card: 86 families

Enrolments: 689 students

Practice Problem (or Aspiration)

Through this project I hope to address the question: Is there an effect on student awareness of Critical and Creative Thinking Skills, brought about by STEM Problem Based Learning lessons?

Practice Model

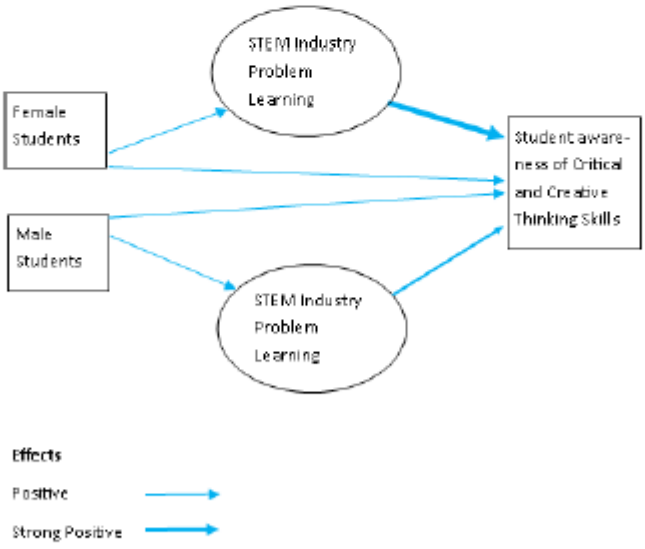


Figure 1 Diagrammatic model of student awareness of the use of Critical and Creative Thinking skills in their learning

Relationship to STEM Strategy

Our project is directly aligned to the first of the 4 Big Questions: In what ways does STEM develop cognitive, interpersonal, and intrapersonal skills and enable students to become inquirers and problem-solvers prepared to be challenged in their learning?

Inquiring, problem solving and being challenged (questioning and generating ideas) are areas of the Critical and Creative capabilities. Our project links to the STEM strategy by showing that even without explicit lessons in how to be a critical thinker, or how to think creatively, just engaging in STEM Problem Based Learning lessons, develops an increased awareness in students about critical and creative thinking skills.

Methods

Since 2015, I have been developing a method of teaching STEM to Primary School students, with my colleague Christie Evans. We have developed a model based on Problem Based Learning, which incorporates elements of the Engineering Design Process, Computational Thinking, and utilises genuine Industry generated problems. As an assessment tool, we have created rubrics aligned with the General Capabilities Learning Continuum.

The Innovation

The STEM Problem Based Learning Project at Brighton Primary School was part of a larger project (The Department for Education STEM Works Leaders Support Project) across four primary schools, designed to develop skills in teachers in STEM PBL Teaching. The three teachers at Brighton Primary School involved in the project were provided with one and a half days of professional development on how to set up and run a school STEM Problem Based Learning project.

Assessment rubrics were developed from the Australian Curriculum General Capabilities and these were shared with the teachers. The possibility of using the rubrics to assess student learning was promoted. The teachers were encouraged to decide for themselves what the focus of their student assessments would be through the project. As they were new to this style of teaching, the teachers decided to plan to assess only the curriculum areas they were teaching in parallel to the STEM PBL Teaching.

By the time this Evaluation Study began, the three teachers and seventy-nine Year 6 students had been introduced to their Industry partner and the problem they were to solve. The instructions provided to the teachers was that they were to talk through the rubrics with the students, briefly explaining the levels in each rubric. There were no specific lessons taught to the students during the period of the project relating to thinking critically or creatively.

The STEM PBL process follows a specific path:

- The Industry Partner presents the problem to the students
- Students investigate the problem by mind mapping the associated issues, identifying and developing smaller questions which they will be able to research
- Students articulate the central question or problem they will be attempting to solve
- The issues developed by the mind mapping are grouped into curriculum or topic groups, and students are divided into groups of 3 to 5 to investigate those smaller questions
- Sharing of information is managed across the groups
- Students rotate through two or more areas of research

- Once sufficient time has been spent, or sufficient information has been gathered, students re-form into new groups, focussing on proposed solutions
- Preliminary solutions are presented to the Industry Partner for feedback
- Final solutions are presented to the Industry Partner

Data Collection Methods

The tool I used to assess whether there had been any change in attitude of the students to Critical or Creative Thinking during the STEM Problem Based Learning process was a set of four rubrics we developed, based on the Australian Curriculum General Capabilities learning continuum.

The rubrics (attached in the Appendices) each have 6 levels, equating to the expectation of students' abilities by the end of Foundation; the end of Year 2; the end of Year 4; the end of year 6; the end of Year 8; and the end of Year 10. The rubrics were designed to be used as a self-assessment tool, with students encouraged to add an example of an occasion on which they believe they demonstrated the skill at the chosen level.

The first set of rubrics was provided to the teachers shortly after the STEM PBL Teaching process had commenced. The teachers were requested to present the rubrics to the students over several days, with a brief verbal description of the task and a brief explanation of the rubrics and what they meant. None of the classes managed to complete all four rubrics.

Following the completion of the STEM PBL Teaching, the four rubrics were again provided to the teachers, with the same instructions. One class managed to complete all four of these rubrics, one returned three, and the third class returned two of the rubrics. All three of the classes completed the pre-teaching and post-teaching rubrics for Rubric 1: Inquiring – identifying, exploring and organising information and ideas, and Rubric 3: Reflecting on thinking and processes. Therefore, these two rubrics form the basis for the STEM evaluation.

Some students included a descriptive comment about an event or opportunity which they felt demonstrated the level of expertise they claimed to have.

Following the completion of the project, the three teachers were asked to complete an assessment against Rubric 1 for each of the students in their class. As the project coincided with the end of the year, the teachers were encouraged to think of this tool as an assessment aligned with reporting, so they were considering the skills each student demonstrated in their learning, not necessarily just relating to the STEM project.

Findings

Of the four rubrics we provided the students with, only Rubric 1 was completed by all the classes, both before and after the intervention. Where comparisons were made between pre-intervention and post-intervention rubrics, the data was only used where a student had completed both the pre-intervention and post-intervention rubric. This provided us with data from a total of 63 students.

Rubric 3 was completed (both pre-intervention and post-intervention) by two classes, providing us with data from a total of 40 students.

As the other rubrics were not consistently completed for both pre and post-teaching, we did not include the data.

Two of the teachers completed Rubric 1 post-intervention providing results for their assessment of a total of 53 students.

In Table 1 the numbers of completed rubrics are shown, both the self-assessments completed by the students, and the assessments of the students completed by the teachers.

Table 1. Numbers of participants who completed the rubrics

	Rubric 1			Rubric 2			Rubric 3			Rubric 4		
	pre	post	Completed both pre and post	pre	post	Completed both pre and post	pre	post	Completed both pre and post	pre	post	Completed both pre and post
Students Class 1	25	24	23	0	23	0	25	22	21	23	0	0
Students Class 2	25	22	19	22	25	21	26	21	19	0	18	0
Students Class 3	24	24	21	0	23	0	22	0	0	0	0	0
Total responses	74	70	63	22	71	21	73	43	40	23	18	0
Teacher Class 1	0	0	0	0	0	0	0	0	0	0	0	0
Teacher Class 2	0	27	0	0	0	0	0	0	0	0	0	0
Teacher Class 3	0	26	0	0	0	0	0	0	0	0	0	0

In Table 2 the results of a statistical analysis of students' scores on pre- and post-intervention administrations of the rubrics are shown.

The rubrics have been scored by assigning a value where 1 = Foundation level achievement, 2 = Year 2, 3 = Year 4 etc. Where the **total** is reported, this has been achieved by adding the scores for the three elements in each rubric, enabling a maximum total score of 18. Where the **mean** is reported, this is the mean of the scores given for all three elements in the rubric, i.e. the maximum total would be six.

The intervention was completed in Term 4, so it could be expected that the students and teachers would rate the students on average at the level of student at the end of Year 6. This would give a total score of 12, and a mean of four.

The Significance column identifies which of the comparisons in the following chart are statistically significant, and therefore of interest. Where $p > 0.05$, the comparison is not statistically significant. Where $p < 0.05$, the comparison is statistically significant – these have been highlighted in grey.

Where the TOTAL is reported, this has been calculated only for students who completed all elements of the rubric. Where the MEAN has been reported, student responses were also included for students who had not provided a response to one or two elements of the rubric. These non-responses were discounted in the calculation of the mean.

Table 2. Statistical analysis of students' scores on pre- and post-intervention administration of rubrics

		Number of students	Mean (average)	Std. Deviation	Significance
Rubric 1: Inquiring TOTAL Pre-intervention	Male	32	7.28	2.888	0.089
	Female	39	8.64	3.616	
Rubric 1: Inquiring TOTAL Post-intervention	Male	31	8.16	3.023	0.019
	Female	39	10.18	3.838	
Rubric 1: Inquiring MEAN Pre-intervention	Male	33	2.3838	0.97937	0.083
	Female	41	2.8374	1.19576	
Rubric 1: Inquiring MEAN Post-intervention	Male	31	2.7204	1.00774	0.019
	Female	39	3.3932	1.27933	
Rubric 3: Reflecting MEAN Pre-intervention	Male	33	2.5354	1.13938	0.093
	Female	40	3.0125	1.23251	

Rubric 3: Reflecting MEAN Post-intervention	Male	21	3.1905	1.27179	0.047
	Female	22	3.9394	1.12516	
Rubric 3: Reflecting TOTAL Pre-intervention	Male	30	7.87	3.451	0.149
	Female	38	9.16	3.753	
Rubric 3: Reflecting TOTAL Post-intervention	Male	21	9.57	3.815	0.047
	Female	22	11.82	3.375	

From the data in Table 2 and Figure 2, we can see that, for Rubric 1. Inquiring – identifying, exploring and organising information and ideas:

- No statistically significant difference was found between male and female students' self-assessments prior to the intervention.
- Male students had a positive increase in their mean scores following the intervention, by 0.34, although this difference was not statistically significant. The effect size was small - Cohen's *d* = 0.30
- Female students had a **statistically significant positive increase** in their mean scores following the intervention, by 0.56. The effect size was small to medium - Cohen's *d* = 0.44
- **A statistically significant difference** was found between the way the male and female students rated themselves after the intervention, females scoring themselves higher than males by 0.67 levels or two points higher in their total. The effect size was medium - Cohen's *d* = 0.60

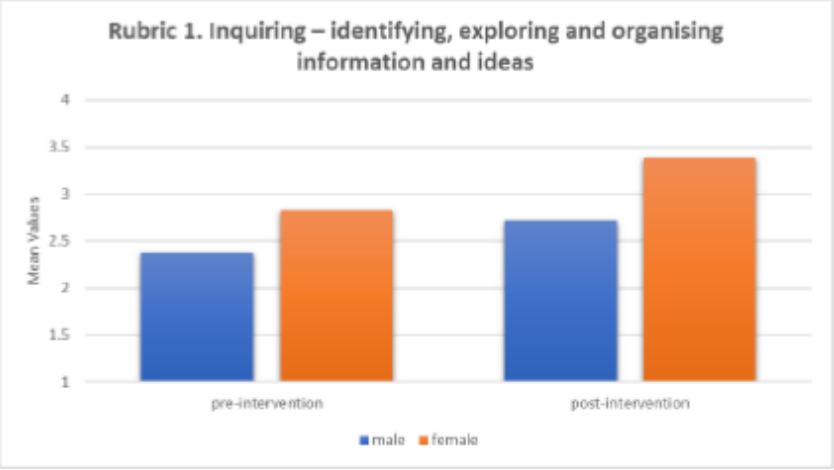


Figure 2 Comparison of Rubric 1 mean values for males and females, pre and post-intervention.

For Rubric 3. Reflecting on Thinking and Processes, we can see that:

- No statistically significant difference was found between male and female students' self-ratings prior to the intervention.
- **A statistically significant difference** was found between the way the male and female students rated themselves after the intervention, in both total and mean. Females scored themselves higher than males by 0.74 levels or two points higher in their total. The effect size was large - Cohen's *d* = 0.79 This is also shown in Figure 3.

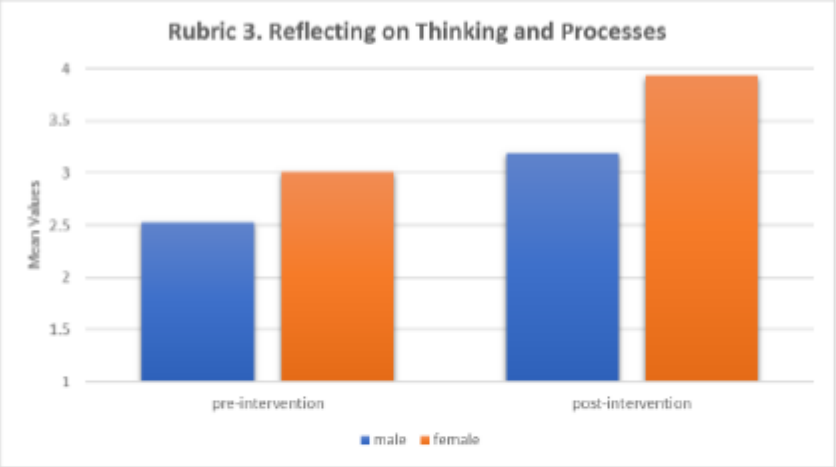


Figure 3 Comparison of Rubric 3 mean values for males and females, pre and post-intervention.

Table 3 shows the mean scores for Rubric 1 provided to the students following the intervention.

These results show that the teacher in Class 2 has given the Class 2 students a mean score 2.2 higher than the teacher in Class 3 has given the Class 3 students.

When compared as a whole cohort, the teachers have given the students a mean rating 0.5 higher than the students rated themselves.

There was no statistically significant difference in the scores given to students based on their gender, when the students were rated by teachers.

Table 3. Teacher Scoring of Rubric 1, post-intervention

		Number of students	Mean	Standard Deviation
Teacher Rubric 1: Inquiring Post-intervention	Class 2	27	4.4198	0.98533
	Class 3	26	2.2051	0.85415
Teacher Rubric 1: Inquiring Post-intervention	Class 2 & 3	53	3.33333	1.44412
Student Rubric 1: Inquiring Post-intervention	Class 2 & 3	45	2.8815	1.09903

Summary and Conclusion

There are at least three possible ways to interpret the results of the change we measured in the students' awareness of Critical and Creative Thinking Skills.

- Students skills in Critical Thinking and Creative Thinking were improved by exposing them to STEM Problem Based Learning lessons.
- Students were more accurately able to assess their own Critical Thinking and Creative Thinking skills when they completed the post-intervention rubric.
- Student understanding of the descriptions of Critical Thinking and Creative Thinking skills in the rubrics were more highly developed when they completed the post-intervention rubric.

The fact that we were able to measure a change in the student's results, indicates that there was change in one of these areas. That signifies that the teaching of STEM Problem Based Learning lessons is influencing student awareness of Critical and Creative Thinking Skills. Further investigation is warranted to discover what the change is that we are actually measuring, and how we can be more intentional in effecting a positive change.

The particularly significant change is in the difference between the way the male and female students have rated themselves, before and after the intervention. In both Rubric 1 and Rubric 3, both male and female students rated themselves at roughly the same level pre-intervention. Post-intervention, scores for both males and females increase, however, the female students have rated themselves significantly higher than their male peers. Interestingly though, the levels at which the students are rating themselves are below the level expected for their Year level.

When the teachers rated the students, compared as a whole cohort, the teachers have given the students a mean rating one level higher than the students rated themselves, i.e. the students rated themselves at a Year 2 level, while the teachers rated the students at a Year 4 level.

When looking at the ratings given to the students in their own classes, the discrepancy is significant. One teacher (Class 3), has assessed the students in their own class as working on average at a Year 2 level in Inquiring, while the other teacher (Class 2), has assessed the students in their class as working at a Year 6 level. Assuming the range of skills and abilities in the 2 classes was fairly evenly spread, the difference in mean score is significant.

Suggestions for Further Research

This discrepancy in the teacher assessments is worth investigating further. One teacher was male, one female, one an experienced teacher, one a recent graduate. Was the difference due to the level of experience of the teachers, their gender, their lack of familiarity with the rubric? Or does it illustrate the difficulties teachers find in trying to assess some of the skills in the General Capabilities?

It would also be interesting to interrogate the data further, to identify such things as:

- How the student self-assessment scores compared with the levels promoted as appropriate for their year level in the Australian Curriculum (Are the AC expectations in line with student abilities?)
- How students scored themselves against the rubrics assessing Critical Thinking, as compared with rubrics assessing Creative Thinking (Do students consistently assess themselves higher in Creative Thinking or Critical Thinking?)
- How much variation there is in scoring across the elements in a single rubric. (Are some elements of rubrics consistently over or under scored by students?)
- What are the variations between how a teacher and student score the student?
- Would some explicit lessons on how to think creatively, and how to think critically, result in more 'accurate' self-assessment scores?
- Some students included a written response to illustrate examples of occasions when they felt they had demonstrated application of the level they chose in the rubric, in their life. How do these comments by students compare with the level they chose?
- What is it about this method of teaching (or assessing) STEM that causes a marked difference between boys' and girls' results?
- How could the rubrics be made more student-friendly?
- Many of the elements in the rubrics are observed behaviours (What sort of questions do I ask? How aware am I of my own thinking processes?) which are not easily recorded during lessons. How can we accurately record the skills to enable qualitative analysis and moderation to occur?

Appendix

The following documents are the four Rubrics presented to the students, based on the Continuum of learning from the Australian Curriculum, General Capabilities.

Rubric 1. Inquiring – identifying, exploring and organising information and ideas

In each of the three areas, choose the statement that best describes what you think you do when you are working on problems at school. In the box below each area, please justify your choice e.g. by writing a statement or giving an example of an activity you have done, or questions you have asked.

1. Asking Questions	I ask questions that are about facts and things I want to know, based on my personal interests and experiences	I ask questions to identify and make sure I understand issues, and to compare information in my world	I ask questions to expand my knowledge about the world	I ask questions to make the meaning of things clear and interpret information, and to search for causes and consequences	I ask questions to identify and understand assumptions, and to investigate complex issues	I ask questions to critically analyse complex issues and abstract ideas
2. Finding and Explaining Information	I can tell you information about things I already know and ideas I have	I collect information and ideas from texts and other sources	I can identify the main ideas in information, and I can select and explain information from a range of sources	I can identify and explain relevant information, and I can organise ideas in order of importance	I can clarify information and ideas from texts of images when exploring challenging issues	I can clarify complex information and ideas drawn from a range of sources
3. Organise and process information	I can sort similar information (in words or pictures) when I am provided with the information	I can organise information that I have been provided with from more than one source, based on similar or relevant ideas	I can collect, compare and organise facts and opinions which come from a range of sources	I can analyse, summarise and combine relevant information from multiple sources	I critically analyse information and evidence by using criteria, e.g. is the information valid and relevant	I critically analyse information that I find myself, to decide the bias and reliability of the information

Rubric 2. Generating ideas, possibilities and actions

In each of the three areas, choose the statement that best describes what you think you do when you are working on problems at school. In the box below each area, please justify your choice. You may like to give an example to support your choice.

1. Imagining possibilities and connecting ideas	I use my imagination to view or create things in new ways and connect two things that seem different	I build on what I know to create new ideas.	I build on what I know and combine ideas to create imaginative solutions	combine ideas in a variety of ways and from a range of sources to create new possibilities	I see how known and new ideas are related and I use this to create new ways of achieving goals	I create and connect complex ideas using imagery, analogies and symbolism
2. Considering alternatives	I suggest alternative and creative ways to approach a given situation or task	I identify and compare creative ideas to think about a given situation or problem	I explore situations using creative thinking strategies to develop different solutions	I identify situations where current solutions do not work, challenge existing ideas and generate alternative solutions	I can come up with different and imaginative solutions and can change my ideas even if the information is limited	I think about creative options to change my ideas when I am presented with new information
3. Seeking solutions and putting ideas into action	I can predict what might happen in a given situation and when I put ideas into action	I investigate options and predict possible outcomes when putting ideas into action	I experiment with a range of options when looking for solutions and putting ideas into action	I assess and test options to identify the most effective solution and to put ideas into action	I can identify what might happen and test my predictions when finding solutions and putting ideas into action	I am aware of other people's views and consider these when I think about the risks and alternatives, when looking for solutions and putting complex ideas into action

Rubric 3. Reflecting on Thinking and Processes

In each of the three areas, choose the statement that best describes what you think you do when you are working on problems at school. In the box below each area, please justify your choice. You may like to give an example to support your choice.

1. Thinking about my thinking (Metacognition)	I can describe what I am thinking and can give reasons for my thinking	I can explain the thinking strategies I use in particular situations and tasks	I think about the way I think, and can explain and check the decisions I make	I think about the assumptions I make which affect what I think, and I listen to and think about reasonable criticism of my work	I think about the assumptions I make which affect my thinking, and I ask others for alternative opinions	I give reasons to support my thinking, I can talk about ideas which are different from mine, and I can identify areas in my own thinking which are not clear
2. Reflecting on Processes	I can identify the main parts of the steps in a thinking process	I can outline the details and order of activities in a task, and I can separate the task into smaller parts	I can identify the important information in an investigation, and separate it into smaller parts or ideas	I can identify the thinking behind choices I have made, and can justify those choices	I can explain why I choose a particular problem-solving strategy, and can discuss if it was a successful strategy to choose	I can balance rational and irrational components of a complex problem to evaluate evidence
3. Transferring knowledge into new contexts	I can make connections between information I learn in different places	I can use information I learned before, to understand or create a new idea	I can use information I learn in one place, to understand something in another place	I can use information I know from one subject or place, to learn and understand something different in another subject or place	I can explain why I choose what previously learned information I should use, when I am working in a different subject area or place.	I identify, plan and justify transference of knowledge to new contexts

Rubric 4. Analysing, synthesising and evaluating reasoning and procedures

In each of the three areas, choose the statement that best describes what you think you do when you are working on problems at school. In the box below each area, please justify your choice. You may like to give an example to support your choice.

1. Apply logic and reasoning	I can talk about the thinking I use when I am solving problems I have been given	I can talk about the reasoning I use when I am making choices or deciding on an action in a specific situation	I use different reasoning and thinking strategies in different situations, and can tell you why I chose them	I can decide whether there is enough reasoning and evidence to justify a claim, conclusion or outcome	I can identify gaps in reasoning and missing elements in information	I can analyse reasoning used in finding and applying solutions, and in choice of resources
2. Draw conclusions and design a course of action	I can share my thinking about possible courses of action	I can suggest different ways to do things, or possible different conclusions when I am presented with new information	I can use things I already know and use evidence when choosing a course of action or drawing a conclusion	I think deeply about ideas or concepts, test conclusions and use that information to change my actions when I am designing a course of action	When I have a plan of action I can explain the different parts of the plan, and can cope with information that may be confusing when I am making conclusions	I use logical and abstract thinking to analyse and synthesise complex information to inform a course of action
3. Evaluate procedures and outcomes	When I finish doing something, I make sure I am satisfied with the results	I can evaluate whether I have managed to do what I set out to achieve	I can explain and justify ideas and outcomes	I can evaluate how effective ideas, products, performances, methods and courses of action are against criteria that I have been given	I can explain what my intentions are; I can give reasons for my ideas, methods and courses of action; and I can explain both the expected and unexpected things that happen against criteria I have identified	I can evaluate the effectiveness of ideas, products and performances and implement courses of action to achieve desired outcomes against criteria I have identified

Budget

Item	How item relates to your evaluation reporting	Amount
13 x TRT days	To provide PD to teachers in STEM Problem Based Learning	$7,085
15 x TRT days	Design of rubrics, collection and analysis of data, preparation of report	$8,175
15 iPads	To enable students to engage in research, and with a digital format of the self-assessment rubric	$14,225
	TOTAL	$29,485

Appendix B Measures of Critical Thinking

Name	Skills	Reference	Age group	Format/ Focus
California Critical Thinking Disposition Inventory (CCTDI)	Truthseeking Open-mindedness Analyticity Systematicity Critical Thinking Self Confidence Inquisitiveness Cognitive Maturity	Facione et al. (1994); Insight Assessment (2023a)	Tertiary	Self-report On-line Disposition to using CT. Does not measure CT ability or skill
California Critical Thinking Skills Test: (CCTST) [1,2,3]	Analysis Evaluation Inference Inductive reasoning Deductive reasoning	Facione and Facione (1994); Insight Assessment (2023b)	Tertiary + upper secondary	Multiple choice Self-report On-line
California Measure of Mental Motivation (CM3)	Disposition towards critical thinking	Giancarlo et al. (2004)	Secondary	
Cornell Class Reasoning Test[1]	Deductive class reasoning	Ennis, Gardiner, Morrow, et al. (1964)	Grades 4-14	
Cornell Conditional Reasoning Test[1]	(Deductive) Conditional reasoning	Ennis, Gardiner, Guzzetta, et al. (1964)	Grades 4-14	
Cornell Critical Thinking Tests Level X [1,4] (CCTT-X)	Induction, deduction, Identification of assumptions, credibility	Ennis et al. (1985)	Grade 5 and up	Multiple choice On-line or paper Requires reading 50 mins

Name	Skills	Reference	Age group	Format/ Focus
Cornell Critical Thinking Tests Level Z [1,4] (CCTT-Z)	Induction, deduction, Identification of assumptions, credibility, semantics, definition, prediction in planning	Ennis et al. (1985)	Grade 5 and up	Multiple choice On-line or paper Requires reading 50 mins
Educate Insight Grades 3-5 (Form of CCTST)	Habits of mind + Inductive Reasoning, Deductive Reasoning, Evaluation, Analysis, Inference, Numeracy	Insight Assessment (2023c)	Grades 3-5	On-line Self report + scenario based 75 mins
Educate Insight Grades 6-8 (Form of CCTST)	Habits of mind + Inductive Reasoning, Deductive Reasoning, Evaluation, Analysis, Inference, Numeracy	Insight Assessment (2023d)	Grades 6-8	On-line Self report + scenario based 75 mins
Ennis-Weir Critical Thinking Essay Test[1]	Getting the point, seeing the reasons and assumptions, stating one's point, offering good reasons, seeing other possibilities, responding to and avoiding equivocation, irrelevance, circularity, reversal of conditional relationships, overgeneralization, credibility	Ennis and Weir (1985)	Grade 7 and Upwards + Tertiary	Paper Requires reading and writing context outdated—letter to newspaper, topic not relevant to children—parking issues, driving, etc.
Judgement: Deductive Logic and Assumption Recognition[1]	Deduction, assumption identification, credibility	Shaffer and Steiger (1971)	Grades 7-12	Not available
Logical Reasoning[1]	Tests for facility with class reasoning	Hertzka and Guilford (1955)	Secondary + tertiary	

Name	Skills	Reference	Age group	Format/ Focus
Metacognitive Awareness Inventory (MAI)	Declarative knowledge, procedural knowledge, conditional knowledge, planning, information management, strategies, monitoring, debugging strategies, evaluation	Schraw and Dennison (1994)	Adults	Self-report Only measures Metacognitive awareness. 10 mins
New Jersey Test of Reasoning Skills[1]	Syllogism, assumption identification, induction, good reasons	Shipman (1983)	Grade 4 and above + tertiary	Paper Multiple choice Requires reading
Online Critical Thinking Basic Concepts Test	Analysis of thought, assessment of thought, dispositions of thought, skills and abilities of thought, obstacles or barriers to critical thought.	Paul and Elder (2007)	High School (Grade 10) and above	On-line
Ross Test of Higher Cognitive Processes[1]	Verbal analogies, deduction, assumption identification, word relationships, sentence sequencing, interpreting answers to questions, information sufficiency and relevance in mathematics problems, analysis of attributes of complex stick figures	Ross and Ross (1976)	Grades 4-6	Not available
Test of Enquiry Skills[1]	Using reference materials, interpreting and processing information, thinking in science	Fraser (1979)	Australian grades 7-10	Not available
Tests of Inference Ability in Reading Comprehension[1]	Ability to infer information, Interpretations from short passages	Phillips (1988)	Grades 6-8	Multiple choice version + Constructed response version
Tests on Appraising Observations[1]	Ability to judge the credibility of statements of observation.	Norris and King (1984)	Grades 7-14	Multiple choice and constructed response versions

Name	Skills	Reference	Age group	Format/ Focus
Watson-Glaser Critical Thinking Appraisal[1] (W-GCTA)	Inference, assumptions, deduction, interpretation, evaluation of arguments	Watson (1980); Watson and Glaser (1980) Pearson (2023)	16 years+	On-line Paper version 30 mins Random generation of questions from item-bank

Notes. [1]Ennis (1993) [2]One of the most used tests of CT according to Davies (2019) [3]Hatcher (2015) [4]Hasinger (2023)

Appendix C Measures of Creativity, Creative Potential, Creative Thinking and Creative Problem Solving

Name	Author - refs	Format	Focus	Scale reliability (alpha coefficients)	Table 14 Cell References
Abedi-Schumacher Creativity Test	O'Neil Jr et al. (1994)	Multiple choice, self-report, 75-items	Measures the underlying traits of fluency, flexibility, originality and elaboration. Translated from Persian to English and Spanish		G
ACER general capabilities assessment	Scoular et al. (2020)	Digital format, mostly constructed-response items. Manually marked	Collaborative problem-solving skills on domain-oriented tasks. Critical thinking, collaboration, creative thinking		E & K
ACL: Adjective Check List	Gough (1979); Gough and Heilbrun (1983); Smith and Schaefer (1969)	Self-report or observer-judged 27-item. Secondary school	Creative personality scale. Subset of adjectives to describe more or less creative high-school students		H
BICB: Biographical Inventory of Creative Behaviors	Batey (2007); Furnham et al. (2008) https://osf.io/4s9p6/	Self-report	Everyday achievements 34 everyday creative activities summed to produce one rating		J
Big Five Personality Model (see NEO)					
Biographical Inventory	Schaefer and Anastasi (1968)	Self-report. 165 items; multiple-choice, open-ended, selecting alternatives	Biographical inventory. Measure of 5 areas: Family background, intellectual and cultural orientation, motivation (pervasive and continuing enthusiasm), breadth of interest, and drive towards novelty and diversity. Scores provided for Artistic Creativity and Scientific Creativity.	0.35-0.64	G & J

Name	Author - refs	Format	Focus	Scale reliability (alpha coefficients)	Table 14 Cell References
CAQ: Creative Achievement Questionnaire	Carson et al. (2005) https://osf.io/4s9p6/	Self-report, 8-point scale	Measure of creative ideation at pro-c level 96 achievements assessed, 10 domains covered: visual arts, music, dance, architectural design, creative writing, humor, inventions, scientific discovery, theatre and film, culinary arts		P
CAT: Consensual Assessment Technique	Amabile (1982); (1983, 1996); Hennessey (1994)	Judged by panel of experts	Experts apply their own subjective understandings of specified qualities to assess a product	.73-.93	R
CBI: Creative Behavioural Inventory	Hocevar (1979)	Self-report, 90 activities	Everyday achievements 6 domains: literature, music, craft, art, maths/science, performing arts. Single score		J
CBI: Creative Behavioural Inventory (Revised version)	Dollinger et al. (2004); Kirschenbaum (1989) 1989	Self-report, 28 items observatory ratings	Everyday achievements rated as one single, general factor.	Original test 0.93	J & K
CCL: Creativity checklist	Johnson (1979)	5-point scale (never to consistently), observatory ratings	Fluency, flexibility, constructional skills, ingenuity, resourcefulness, independence, positive self-referencing, preference for complexity.		Q
CDQ: Creativity Domain Questionnaire	Kaufman et al. (2009)	Self-report, Creativity in domains rated on a 6-point scale 56 items	interpreted as 5 factors: Science, Social, Visual Arts, Verbal Arts, and Sports; or 7 factors: Artistic-Verbal, Artistic-Visual, Entrepreneur, Interpersonal, Math/Science, Performance, and Problem-Solving		P
CDQ (revised): Creativity Domain Questionnaire – revised	Kaufman et al. (2010)	Self-report, 21 items	Math/Science, Drama, Interaction, and Visual Arts		P

Name	Author - refs	Format	Focus	Scale reliability (alpha coefficients)	Table 14 Cell References
Collaborative problem-solving online assessment	Australian Curriculum Assessment and Reporting Authority (ACARA) (2018)	Digital tool. 4 tasks	Problem-solving ability, Self-management, social management, cognitive – inquiring, cognitive – analysing, cognitive – generating		K
CPES: Creative Process Engagement Scale	Zhang and Bartol (2010)	Self-report	Frequency of problem identification, information searching, and idea generation		A
CPI: California Psychological Inventory (revised)	Gough (1987)	434 true-false questions, self-report	A broadband psychological measure designed to assess personal attributes relevant to everyday life. 20 scales including: flexibility, well-being, self-control, empathy, and independence		J
Cps: Creative Personality Scale	Gough (1992); Gough and Heilbrun (1983)	Self-report, 18 adjectives (positive weight) and 12 (negative weight). For use with adults.	Based on the ACL. Differentiates between creative and less creative adults.	0.80	J
Creative Activities Checklist	Runco (1987b)	Self-report. Grades 5-8	Biographical inventory. Frequency of engagement in: literature, music, drama, arts, craft, science. OR Three most creative achievements listed and rated for degree of creativity	0.50-0.90	J & P
Creative Product Inventory	Taylor (1975)		Product assessed on: Generation, Reformulation, Originality, Relevancy, Hedonics, Complexity and Condensation		Q
Creative Product Semantic Scale	Besemer and O'Quin (1999)	Assessed by raters using a semantic-differential rating scale–43 items	Product assessed on: Novelty, Resolution, Elaboration and Synthesis	0.69-0.87	Q

Name	Author - refs	Format	Focus	Scale reliability (alpha coefficients)	Table 14 Cell References
Creative Response Evaluation at Work	Kaufman and Reiter-Palmon (2017)	Self-report	Creative initiation and creativity styles in work-based scenarios		M
Creative Self-Efficacy	(Tierney & Farmer, 2002)	Self-report, 5-point scale	Self-belief e.g., 'I am good at coming up with new ideas'; 'I have a good imagination'		J
Creativity Tests for Children	Guilford (1976)	10 tests either semantic (verbal) or visual and figural (non verbal)	Suitable for grades 4-6, Based on SI (Structure of Intellect) model of intelligence, focus on divergent production (fluency)	0.42-0.97	B
C&RT: The Creatrix Inventory	Byrd (1986)	Self-report, two blocks of 28 items, 9-point scale	Incorporates cognitive and noncognitive dimensions (thinking and motivation) of creativity, to identify style of idea production: reproducer, modifier, challenger, practicalizer, innovator, synthesizer, dreamer, planner.		A & G
CRT: Creative Reasoning Test	Doolittle (1990)	Level A: grades 3-6; Level B: secondary and college. Forms A & B available. 20 items	Presented as riddles. Requires associative, inductive and divergent thinking	0.63-0.99	E
CSDS: Creative Solution Diagnosis Scale	Cropley and Kaufman (2012)	Paper and digital forms, 24 items, 5-point scale	Assessment of functional product creativity Relevance & effectiveness; novelty; elegance; genesis. Consensual assessment with non-expert judges		Q
CSQ: Creativity Styles Questionnaire	Kumar et al. (1997)	72-items, 3-point scale.	7 dimensions: Belief in unconscious processes, Use of techniques, Use of other people, final product orientation, environmental control, superstition, use of senses.	0.45-0.83	J

Name	Author - refs	Format	Focus	Scale reliability (alpha coefficients)	Table 14 Cell References
Domino Creativity Scale (subscale of ACL)	Domino (1994)	59-item, observer-rated, tertiary	Discriminates between inventors and non-inventors	0.88-0.91	H
EPoC: Évaluation du potential créatif (also e-EPoC)	Barbot et al. (2016); Barbot et al. (2011); Coulaud (2013); Lubart et al. (2011)	2 sessions, one week apart. Child to produce original, relevant ideas	Creative thinking and production through divergent-exploratory thinking and convergent-integrative thinking (creative synthesis). Domains for testing include: artistic-graphic; verbal-literary; social problem-solving; scientific, mathematics and music composition. Focus is on the potential for creativity.		B
GIFT: Group Inventory for Finding Creative Talent	Davis and Rimm (1982); Rimm and Davis (1980)	K-grade 2 (32-item), grade 3-4 (34-item), grade 5-6 (33-item) Self-report (yes/no)	Also GIFFI 1 and GIFFI 2: Group Inventory for Finding Interests. Measures of attitudes and interests associated with creativity: curiosity, originality, independence, flexibility, risktaking	0.80-0.88	H
ICAA: Inventory of creative activities and achievements	Diedrich et al. (2018) https://osf.io/4s9p6/	Self-report, 48 items	Creative achievement and activity in personal and public domains 8 domains: Literature, music, arts and crafts, creative cooking, sports, visual arts, performing arts, science and engineering		P
ICID: Inventory of Children's Individual Differences	Deal et al. (2007)	Parental ratings	Halverson et al 2003		H
III: Iowa Inventiveness Inventory	Colangelo et al. (1992)	Self-report 61 items, 5-point scale Revised version-grade 5 to adult	Distinguishes between acknowledged creative individuals, young inventors and non-inventive academically-talented adolescents. Measures biographical, personality, and vocational variables associated with mechanical inventiveness	0.66 – 0.72	P

Name	Author - refs	Format	Focus	Scale reliability (alpha coefficients)	Table 14 Cell References
KAI: Adaption-Innovation Inventory	Kirton (1989); Puccio et al. (1995)	Self-report 32 items, 5-point scale	Distinguishes between people who use what they know and can do (adaptors) and those who try to reorganise and restructure the problem (innovators). Provides overall score and 3 subscales: originality, conformity, efficiency.	Total score: 0.86-0.88 Sub-scales: 0.61-0.83	J
K-DOCS: Kaufman Domains of Creativity Scale	Kapoor et al. (2021); Kaufman (2012); https://osf.io/4s9p6/	Self-report of creativity compared to peers	50 activities, 5 factors reported: Every-day, Scholarly, Performance, Math/ Science, and Artistic		J & P
KEYS: Assessing the Climate for Creativity	Amabile et al., 1996	78-item questionnaire	10 scales reported, including Environmental Stimulants, Environmental Obstacles, Creativity and Productivity		M
LCS: Lifetime Creativity Scale	Richards et al. (1988)	Self-report (informal interview)	Extent of involvement in creative activities, level of creative achievement Creative activities and achievements		J & P
LEI: Life Experience Inventory	Michael and Colson (1979)	Self-report, 100 items	Biographical inventory. Four areas covered: self-striving or self-improvement, parental striving, social participation and social experience, independence training	0.62	J
Life-space questionnaire	Ivcevic (2007, 2009); Ivcevic and Mayer (2009)	Self-report, 244 item (long version) 59 item (short version)	Frequency of engagement in: everyday creativity/creative life style, artistic creativity, intellectual Achievement/ creativity 13 types of activities: crafts, cultural refinement, self-expressive creativity, interpersonal creativity, sophisticated media use, visual arts, music, dance, drama/theatre, writing, science, technology, teaching/academic		G

Name	Author - refs	Format	Focus	Scale reliability (alpha coefficients)	Table 14 Cell References
MSI: Mode Shifting Index	Pringle and Sowden (2017) https://osf.io/4s9p6/	Self-report, 14 items	Ability to shift between analytic and associative thought in professional or everyday life 4 scales: shifting competence in everyday or professional context; metacognitive awareness of shifting in everyday or professional context		A
NEO-PI-R: The Revised NE0 Personality Inventory	Costa Jr and McCrae (1992); (Costa Jr & McCrae, 2008); McCrae and Costa Jr (1997)	Self-report, 240 items, 5-point scale	Neuroticism, Extraversion, Openness, Agreeableness and Conscientiousness 5 factors (neuroticism, extraversion, openness to experience, agreeableness, conscientiousness), 6 sub-components to each		G
PSI: Problem-Solving Inventory	Heppner and Petersen (1982)	Self-report, 32 items, 6-point scale	Problem-solving confidence, Approach avoidance style, Personal control		J
(Problem-solving)	Mumford et al. (1997)	5-point scale rated by 5 judges. Responses to problem solving tasks rated	Divergent and convergent thinking. Problem construction, Information encoding, Category selection, Category Combination, Reorganization.	0.32-0.81	K
Problem-Solving Inventory (modified) (PSI)	Curtis and Denton (2003)	self-report, 32 items, 4-point scale. Demonstration of problem solving performance (self-report and facilitator validation): 18 items with multiple level descriptions	Problem-solving ability Defining the problem, planning an approach, carrying out a plan, monitoring progress, reflecting on the result. For use as summative assessment and instructional tool		J & K

Name	Author - refs	Format	Focus	Scale reliability (alpha coefficients)	Table 14 Cell References
RAT: Remote Associates Test	Mednick (1962); Mednick and Mednick (1967)	30 items, 40 minute time limit	Divergent and convergent thinking by identifying remote associates to stimulus words. Out of print.	0.91-.092	H
RDCA: Reisman Diagnostic Creativity Assessment	Reisman et al. (2016)	Mobile phone self-report app, 40 items	divergent thinking (fluency, flexibility, elaboration, originality, resistance to premature closure)11 factors: fluency, flexibility, elaboration, originality, resistance to premature closure, tolerance of ambiguity, convergent and divergent thinking, risk taking, intrinsic and extrinsic motivation. Identifies creative strengths rather than potential		J
RIBS: Runco Ideation Behavior Scale revised and short form	Runco et al. (2001); Runco et al. (2014) https://osf.io/4s9p6/	Self-report, 74 items (revised), 19 items (short form), 5-point scale	Creative ideation through divergent thinking, including originality, fluency, and flexibility Single factor generated	.92	J
SOI: Structure of the Intellect	Guilford (1967); Guilford and Hoepfner (1966); Plucker et al. (2019)		Divergent production: 24 types of DT		B
SOI:ELCT: Structure of the Intellect Learning Abilities Test: Evaluation, Leadership, and Creative Thinking	Meeker (1985)		8 cognitive activities involving divergent thinking: divergent symbolic relations, divergent symbolic units, divergent figural units, divergent semantic units, divergent semantic relations, divergent semantic transformations, divergent figural relations, divergent figural transformations	Up to 0.99	B

Name	Author - refs	Format	Focus	Scale reliability (alpha coefficients)	Table 14 Cell References
SOQ: Situational Outlook Questionnaire	Isaksen et al. (1999)	Self-report, 49 items, 4-point scale	Assesses creative climate and change, 9 dimensions		M
SSCS: The Short Scale of Creative Self-efficacy and Identity	(Karwowski et al., 2019); Karwowski et al. (2018) https://osf.io/4s9p6/	Self-report, 11 items	creative self-efficacy and creative personal identity		J
SSSI: Seigel Scale of Support for Innovation	Siegel and Kaemmerer (1978)	Self-report, 61 items, Likert-style responses	To enable comparison between organisations and units within an organisation, on the organizational climate, particularly pertaining to innovation. Reports on 3 factors: support of creativity, tolerance of differences and personal commitment.		M
TCI: Team Climate Inventory	Anderson and West (1998)	5-factor, 38 items Self-report	Climate for innovation with focus on the proximal work group rather than the organisation as a whole		M
TCT-DP: Test for Creative Thinking - Drawing Production	Urban (2005); Urban (n.d.); Urban and Jellen (1996)	Drawn product (image production) as response to figural fragments. Forms A & B available	Creative potential including divergent thinking. 14 evaluation criteria (continuations, completion, new elements, connections with line, connections for a theme, fragment dependent boundary, fragment independent boundary breaking, perspective, humour and affectivity, unconventionality (4 types), speed) To identify high creative potentials as well as low creative, neglected, and poorly developed ones. Single score provided	0.70-0.90	B

Name	Author - refs	Format	Focus	Scale reliability (alpha coefficients)	Table 14 Cell References
TTCT: Torrance Tests of Creative Thinking	Torrance (1966, 1999); (Torrance & Safter, 1989)	Picture construction, Verbal section (6 activities) and nonverbal or figural section (3 activities). Forms A & B available	Divergent thinking: Fluency, flexibility, originality, abstractness of titles, resistance to premature closure. Creative potential	.60-.90	B
WEI: Work Environment Inventory. Renamed KEYS	Amabile and Gryskiewicz (1989)	Paper form	An organizational development instrument to improve the climate for creativity. Assesses stimulants and obstacles to creativity in the work environment.		M
WES: Work Environment Scale	Moos (1986); Moos and Insel (1986)	Self-report	comprises 10 subscales. Used to evaluate productivity, assess employee satisfaction, and clarify employee expectations to ensure a healthy work environment.		M
WKCT: Wallach and Kogan Creative Thinking Test (p-WKCT & e-WKCT)	Lau and Cheung (2010); Wallach and Kogan (1965)	Paper and digital forms. 3 verbal sub-tests & 2 sub-tests with verbal stimuli	Divergent thinking skills: Verbal and figural fluency, flexibility, unusualness, and uniqueness	0.83-0.92	B

Appendix D Measures of Motivation and Engagement

Name	Motivation/ Engagement	Author - refs	Age group	Format	Focus
AMS: Academic Motivation Scale	M	Fortier et al. (1995)	Secondary	Self-report	General motivational orientation toward school. Measures: perceived academic competence, perceoived academic self-determination, autonomous academic motivation,
CAIMI: Children's Academic Intrinsic Motivation Inventory	M	Gottfried (1986); Gottfried (2019)	Elementary and middle school		Subject specific scales (Reading, Math, Social Studies, and Science + school in general)
ESM: Experience Sampling	E	Csikszentmihalyi (1990); Hektner et al. (2007); Shernoff et al. (2003); Shernoff and Schmidt (2008)	Secondary	9-point scale	Data collected on what students think and feel while instruction is occurring (i.e., 'in the moment' rather than retrospectively as with self-report). Randomly spaced responses not less than 20mins apart.
EvsD: Engagement vs. Dissafection with Learning – Student Report	E (behavioural + emotional)	Furrer and Skinner (2003); Skinner et al. (2008); Skinner et al. (2009)	Grades 3 - 6	Student self-report, (24 items), Teacher rating scales (16 items)	Provides a motivational conceptualization of engagement and disaffection
IMI: Intrinsic Motivation Inventory	M	McAuley et al. (1989)	Secondary		Intrinsic motivation and self-regulation, Sport and exercise, PjBL

Name	Motivation/ Engagement	Author - refs	Age group	Format	Focus
ISQ: Identification with School Questionnaire	E	Voelkl (1996)	secondary	Self-report, 17 items, 4-point scale	Assesses the extent to which students identify with or disengage from school
MES: Motivation and Engagement Scale MES-JS (Junior School)	M + E (behavioural, cognitive and emotional)	Martin (1999-2022, 2009); Martin et al. (2021)	SMES-JS (Junior School)	Self-report, 44 items, 5-point scale	Measures positive motivation, positive engagement, negative motivation, and negative engagement.
MMRP: Me and My Reading Profile	M	Marinak et al. (2015)	Kindergarten to grade 2	Multiple choice, Likert scale (3 responses), oral presentation	Reading motivation
Motivations for Reading Questionnaire	M	Wigfield and Guthrie (2004)	Elementary and middle school	Self-report	Reading motivation
MRP: Motivation to Read Profile	M	Gambrell et al. (1996); Malloy et al. (2013)	Elementary	Self-report, oral administration + conversational 1:1 interview	an approach that incorporates some of the principles of authentic assessment, combining with some of the principles of phenomoneology in the form of an open-ended interview.

Name	Motivation/ Engagement	Author - refs	Age group	Format	Focus
MRP-F MRP-NF	M	Parsons et al. (2018)	Grades 3 to 6	Oral presentation, 4-point ordinal scale response	motivation for reading fiction and non-fiction. Developed from the MRP, MRP-R, and Me and My Reading Profile
MRP-R (Revisions from the original MRP include updating language)	M	Malloy et al. (2013)	Grades 2 through 6	Self-report, oral administration + conversational 1:1 interview	students' perceived value of reading, self-concept as a reader, motivation to read
MSLQ: Motivated Strategies for Learning Questionnaire	E (cognitive) + M	Pintrich and De Groot (1990)	Middle school	Self-report, 56 items, 7-point scale	self-efficacy, intrinsic value, test anxiety, cognitive strategy use, self-regulation
PALS: Patterns of Adaptive Learning Scales	M	Midgley et al. (2000)	Elementary and middle school	5-point scale, verbal presentation, self-report	social and cognitive aspects of motivation, such as achievement goals, achievement-related beliefs, and academic behaviors and strategies
RSCS: Reading Self-Concept Scale Also RSCS-30	M/E	Chapman and Tunmer (1995, 1997) 1992	Elementary	Self-report, 50 items, 5-point scale	self-concept and self-efficacy particular to the domain of reading

Name	Motivation/ Engagement	Author - refs	Age group	Format	Focus
RSPS: Reader Self-Perception Scale	M/E	Henk and Melnick (1995)	Elementary	Self-report, 33 items, 4 x 5-point scales	Measures reading self-efficacy. Reports on 4 scales: Progress, Observational Comparison, Social Feedback, Physiological States
SAAS: School Attitude Assessment Survey	M	McCoach (2002)	Secondary	Self-report, 20 items	Self-concept, self-motivation, self-regulation, attitude toward school
SchMOT: developed from the SMQ	M	Conradty and Bogner (2022)	Secondary/ Elementary	Self-report, 5-point scale,	Intrinsic and extrinsic motivation with focus on general school motivation rather than domain specific. 5 subscales: self-determination, self-efficacy, intrinsic motivation, grade motivation, career motivation
SEI: Student Engagement Instrument	E (emotional and cognitive) + M (intrinsic)	Anderson et al. (2004); Appleton et al. (2006)	Grade 6 – 12 (middle and high school)	Self-report, oral administration	Cognitive and psychological engagement by expanding on behavioral and academic indicators
SEI-E: Student Engagement Instrument	E	Carter et al. (2012)	Grades 3 – 12	Self-report, oral administration	Cognitive and affective engagement, includes indirect measures e.g., student's relationship with teachers and peers
SEM- MacArthur: School Engagement Measure (SEM) - MacArthur	E	Fredricks et al. (2005)	Grade 3 - 5	Self-report, 5-point scale, oral administration, 19 items	Behavioural, cognitive and emotional engagement reported separately.

Name	Motivation/ Engagement	Author - refs	Age group	Format	Focus
SEQ: School Engagement Scale/Questionnaire	E	Dornbusch and Steinberg (1990); Perry (2008); Steinberg (1997)	Secondary	self-report, 5 or 6 point scale consisting of 12 items	assesses the degree to which students are committed to or express school behavior according to four factors: homework, classroom attention, attendance, classroom concentration
SMOT: Scientific motivation in classrooms (developed from the SMQ)	M	Schumm and Bogner (2016)	Secondary	Self-report, 25 items, 5-point scale.	Assesses motivation in science. 5 subscales
SMQ: School/ science Motivation Questionnaire	M	Marsh et al. (2003)	Secondary	44 items, oral, read-aloud presentation, 5-point scale	2 higher-order factors (Learning, performance) + 8 first-order motivation factors: Mastery, Intrinsic, Cooperative, Individual, Ego, Competition, Achieve Success, Avoid Failure

Appendix E Search Outlines

Search: Problem Based Learning and its effect on Critical and Creative Thinking				
Date	Data Base	Search Terms	Number found	Number selected
28.10.19	ProQuest (Educational database, Eric, Psychology database)	(“problem based learning” OR PBL) AND Industry AND (“critical thinking” OR “creative thinking”)	1900	
		(“problem based learning” OR PBL) AND Industry AND ab(“critical thinking” OR “creative thinking”) and subt.exact(“elementary education”)	69	1
		(“problem based learning” OR PBL) AND Industry AND (“critical thinking” OR “creative thinking”) (subject: problem solving)	88	1
1.12.19	ProQuest	ab(The Probe Method) AND ab(Education) dissertations and theses	37	1
		ab(The Probe Method) AND ab(skills)	302	1

Search: effect of motivation and engagement on learning and skills				
6.11.19	ProQuest	(student engagement and motivation) AND (effect on learning)	134,611	
		Filters: scholarly journals; article; feature; report	90	2
		(effect on learning skills)	74	2
		Ab(student engagement AND motivation AND ab(effect on learning) AND ab(elementary school)	11	1
		pedagogy	1,487	5
	Scopus	TITLE-ABS-KEY (engagement) OR TITLE-ABS-KEY (motivation) AND TITLE-ABS-KEY ("effect on learning")	223	7
		TITLE-ABS-KEY ("student engagement") OR TITLE-ABS-KEY ("student motivation") AND TITLE-ABS-KEY ("effect on learning")	34	6
27.11.19	ProQuest ERIC	ti(motivation) AND ti(assessing) AND (primary school) peer reviewed	2	0
		ab(assessing motivation) AND ab(primary school)	7	0
	ProQuest central	ti(assessing motivation) AND ab(primary school) peer reviewed	3	0
		ti(assessing motivation) AND ab(elementary school) peer reviewed	1	0
		su(assessing motivation) AND su(elementary school) AND (student motivation)	0	
28.11.19	ProQuest central	ab(stem education) AND ab(skills) and (primary school)	132	5
		ab(motivation) AND ab(skills) AND ab(primary school)	130	1

2.12.19	ProQuest central	ab(difference) AND ab(motivation) AND ab(engagement) filter: students	225 23	1
4.12.19		ab(engagement) AND ab(motivation) AND ab(creativity/)	57	3

Search: Development of skills with PBL and industry context				
8.11.19	ProQuest central	ab("acquisition of skills") OR ab("development of skills") AND entrepreneur* AND ab("elementary school") OR ab("primary School") AND ab(PBL) OR ab("Problem Based Learning") NOT nurse*	2,179	
		Article, peer reviewed	1,676	
		schools	20	2
		Skill development	41	1
		Cognition & reasoning	43	1
		Critical thinking	60	3
		Problem Based Learning	461	7

Search: Development of Critical and Creative Thinking in the Australian Curriculum, assessment of Critical and Creative Thinking skills				
13.12.19	ProQuest central	ab(Critical and creative thinking) AND ab(Australian curriculum)	12	1

From 1.1.2020, searches were carried out using the 'snowballing' technique. (Badampudi, D., Wohlin, C., & Petersen, K. (2015). Experiences from using snowballing and database searches in systematic literature studies. Proceedings of the 19th international conference on evaluation and assessment in software engineering, Nanjing, China

Search: Problem Based Learning and its effect on Critical and Creative Thinking				
Date	Data Base	Search Terms	Number found	Number selected
11.05.23	ProQuest Central After 2019	('problem based learning' OR PBL) – in abstract or summary text AND ('critical thinking' OR 'creative thinking') – in abstract or summary text NOT 'tertiary students' OR 'university students' – in abstract	27.986	
	ProQuest Central After 2019 -peer reviewed, Full text, Scholarly journals	'problem based learning' OR PBL – in abstract or summary text AND 'critical thinking' OR 'creative thinking' – in abstract or summary text AND 'elementary school' OR 'primary school' – in abstract NOT 'tertiary students' OR 'university students' OR 'medical' – anywhere	8,336	1 (Heri)

Appendix F Study One Supporting Documents

Appendix F-1. Study One. Assessment Task One Script

Hi,

I'm doing some research at UniSA about what students think about themselves as learners and creative thinkers.

Who knows what research is? *(Accept responses from students)*
Have any of you ever done any research? *(Accept responses from students)*
Sometimes when you are doing science you do research - you might come up with a question or a hypothesis, and then you do some tests or experiments to see if you're right.

That's what I'm doing. I'm trying to find out if some of the lessons you are doing are helping you to be better at thinking creatively and critically, which might be helping you to learn better. I'm doing this because I want to help teachers to find the best ways of teaching you.

I want to make sure first of all that you know what I mean by critical thinking and creative thinking, so we are going to do a short activity together.
I'm going to put you into groups of 3, and you will have a quick activity to do together, which we'll talk about when you are done, then I'm going to give you a questionnaire, which I'll explain in a minute.

Activity
Students are required to build a simple Lego figure, using a set of instructions, in less than 8 minutes. Each set has some parts missing, and some extra blocks.

Instructions
I'm going to give you a bag of Lego, and a set of instructions. I want you to work together as a team to build the thing that's in the picture. Now you aren't the first people to use this set of Lego blocks, so I hope the last people put everything back but if they didn't, just do the best you can to build the thing in the picture. Don't take anything out of the bag until I say. You are welcome to talk to each other, and you can ask me any questions – Do you have any questions now?
Identify teams of 3 or 4 and hand out Lego
OK, I'm going to time you, to see which kit can be done the fastest. Are you ready? – 3,2,1, go.
After 6 minutes stop activity.
Did anyone have to ask any questions during the activity? (establish the sort of questions asked – clarifying about the activity, more general questions about the purpose or to get more general information etc)
Were you asking questions about really complicated issues, or theories, like politics, or global warming, or anything like that? *(Accept responses from students)*
Did you realise that there are different types of questions that we sometimes ask, in different situations? It's one of the things I'll be asking you about in a minute – what sort of questions you ask.

Who had a problem they needed to solve? *(Accept responses from students)* What sort of problems did you have?
Sometimes when we solve problems, we can use something we tried before, to help us come up with a solution – did that help anyone? *(Accept responses from students – prompt if necessary "Have*

you ever had to use different Lego block because you couldn't find the one you wanted? How did you decide what to do, or which one to use – colour? Shape? Size?")
Did anyone think about what they were going to do and then have to change it because it didn't work? *(Accept responses from students)*
Sometimes these are things we have to do when we are thinking creatively to solve problems, and that's something else I'm going to ask you about in a minute.
Ok, please put the Lego back in the bag and I'll collect it from you as I hand out the sheets.

So, to find out what sort of critical thinking and creative thinking you do, I need to get you to listen to some statements from me, and then you are going to decide whether or not you do those things. Does that sound OK?

Let's start:

You each have a sheet that says Critical and Creative Thinking on the front. You can also see where there is a space to put your class and name – do that now please. (*allow time for writing).*
I need your name on there now so I can match up the one you do today, and the one your teacher is doing but when I look at your answers, I'm actually going to remove your name and replace it with a number. That means your answers will be anonymous, so its really important that you give me the answer that best fits what you really think you do – so you shouldn't give me the answer that you **wish** you could do, or that you **think you ought** to be able to do. Just the one that describes what you **actually do most of the time**.

I don't mean just what you did in the Lego activity, but what you do in other lessons at school.

If you turn your page over you'll see there is a box with **example** written in it. We're going to use that box to practice what I want you to do with the other questions on the sheet.

You can see there are 5 boxes on a line. The boxes are called box 1, box 2, box 3, box 4 and box 5.

The example is about Catching and throwing balls. I'm going to read you a statement, and if you think you do that most of the time, I want you to put a tick in box 1. If you think you don't do it, or have never tried to do it, put a cross in the box.

(Box 1): **I accurately throw balls a long way** (repeat the statement)

If you do that, put a tick in box 1. If not, put a cross. (allow about 5 seconds)

I'm going to do this example with you.

I'm OK at throwing balls, but I'm not brilliant, so when I throw a ball a long way, it definitely isn't always very accurate so I'm going to put a cross in that box.

Box 2: **I choose when I want to throw a ball overarm or underarm**. (repeat the statement)

Please put a mark in box 2. (allow about 5 seconds)

Well I can't really throw overarm, so most of the time I don't really choose. So that box gets another cross from me.

Box 3: I catch balls that are thrown to me from less than 2 metres away. (repeat the statement)

Please put a mark in box 3. (allow about 5 seconds)

We can assume that it's a good throw, and its being thrown to you, or me. I would expect to catch it every time, so I'll tick that one.

Box 4: I throw balls accurately so that someone else can catch them. (repeat the statement) (allow about 5 seconds)

I would say usually I can, if they're not too far away from me, so I'll tick that one

Box 5: **I can catch a ball that is thrown to me from a long distance**. (allow about 5 seconds).

I think a long distance is like from the other side of the oval or the cricket pitch. I'm pretty good at catching balls, but I don't always manage it, but I mostly do so I'll tick that one too.

Were there any you found difficult to decide? What would you do if you'd never tried to throw a ball a long way? (elicit answer 'put a cross')

As you can see, the statements are all about the same sort of thing, but they are not in any sort of order. Also, some statements might seem to be asking you the same thing – like the two statements asking about if you can throw a ball accurately, but your answer will give me important information, so it's important that you think about each question on its own, and answer it as truthfully as you can.

I'm now going to read you out a statement for each of the boxes on your sheet, and you will put a tick or cross in the boxes, just like we've practiced. I'll read each statement twice so you can think about it and be sure you've heard it correctly.

Do you have any questions about doing this sheet now?

Please turn over the page to the front again

The first part is about Asking Questions, so you need to think about questions you might ask in different situations.
I'll read each statement twice, so think carefully about if you do each thing most of the time.

Box 1: **I ask questions to compare information about things around me**. Do you often do that? (repeat the statement) (allow about 5 seconds)

Box 2: **I ask questions to help me learn about complex or complicated issues**. (repeat the statement) (allow about 5 seconds)

Box 3: **I ask questions about things I'm interested in**. (repeat the statement) (allow about 5 seconds)

Box 4: **I ask questions to understand why and how things happen**. (repeat the statement) (allow about 5 seconds)

Box 5: **I ask questions to understand about things I notice in the world**. (repeat the statement) (allow about 5 seconds)

Box 6: **I ask questions to understand and analyse very complicated issues and abstract ideas or theories**. (repeat the statement) (allow about 5 seconds)

How are you going? Do you have any questions? Wriggle your fingers and your shoulders.

The next part is about how you choose which bits are useful information, and how you organize information once you've found it.

Box 1: **I find and organise information by identifying and ordering it**. (repeat the statement) (allow about 5 seconds)

Box 2: **I find and organise information by collecting and sorting things that are the same or similar**. (repeat the statement) (allow about 5 seconds)

Box 3: **I find and organise information by comparing it against criteria I have developed to make sure that it is reliable**. (repeat the statement) (allow about 5 seconds)

Box 4: **I find and organise information by choosing the relevant part to use and knowing which are the facts**. (repeat the statement) (allow about 5 seconds)

Box 5: **I find and organise information by carefully selecting and analysing it against criteria I have been given**. (repeat the statement) (allow about 5 seconds)

Box 6: **I find and organise information by collecting, comparing and putting it into categories**. (repeat the statement) (allow about 5 seconds)

Part 3 is called Connecting Information. It's about whether you make connections with information you might have got in a different subject, or even from things you've done out of school. Some of you might have done this with the Lego activity.

Box 1: **I can connect information about different things**. (repeat the statement) (allow about 5 seconds)

Box 2: **I can connect information from different places, and use it to help me learn and understand something completely different**. (repeat the statement) (allow about 5 seconds)

Box 3: **I can connect information from different places, even when it is about something completely different.** (repeat the statement) (allow about 5 seconds)

Box 4: **I can connect information from different places, I plan to do it, and I can explain why it improves my ideas.** (repeat the statement) (allow about 5 seconds)

Box 5: **I can connect information I learned somewhere else, and I can explain why I choose to use it when I'm working in a different subject area**. (repeat the statement) (allow about 5 seconds)

Box 6: **I can connect information I learned from doing something else, and use it to come up with a new idea**. (repeat the statement) (allow about 5 seconds)

Please stand up demonstrate …and SNAP…
The fourth topic is called Creating New Ideas, and it's about where you get new ideas from.

Box 1: **I create new ideas by looking closely at things around me to give me ideas**. (repeat the statement) (allow about 5 seconds)

Box 2: **I create new ideas by seeing how my ideas are related to existing ideas and making them more imaginative**. (repeat the statement) (allow about 5 seconds)

Box 3: **I create new ideas by coming up with more than one idea and then choosing the best.** (repeat the statement) (allow about 5 seconds)

Box 4: **I create new ideas by combining ideas from more than one place, and in more than one way**. (repeat the statement) (allow about 5 seconds)

Box 5: **I create new ideas by coming up with really complex solutions that combine ideas and are imaginative and unusual**. (repeat the statement) (allow about 5 seconds)

Box 6: **I create new ideas by adding to things I already know**. (repeat the statement) (allow about 5 seconds)

Pat your head
Now rub your tummy at the same time.
Now swap hands. Can you still do it?

This next section is asking you to think about how you evaluate ideas that you've come up with. It's called coming up with ideas.

Box 1: **When I've come up with an idea I choose the best option after I've carefully tested a few.** (repeat the statement) (allow about 5 seconds)

Box 2: **When I've come up with an idea I test options, analyse solutions and justify choices.** (repeat the statement) (allow about 5 seconds)

Box 3: **When I've come up with an idea I compare it with other people's ideas and sometimes add some of theirs to mine to make it better**. (repeat the statement) (allow about 5 seconds)

Box 4: **When I've come up with an idea I just go with it**. (repeat the statement) (allow about 5 seconds)

Box 5: **When I've come up with an idea I can describe what the problems might be, and how I might change my idea to avoid them.** (repeat the statement) (allow about 5 seconds)

Box 6: **When I've come up with an idea I see if I can make it better by coming up with more ideas**. (repeat the statement) (allow about 5 seconds)

The last part is about being able to explain the thinking you had, when you were making your plan.

Box 1: **When I've made a plan to do something I can tell you what decisions and choices I had to make**. (repeat the statement) (allow about 5 seconds)

Box 2: **When I've made a plan to do something I can change my plan if the situation changes or if I find new information**. (repeat the statement) (allow about 5 seconds)

Box 3: **When I've made a plan to do something I can compare my plan to someone else's, and explain which will be the best plan.** (repeat the statement) (allow about 5 seconds)

Box 4: **When I've made a plan to do something I can explain the decisions and choices I had to make, and why I made them**. (repeat the statement) (allow about 5 seconds)

Box 5: **When I've made a plan to do something I can discuss how reliable my information is and tell you where there might be mistakes in my reasoning**. (repeat the statement) (allow about 5 seconds)

Box 6: **When I've made a plan to do something I can see where more information is needed, or where parts of the plan may need more thought**. (repeat the statement) (allow about 5 seconds)

Do you have any questions? Please make sure your class and name are at the top.

That's the end now. Thank you.

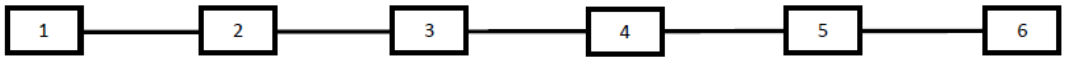
(collect sheets and ensure they are named)

Appendix F-2. Study One. Student Assessment Task One

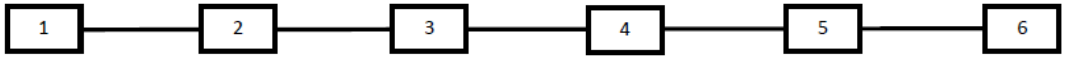
Class: _____ Name: _____

Critical and Creative Thinking

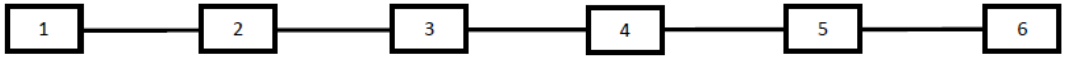
1. Asking Questions



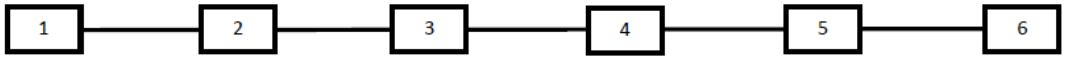
2. Finding and organising information



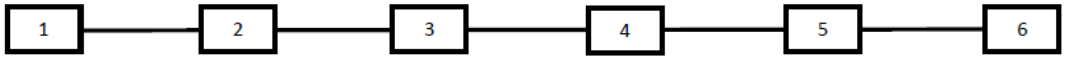
3. Connecting information



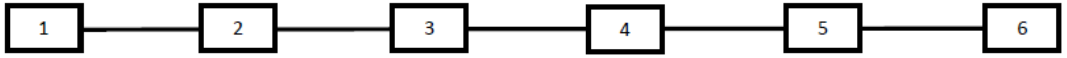
4. Creating new ideas



5. Coming up with ideas

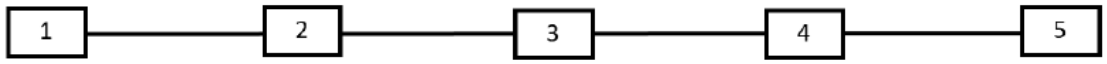


6. Thinking about a plan



Example

Catching and throwing balls



Appendix F-3. Study One. Student Assessment Task Two

Class: Name:

Critical and Creative Thinking

Example	I do this...				
	Never	Occasionally	Sometimes	Often	Always
1. I accurately throw balls a long way.					
2. I choose when I want to throw a ball overarm or underarm.					
3. I catch balls that are thrown to me from less than 2 metres away.					
4. I throw balls accurately so that someone else can catch them.					
5. I can catch a ball that is thrown to me from a long distance.					

Asking questions	I do this...				
	Never	**Occasionally**	**Sometimes**	**Often**	**Always**
1. I ask questions to compare information about things around me.					
2. I ask questions to help me learn about complex (complicated) issues.					
3. I ask questions about things I'm interested in.					
4. I ask questions to understand why and how things happen.					
5. I ask questions to understand about things I notice in the world .					
6. I ask questions to understand and analyse very complicated issues and abstract ideas (theories).					

Finding and organising information	I do this...				
	Never	**Occasionally**	**Sometimes**	**Often**	**Always**
1. I find and organise information by identifying and ordering it.					
2. I find and organise information by collecting and sorting things that are the same or similar.					
3. I find and organise information by comparing it against criteria I have developed to make sure that it is reliable.					
4. I find and organise information by choosing the relevant part to use and knowing which are the facts.					
5. I find and organise information by carefully selecting and analysing it against criteria I have been given.					
6. I find and organise information by collecting, comparing and putting it into categories.					

Connecting information	I do this...				
	Never	Occasionally	Sometimes	Often	Always
1. I can connect information about different things.					
2. I can connect information from different places, and use it to help me learn and understand something completely different.					
3. I can connect information from different places, even when it is about something completely different.					
4. I can connect information from different places, I plan to do it, and I can explain why it improves my ideas.					
5. I can connect information I learned somewhere else, and I can explain why I choose to use it when I'm working in a different subject area.					
6. I can connect information I learned from doing something else, and use it to come up with a new idea.					

Creating new ideas	I do this...				
	Never	Occasionally	Sometimes	Often	Always
1. I create new ideas by looking closely at things around me to give me ideas .					
2. I create new ideas by seeing how my ideas are related to existing ideas, and making them more imaginative.					
3. I create new ideas by coming up with more than one idea and then choosing the best.					
4. I create new ideas by combining ideas from more than one place, and in more than one way.					
5. I create new ideas by coming up with really complex solutions that combine ideas and are imaginative and unusual.					
6. I create new ideas by adding to things I already know.					

Coming up with ideas	I do this...				
	Never	**Occasionally**	**Sometimes**	**Often**	**Always**
1. When I've come up with an idea I choose the best option after I've carefully tested a few.					
2. When I've come up with an idea I test options, analyse solutions and justify choices.					
3. When I've come up with an idea I compare it with other people's ideas and sometimes add some of theirs to mine to make it better.					
4. When I've come up with an idea I just go with it.					
5. When I've come up with an idea I can describe what the problems might be, and how I might change my idea to avoid them.					
6. When I've come up with an idea I see if I can make it better by coming up with more ideas.					

Thinking about a plan	I do this...				
	Never	**Occasionally**	**Sometimes**	**Often**	**Always**
1. When I've made a plan to do something I can tell you what decisions and choices I had to make.					
2. When I've made a plan to do something I can change my plan if the situation changes or if I find new information.					
3. When I've made a plan to do something I can compare my plan to someone else's, and explain which will be the best plan.					
4. When I've made a plan to do something I can explain the decisions and choices I had to make, and why I made them.					
5. When I've made a plan to do something I can discuss how reliable my information is and tell you where there might be mistakes in my reasoning.					
6. When I've made a plan to do something I can see where more information is needed, or where parts of the plan may need more thought.					

Appendix F-4. Study One. Teacher Task

Student:

Critical and Creative Thinking

Teacher: Class:	Score: 0 = never, 1 = occasionally, 2 = sometimes, 3 = often, 4 = always Confidence (how confident you are to score this section): 1 = a little, 2 = moderately, 3 = very	**Score**	**Confidence**
I ask questions:	to compare information about things around		
	me to help me learn about complex (complicated) issues		
	about things I'm interested in		
	to understand why and how things happen		
	to understand about things I notice in the world		
	to understand and analyse very complicated issues and abstract ideas (theories)		
I find and organise information by:	identifying and ordering it		
	collecting and sorting things that are the same or similar		
	comparing it against criteria I have developed to make sure that it is reliable		
	choosing the relevant part to use and knowing which are the facts		
	carefully selecting and analysing it against criteria I have been given		
	collecting, comparing and putting it into categories		
I can connect information:	about different things		
	from different places, and use it to help me learn and understand something completely different		
	from different places, even when it is about something completely different		
	from different places, I plan to do it, and I can explain why it improves my ideas		
	I learned somewhere else, and I can explain why I choose to use it when I'm working in a different subject area		
	I learned from doing something else, and use it to come up with a new idea		

Student:	Score: 0 = never, 1 = occasionally, 2 = sometimes, 3 = often, 4 = always Confidence (how confident you are to give this score): 1 = a little, 2 = moderately, 3 = very	**Score**	**Confidence**
I create new ideas by:	looking closely at things around me to give me ideas		
	seeing how my ideas are related to existing ideas, and making them more imaginative		
	coming up with more than one idea and then choosing the best		
	combining ideas from more than one place, and in more than one way		
	coming up with complex solutions that combine ideas and are imaginative and unusual		
	adding to things I already know		
When I've come up with an idea I:	choose the best option after I've carefully tested a few		
	test options, analyse solutions and justify choices		
	compare it with other people's ideas and sometimes add some of theirs to mine to make it better		
	just go with it		
	can describe what the problems might be, and how I might change my idea to avoid them		
	see if I can make it better by coming up with more ideas		
When I've made a plan to do something I can:	tell you what decisions and choices I had to make		
	change my plan if the situation changes or if I find new information		
	compare my plan to someone else's, and explain which will be the best plan		
	explain the decisions and choices I had to make, and why I made them		
	discuss how reliable my information is and tell you where there might be mistakes in my reasoning		
	see where more information is needed, or where parts of the plan may need more thought		

Appendix F-5. Study One. Teacher Survey

I am currently doing research to investigate effective ways to develop Critical and Creative Thinking skills in Primary School students.

The first thing I need to do is to find or create effective assessment tools for assessing the Critical and Creative Thinking (C&CT) skills, as they are described in the Australian Curriculum. You have been suggested to me as someone who might be able to help me with my research.

I am looking for predominantly year 4, 5 and 6 teachers who:

- have some familiarity with the Critical and Creative Continuum from the Australian Curriculum.
- would be happy to have me guide students through a self-assessment task for 1 lesson during Term 2, 2020.
- would be happy to also complete the assessment of their pupils C&CT skills to compare with the student self-assessments.

Please complete the attached brief questionnaire as soon as possible before the end of Term 1. I will contact you by Friday 17th April to let you know the results of the survey.

Kind regards,

Sue Gaardboe

Critical and Creative Thinking

Teacher Survey

(scoring to be hidden from respondents)

1. Which year levels are you teaching this year?
   (score: 0 = year 1,2 and/or 3; 2 = year 4,5 and/or 6; 1 = year 7)

2. How many years have you been teaching your current year level?
   (score: 0 < 2 years; 1 ≥ 2 years)

3. Have you ever used the Australian Curriculum, Critical and Creative Thinking continuum to directly assess your students?
   (score: 0 = no; 1 = yes; 0 = not sure)

4. On a scale of 1 to 5, where 1 = not comfortable, 5 = very comfortable, how comfortable do you feel using the Critical and Creative Thinking continuum to assess your students?
   (score: 0 < 3; 1 ≥ 3)

5. Does your school identify as having an Inquiry Learning, IB or PBL focus in years 3–7?
   (score: 1 = yes; 0= no/ not sure)

6. Do you have other formal assessment methods you have used to assess Critical Thinking or Creative Thinking?
   (score: 0 = no; 1 = yes)

7. How many students do you have in your class this year?
   (score: 0 < 25; 1 ≥ 25)

Appendix F-6. Study One. Teacher Information

Identifying an Assessment Method for Critical and Creative Thinking

I am carrying out research for the purpose of my PhD, to find out whether using an industry partner and their genuine, messy, unsolved problem in STEM Problem-Based Learning, has any effect on helping primary school students develop critical and creative thinking skills.

For this study I am looking for teachers who have:

- experience with the Critical and Creative Thinking Continuum in the Australian Curriculum.
- a year 4, 5 or 6 class in 2020
- have taught their current year level for at least two years previously (not necessarily last year)

Study One will be testing different types of assessment for Critical and Creative Thinking Skills, to determine which is the most effective.

You have been contacted as a potential collaborator because you have either worked with me in the past or have been recommended to me by a colleague. **Your involvement in Study One will consist of enabling the researcher to take your class for one lesson and providing information about the students in your class.**

This Participant Information Sheet tells you about the research project. It explains the processes involved with taking part. Knowing what is involved will help you decide if you want to take part in the research.

Please read this information carefully. Ask questions about anything that you don't understand or want to know more about.

Participation in this research is voluntary. If you don't wish to take part, you don't have to.

If you decide you want to take part in the research project and fit the required criteria, you will be asked to sign the consent section. By signing it you are telling us that you:

- Understand what you have read
- Consent to take part in the research project
- Consent to the use of your information as described.

This Participant Information Sheet is for you to keep.

Aim of the Project

This project aims to identify an effective method for assessing primary school students Critical and Creative Thinking, as it is described in the Australian Curriculum General Capabilities. Very little research has been done on assessing these skills, particularly at a primary school level, and as you probably know, there are no standardised tests for assessing the Skills as described in the Australian Curriculum.

Study One is part of a larger project to investigate how much of an effect using an industry partner has on the development of students' skills. Study One is designed to identify a method for assessing the skills. The next step, Study Two, is to identify the effect on student's learning of using an industry partner. In fact, there is very little research done on Primary School students and their learning as a whole, so this research is important as it helps fill this gap in world-wide research.

You may be aware that the use of industry partners and their problems are currently being promoted as an effective way to teach STEM in schools. This research will help us understand more about the effect this has on student learning, and whether the investment of the time and effort it takes to engage an industry partner in our student's education is worthwhile.

Procedures and activities

1. Teachers eligible for Study One will be randomly assigned an assessment task.
2. The researcher and teacher will agree on a suitable date and time in Term 2, 2020 to run the assessment task.
3. All students will be provided with information and a parental consent form to complete prior to the assessment task being run. (Teacher to distribute and collect for researcher)
4. The researcher will present the assessment task to the students during one lesson (approximately 50 minutes). At conclusion of the lesson the researcher will collect the completed assessments for analysis.
5. The teacher will be required to also score all students in the class in line with the assessment task.
6. Standard student information (name, age, gender, year group, NEP/One Plan/ILP, year level for English instruction) will be required for each student. All student data will be made anonymous once it has been collated with the teacher assessments.

This research will compare the effectiveness of 2 assessments, each assessment being tested by a minimum of 64 students. Both tasks are self-assessments, differing in the format of the task and its presentation to the students. All testing will be carried out in primary schools in South Australia.

The benefit for you, of being involved in this research, will be to provide you with data on your students which you may find helpful in your reporting to parents for 2020. In the longer term, the benefits of this research will be in helping us develop effective methods of teaching and assessing of student skills.

It is not anticipated that there are any risks to participation in this study beyond those encountered in everyday life.

If you do consent to participate, you may withdraw at any time. Your decision to take part or not to take part (or to take part and then withdraw), will not affect your relationship with your employer, the researcher or the University of South Australia, now or in the future. Data will be excluded from participants who wish to withdraw, however it may not be possible to withdraw your data from the study results if these have already had your identifying details removed. If you decide to withdraw from the project, please notify a member of the research team before you withdraw. A member of the research team will inform you if there are any special requirements linked to withdrawing.

What will happen to the information?

By signing the consent form you consent to the research team collecting and using information from you for the research project. Any information obtained in connection with this research project that can identify you or your students will remain confidential. All student information will be de-identified immediately after it is collated. De-identified information collected or used will be stored as digital data on a secure UniSA server for 7 years (standard University of South Australia policy). Hard copies will be kept in a locked facility and destroyed on completion of the research.

The information collected will only be used for the purpose of this research project. It will only be disclosed with your written consent, except as required by law.

How do I find out about the research?

A summary of the Study One project results will be available from Sue Gaardboe susan.gaardboe@mymail.unisa.edu.au" in December 2020. Individual class results will be provided to the teachers two weeks after the data collection. A summary of the results from the complete research project is expected to be available in July 2022.

Who has reviewed the research project?

The ethical aspects of this research project have been approved by the Human Research Ethics Committee (HREC) of the University of South Australia as required by the Australian government research requirements, specified in the National Statement on Ethical Conduct in Human Research (2007 - updated 2018). This statement has been developed to protect the interests of people who agree to participate in human research studies.

Research contact person

Name	Mrs Sue Gaardboe
Position	PhD candidate
Telephone	0405149933
Email	susan.gaardboe@mymail.unisa.edu.au

If you have any complaints about any aspect of the project, the way it is being conducted or any questions about being a research participant in general, please contact:

Reviewing HREC approving this research and HREC Executive Officer details

Reviewing HREC name	University of South Australia Human Research Ethics Committee
HREC Executive Officer	Ms Vicki Allen
Telephone	+618 8302 6330
Email	humanethics@unisa.edu.au

Appendix F-7. Study One. Participant Information

Student/parent/guardian

Identifying an Assessment Method for Critical and Creative Thinking

I am carrying out research for the purpose of my PhD, to find out whether using an industry partner and their genuine, messy, unsolved problem in STEM Problem-Based Learning, has any effect on helping primary school students develop critical and creative thinking skills.

This part of the study will be testing different types of assessment for Critical and Creative Thinking Skills, to determine which is the most effective.

This Participant Information Sheet tells you about the research project. It explains the processes involved with taking part. Knowing what is involved will help you decide if you want to take part in the research.

Please read this information carefully. Ask questions about anything that you don't understand or want to know more about.

Participation in this research is voluntary. If you don't wish to take part, you don't have to.

If you decide you want to take part in the research project and fit the required criteria, you will be asked to sign the consent section. By signing it you are telling us that you:

- Understand what you have read
- Consent to take part in the research project
- Consent to the use of your information as described.

This Participant Information Sheet is for you to keep.

Aim of the Project

This project aims to identify an effective method for assessing primary school students Critical and Creative Thinking, as it is described in the Australian Curriculum General Capabilities. Very little research has been done on assessing these skills, particularly at a primary school level.

Study One is part of a larger project to investigate how much of an effect using an industry partner has on the development of students' skills. Study One is designed to identify a method for assessing the skills. The next step, Study Two, is to identify the effect on student's learning of using an industry partner. In fact, there is very little research done on Primary School students and their learning as a whole, so this research is important as it helps fill this gap in world-wide research.

Procedures and activities

7. The researcher will present the assessment task to the students during one lesson (approximately 50 minutes). At conclusion of the lesson the researcher will collect the completed assessments for analysis.
8. The teacher will be required to also score all students in the class in line with the assessment task.

This research will compare the effectiveness of 2 assessments, each assessment being tested by a minimum of 64 students. Both tasks are self-assessments, differing in the format of the task and its presentation to the students. All testing will be carried out in primary schools in South Australia.

The benefit for you, of being involved in this research, is to know that you are contributing to knowledge that helps us design better ways of assessing students in the critical and creative thinking skills. In the longer term, the benefits of this research will be in helping us develop effective methods of teaching and assessing of student skills.

It is not anticipated that there are any risks to participation in this study beyond those encountered in everyday life.

If you choose not to participate in the project you may still take part in the assessment task, but your results will not be used by the researcher. If you do consent to participate, you may withdraw at any time. Data will be excluded from participants who wish to withdraw, however it may not be possible to withdraw your data from the study results if these have already had your identifying details removed. If you decide to withdraw from the project, please notify a member of the research team before you withdraw. A member of the research team will inform you if there are any special requirements linked to withdrawing.

What will happen to the information?

By signing the consent form, you consent to the research team collecting and using information from you for the research project. Any information obtained in connection with this research project that can identify you will remain confidential. All student information will be de-identified immediately after it is collated. De-identified information collected or used will be stored as digital data on a secure UniSA server for 7 years (standard University of South Australia policy). Hard copies will be kept in a locked facility and destroyed on completion of the research.

The information collected will only be used for the purpose of this research project. It will only be disclosed with your written consent, except as required by law.

How do I find out about the research?

A summary of the Study One project results will be available from Sue Gaardboe susan.gaardboe@mymail.unisa.edu.au in December 2020. Individual class results will be provided to the teachers two weeks after the data collection. A summary of the results from the complete research project is expected to be available in July 2022.

Who has reviewed the research project?

The ethical aspects of this research project have been approved by the Human Research Ethics Committee (HREC) of the University of South Australia as required by the Australian government research requirements, specified in the National Statement on Ethical Conduct in Human Research (2007 - updated 2018). This statement has been developed to protect the interests of people who agree to participate in human research studies.

Research contact person

Name	Mrs Sue Gaardboe
Position	PhD candidate
Telephone	0405149933
Email	susan.gaardboe@mymail.unisa.edu.au

If you have any complaints about any aspect of the project, the way it is being conducted or any questions about being a research participant in general, please contact:

Reviewing HREC approving this research and HREC Executive Officer details

Reviewing HREC name	University of South Australia Human Research Ethics Committee
HREC Executive Officer	Ms Vicki Allen
Telephone	+618 8302 6330
Email	humanethics@unisa.edu.au

Teacher Consent Form

This project has been approved by the University of South Australia's Human Research Ethics Committee. If you have any ethical concerns about the project or questions about your rights as a participant please contact the Executive Officer of this Committee, Tel: +618 8302 6330; Email: humanethics@unisa.edu.au.

SECTION 1: CONTACT AND PROJECT DETAILS	
Researcher's Full Name	Susan Gaardboe
Contact Details	0405149933
Supervisor's Name *(students only)*	Prof David Cropley
Contact Details	83025344
Project Number	202831
Project Title	Study One: Assessment Evaluation

SECTION 2: PARTICIPANT CERTIFICATION

In signing this form, I confirm that:

- I have read the Participant Information Sheet, and the nature and the purpose of the research project has been explained to me. I understand and agree to take part.
- I understand the nature of my involvement in the project.
- I understand that I may not directly benefit from taking part in the project.
- I understand that I can withdraw from the project at any stage and that this will not affect my status now or in the future.
- I confirm that I am over 18 years of age.
- I understand that while information gained during the project may be published, I will not be identified and personal results will remain confidential, unless required by law.
- I understand that data collected will be stored for 7 years after completion of the project, as required by university policy for general research involving a government department. Data will be stored at UniSA, as paper copies in a locked filing cabinet in a locked office, and in digital format on a computer file on the UniSA server.

Participant's Signature	*Printed Name*	*Date*

SECTION 3: RESEARCHER CERTIFICATION

I have explained the study to the participant and consider that he/she understand what is involved.

Researcher Signature	*Printed Name*	*Date*

Student Consent Form

This project has been approved by the University of South Australia's Human Research Ethics Committee. If you have any ethical concerns about the project or questions about your rights as a participant please contact the Executive Officer of this Committee, Tel: +618 8302 6330; Email: humanethics@unisa.edu.au.

SECTION 1: CONTACT AND PROJECT DETAILS	
Researcher's Full Name	Susan Gaardboe
Contact Details	0405149933
Supervisor's Name *(students only)*	Prof David Cropley
Contact Details	83025344
Project Number	202831
Project Title	Study One: Assessment Evaluation

SECTION 2: PARTICIPANT CERTIFICATION

In signing this form, I confirm that:

- I have read the Participant Information Sheet, and the nature and the purpose of the research project has been explained to me. I understand and agree to take part.
- I understand the nature of my involvement in the project.
- I understand that I may not directly benefit from taking part in the project.
- I understand that I can withdraw from the project at any stage and that this will not affect my status now or in the future.
- I understand that while information gained during the project may be published, I will not be identified and my personal results will remain confidential, unless required by law.
- I understand that data collected will be stored for 7 years after completion of the project, as required by university policy for general research involving a government department. Data will be stored at UniSA, as paper copies in a locked filing cabinet in a locked office, and in digital format on a computer file on the UniSA server.

Participant's Signature	*Printed Name*	*Date*
Parent/ Guardian's Signature	*Printed Name*	*Date*

SECTION 3: RESEARCHER CERTIFICATION

I have explained the study to the participant and consider that he/she understand what is involved.

Researcher Signature	*Printed Name*	*Date*

Appendix G Study Two Supporting Documents

Appendix G-1. *Study Two. Critical and Creative Thinking measure Script*

Hi,

I'm doing some research at UniSA about what students think about themselves as learners and creative thinkers.

Who knows what research is? *(Accept responses from students)*
Have any of you ever done any research? *(Accept responses from students)*
Sometimes when you are doing science you do research - you might come up with a question or a hypothesis, and then you do some tests or experiments to see if you're right.

That's what I'm doing. I'm trying to find out if some of the lessons you are doing are helping you to be better at thinking creatively and critically, which might be helping you to learn better. I'm doing this because I want to help teachers to find the best ways of teaching you.

I want to make sure first of all that you know what I mean by critical thinking and creative thinking, so we are going to do a short activity together.
I'm going to put you into groups of 3, and you will have a quick activity to do together, which we'll talk about when you are done, then I'm going to give you a questionnaire, which I'll explain in a minute.

Activity
Students are required to build a simple Lego figure, using a set of instructions, in less than 8 minutes. Each set has some parts missing, and some extra blocks.

Instructions
I'm going to give you a bag of Lego, and a set of instructions. I want you to work together as a team to build the thing that's in the picture. Now you aren't the first people to use this set of Lego blocks, so I hope the last people put everything back but if they didn't, just do the best you can to build the thing in the picture. Don't take anything out of the bag until I say. You are welcome to talk to each other, and you can ask me any questions – Do you have any questions now?
Identify teams of 3 or 4 and hand out Lego
OK, I'm going to time you, to see which kit can be done the fastest. Are you ready? – 3,2,1, go.
After 6 minutes stop activity.

I want to talk to you for a minute about some of the thinking you were doing while you were doing that activity. The sort of thinking that we call critical thinking, and the sort of thinking that we call creative thinking.

Did any of you have to ask any questions during the activity? (establish the sort of questions asked – clarifying about the activity, more general questions about the purpose or to get more general information etc)

Were you asking questions about really complicated issues, or theories, like politics, or global warming, or anything like that? *(Accept responses from students)* Did any of you have to ask any questions during the activity? *(establish the sort of questions asked – clarifying about the activity, more general questions about the purpose or to get more general information etc)*
Were you asking questions about really complicated issues, or theories, like politics, or global warming, or anything like that? *(Accept responses from students)* I would be really surprised if you had been asking questions like that, because they wouldn't have been helpful in this activity. Sometimes though you might ask questions like that, because you want to know more about complex issues in the world.

Did you realise that there are different types of questions that we sometimes ask, in different situations? Asking questions is one of the critical thinking skills. It's one of the things I'll be asking you about in a minute – **When** you are asking questions, what **sort** of questions you ask.

Let's talk about a different critical thinking skill:

Who had a problem they needed to solve? *(Accept responses from students)* What sort of problems did you have?
Sometimes when we solve problems, we can use something we tried before, to help us come up with a solution – did that help anyone? *(Accept responses from students – prompt if necessary "Have you ever had to use different Lego block because you couldn't find the one you wanted? How did you decide what to do, or which one to use – colour? Shape? Size?")*
Did anyone think about what they were going to do and then have to change it because it didn't work? *(Accept responses from students)* Perhaps you didn't quite follow the instructions because you had a better idea of how to make it look better with the bricks you've got, so you did it a bit differently.

Sometimes these are things we have to do when we are thinking creatively to solve problems, and that's something else I'm going to ask you about in a minute – how you come up with creative ideas.
Ok, please put the Lego back in the bag and I'll collect it from you as I hand out the sheets.

Have you noticed that I'm reading from a script? I'm doing this activity with lots of different schools, and I have to make sure I give the same information to every class, so that's why I'm reading it!

So, to find out what sort of critical thinking and creative thinking you do, I need to get you to listen to some statements from me, and then you are going to decide whether or not you do those things. Does that sound OK?

Let's start:

You each have a sheet that says Critical and Creative Thinking and has a Lego picture on the front. You can also see where there is a space to put your class and name – do that now please. (*allow time for writing).*
I need your name on there now so I can match up the one you do today, and the one your teacher is doing but when I look at your answers, I'm actually going to remove your name and replace it with a number. That means your answers will be anonymous, so its really important that you give me the answer that best fits what you really think you do – so you shouldn't give me the answer that you **wish** you could do, or that you **think you ought** to be able to do. Just the one that describes what you **actually do most of the time**.

I don't mean just what you did in the Lego activity, but what you do at other times too.

If you turn your page over you'll see there is a box with **example** written in it. We're going to use that box to practice what I want you to do with the other questions on the sheet.

You can see there are 6 boxes on a line. The boxes are called box 1, box 2, box 3, 4, 5, and 6.

The example is about Catching and throwing balls. I'm going to read you a statement, and if you think you do that most of the time, I want you to put a tick in box 1. If you think you don't do it, or have never tried to do it, put a cross in the box.

(Box 1): **I accurately throw balls a long way** (repeat the statement)

If you do that, put a tick in box 1. If not, put a cross. (allow about 5 seconds)

I just want to point out the wording in the statements to you. The first one is: I accurately throw balls a long way. It doesn't say I can accurately throw balls a long way if I want to, or I think I can accurately throw balls a long way if I try. It very clearly makes the statement "I do this". You may never have thrown a ball a long way because you never had the opportunity. Even if you're the strongest kid in the school and you know you could do it if you tried – if you haven't tried, you NEED to put a cross as your answer, because its not something that you do. If you've tried to do it lots of times, and you once managed to throw it straight to the person you were aiming for, but mostly you aren't that accurate, then you also need to put a cross, because mostly when you're throwing balls a long way you're not accurate. If you can do it most of the time when you try, then you can put a tick.

Also, If you don't understand the question, it's important that you ask me to explain it. If you don't understand the question, there will probably be others in the class who don't understand it either, so just put your hand up and I'll try to explain it more.

Do you have any questions now?

I'm going to do this example with you.

I haven't tried to throw a ball a long way since I was in high school, so I'm going to put a cross in that box.

Box 2: **I choose when I want to throw a ball overarm or underarm**. (repeat the statement)

Please put a mark in box 2. (allow about 5 seconds)

I don't really know if I choose. I can do both, but I don't know if I make a choice, so I'll mark another cross.

Box 3: **I catch balls that are thrown to me from less than 2 metres away**. (repeat the statement)

Please put a mark in box 3. (allow about 5 seconds)

We can assume that it's a good throw, and its being thrown to you. I would expect to catch it every time, so I'll tick that one.

Box 4: **I catch balls thrown to me from a long distance when I wear a catcher's mitt.** (allow about 5 seconds)

Do you know what a catcher's mitt is? I've never worn one, so I'll put a cross.

Box 5: **I throw balls accurately across a short or medium distance so that someone else can catch them**

Lastly, Box 6: **I catch balls that are thrown to me from a long distance.** A long distance is like from the other side of the oval or the other end of the netball court. Please put make your mark (allow about 5 seconds) I'll put a cross for that one too. (allow about 5 seconds).

Were there any you found difficult to decide? What would you do if you'd never tried to throw a ball a long way? (elicit answer 'put a cross')

As you can see, the statements are all about the same sort of thing, but they are not in any sort of order. Also, some statements might seem to be asking you the same thing – like the two statements asking about if you can throw a ball accurately, but your answer will give me important information, so it's important that you think about each question on its own, and answer it as truthfully as you can.

I'm now going to read you out a statement for each of the boxes on your sheet, and you will put a tick or cross in the boxes, just like we've practiced. I'll read each statement twice so you can think about it and be sure you've heard it correctly.

Do you have any questions about doing this sheet now?

Please turn over the page to the front again

The first part is about Asking Questions, so you need to think about questions you might ask in different situations.
I'll read each statement twice, so think carefully about if you do each thing most of the time.

Box 1: **I ask questions to compare information about I already know something about.** Do you do that? (repeat the statement) (allow about 5 seconds)

Box 2: **I ask questions to help me learn about complex or complicated issues**. (repeat the statement) (allow about 5 seconds)

Box 3: **I ask questions about things I'm interested in**. (repeat the statement) (allow about 5 seconds)

Box 4: **I ask questions to understand more about things I notice in the wider world** (repeat the statement) (allow about 5 seconds)

Box 5: **I ask questions to understand why and how things happen.** (repeat the statement) (allow about 5 seconds)

Box 6: **I ask questions to understand and analyse very complicated issues and abstract ideas or theories**. (repeat the statement) (allow about 5 seconds)

How are you going? Do you have any questions? Wriggle your fingers and your shoulders.

The next part is called **Finding and Organising Information.** What have you had to find information about recently? *(accept responses)* (looking at web pages? In a book) Did you have to do something with the information? (find the facts, sort it under topic headings?)

I want you to think about about if you do some of the following things, WHEN you find and organize information.

Box 1: **When I find and organise information I identify the parts I need and sort them in order of importance** (repeat the statement) (allow about 5 seconds)

Box 2: **When I find and organise information I collect and sort things that are the same or similar**. (repeat the statement) (allow about 5 seconds)

Box 3: **When I find and organise information I compare it against criteria I have developed to make sure that it is reliable**. (repeat the statement) (allow about 5 seconds)

Box 4: **When I find and organise information I use more than one source and choose the relevant parts to use.** (repeat the statement) (allow about 5 seconds)

Box 5: **When I find and organise information I carefully select and analyse it against criteria I have been given**. (repeat the statement) (allow about 5 seconds)

Box 6: **When I find and organise information I put it into categories, like 'facts' and 'opinions'.** (repeat the statement) (allow about 5 seconds)

Part 3 is called Connecting Information. It's about whether you make connections with information you might have got in a different subject, or even from things you've done out of school. Some of you might have done this with the Lego activity. , remembering when

you'd had a similar problem at home. Or have you ever used something you learned in one lesson, to help you do something or understand something in another lesson? (Accept answers. Prompt: Perhaps you realized something you learned in maths would help you to understand something in science?...can you think of a time when something like that has happened?)

Box 1: **I make connections between information I learned in different places (for example at home and at school).** (repeat the statement) (allow about 5 seconds)

Box 2: **I make connections between information I learned at different times (for example lessons last year or last week, and a lesson today).** (repeat the statement) (allow about 5 seconds)

Box 3: **I make connections between information I learned in different places or times and use this to help me understand things.** (repeat the statement) (allow about 5 seconds)

Box 4: **I make connections between information I learned in different subjects, places and times. I can explain the connections and why they improve my ideas.** (repeat the statement) (allow about 5 seconds)

Box 5: **I make connections between information I learned in different subjects. I can explain the connections and why they are useful.** (repeat the statement) (allow about 5 seconds)

Box 6: **I make connections between information I learned in different subjects, and use this to help me understand things in new ways.** (repeat the statement) (allow about 5 seconds)

(more than half way – 15 mins remaining) We've got time for a little game - Please stand up (demonstrate …and SNAP…)
The fourth topic is called Creating New Ideas, and it's about where you get new ideas from.

Box 1: **I create new ideas by copying things I'm familiar with.** (repeat the statement) (allow about 5 seconds)

Box 2: **I create new ideas by seeing how my ideas are related to existing ideas and changing them to make them more imaginative**. (repeat the statement) (allow about 5 seconds)

Box 3: **I create new ideas by coming up with more than one idea and then choosing the best.** (repeat the statement) (allow about 5 seconds)

Box 4: **I create new ideas by combining ideas from different places, and using them in new and different ways.** (repeat the statement) (allow about 5 seconds)

Box 5: **I create new ideas by coming up with complex solutions that combine ideas and are imaginative and unusual**. (repeat the statement) (allow about 5 seconds)

Box 6: **I create new ideas by adding to things I already know**. (repeat the statement) (allow about 5 seconds)

Pat your head
Now rub your tummy at the same time.
Now swap hands. Can you still do it?

This is the last section and it asks you to think about how you evaluate ideas that you've come up with. It's called evaluating a plan or idea. Evaluating is about thinking if something is good or works well. What makes an idea good? Do you ever get asked to explain why your idea is a good one? (accept responses and prompt as necessary) That's what these questions are about.

Box 1: **When I come up with a plan or an idea I am able to tell you what decisions and choices I had to make.** (repeat the statement) (allow about 5 seconds)

Box 2: **When I come up with a plan or an idea I am able to change my plan if the situation changes or if I find new information.** (repeat the statement) (allow about 5 seconds)

Box 3: **When I come up with a plan or an idea I am able to compare my plan to someone else's, and explain which will be the best plan.** (repeat the statement) (allow about 5 seconds)

Box 4: **When I come up with a plan or an idea I am able to explain the decisions and choices I had to make, and why I made them.** (repeat the statement) (allow about 5 seconds)

Box 5: **When I come up with a plan or an idea I am able to discuss how reliable my information is and tell you where there might be mistakes in my reasoning.** (repeat the statement) (allow about 5 seconds)

Box 6: **When I come up with a plan or an idea I am able to see where more information is needed, or where parts of the plan may need more thought.** (repeat the statement) (allow about 5 seconds)

Do you have any questions? Please make sure your class and name are at the top.

That's the end now. Thank you.

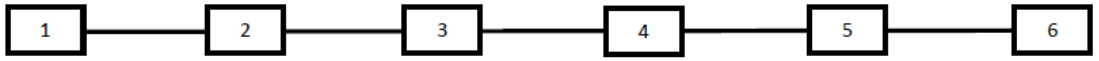
(collect sheets and ensure they are named)

Appendix G-2. *Study Two. Student Assessment. Presentation Format One*

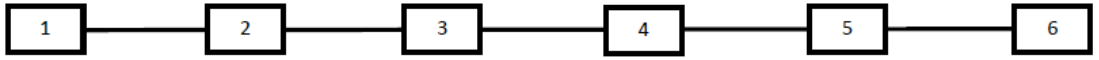
Class: _____ Name: _____

Critical and Creative Thinking

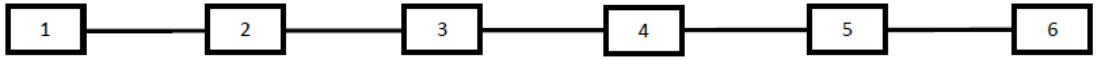
1. Asking Questions



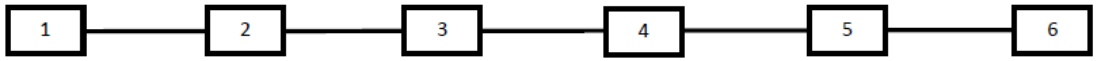
2. Finding and organising information



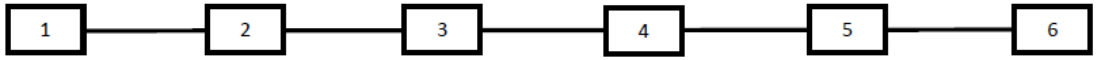
3. Connecting information



4. Creating new ideas



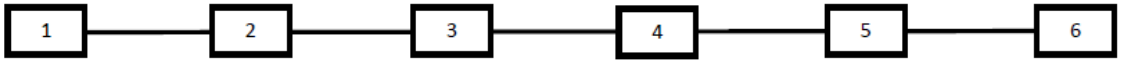
5. Evaluating a plan or idea



Yes ✓ No X Unsure if I do 1

Example

Catching and throwing balls



Yes ✓ No X Unsure if I do 1

Appendix G-3. *Study Two. Student Assessment. Presentation Format Two*

Name:

Critical and Creative Thinking

Example	Yes	No	Unsure if I do
1. I accurately throw balls a long way.			
2. I choose when I want to throw a ball overarm or underarm.			
3. I catch balls that are thrown to me from less than 2 metres away.			
4. I catch balls thrown to me from a long distance when I wear a catcher's mitt.			
5. I throw balls accurately across a short or medium distance so that someone else can catch them.			
6. I catch balls that are thrown to me from a long distance.			



https://www.parentingscience.com/toy-blocks.html CC BY-ND 2.0

Asking questions

	Yes	No	Unsure if I do
1. I ask questions to compare information about things I already know something about.			
2. I ask questions to help me learn about complex or complicated issues.			
3. I ask questions about things I'm interested in.			
4. I ask questions to understand more about things I notice in the wider world.			
5. I ask questions to understand why and how things happen.			
6. I ask questions to understand and analyse very complicated issues and abstract ideas or theories.			

Finding and organising information

	Yes	No	Unsure if I do
1. When I find and organise information I identify the parts I need and sort them in order of importance.			
2. When I find and organise information I collect and sort things that are the same or similar.			
3. When I find and organise information I compare it against criteria I have developed to make sure that the information is reliable.			
4. When I find and organise information I use more than one source and choose the relevant parts to use.			
5. When I find and organise information I carefully select and analyse it against criteria I have been given.			
6. When I find and organise information I put it into categories, like 'facts' and 'opinions'.			

Connecting information

	Yes	No	Unsure if I do
1. I make connections between information I learned in different places (e.g. at home and at school).			
2. I make connections between information I learned at different times (e.g. lessons last year or last week, and a lesson today).			
3. I make connections between information I learned in different places or times and use this to help me understand things.			
4. I make connections between information I learned in different subjects, places and times. I can explain the connections and why they improve my ideas.			
5. I make connections between information I learned in different subjects. I can explain the connections and why they are useful.			
6. I make connections between information I learned in different subjects, and use this to help me understand things in new ways.			

Creating new ideas

	Yes	No	Unsure if I do
1. I create new ideas by copying things I'm familiar with.			
2. I create new ideas by seeing how my ideas are related to existing ideas, and changing them to make them more imaginative.			
3. I create new ideas by coming up with more than one idea and then choosing the best.			
4. I create new ideas by combining ideas from different places, and using them in new and different ways.			
5. I create new ideas by coming up with complex solutions that combine ideas and are imaginative and unusual.			
6. I create new ideas by adding to things I already know.			

Evaluating a plan or idea	Yes	No	Unsure if I do
1. When I come up with a plan or an idea I am able to tell you what decisions and choices I had to make.			
2. When I come up with a plan or an idea I am able to change my plan if the situation changes or if I find new information.			
3. When I come up with a plan or an idea I am able to compare my plan to someone else's, and explain which will be the best plan.			
4. When I come up with a plan or an idea I am able to explain the decisions and choices I had to make, and why I made them.			
5. When I come up with a plan or an idea I am able to discuss how reliable my information is and tell you where there might be mistakes in my reasoning.			
6. When I come up with a plan or an idea I am able to see where more information is needed, or where parts of the plan may need more thought.			



https://www.flickr.com/photos/wwworks/4759535950 CC BY 2.0

Appendix G-4. *Study Two. Teacher Task*

Teacher (initials): **Critical and Creative Thinking** Student:

Asking questions	Score: (When the student applies this skill, how often does the student demonstrate this level of skill?) 0 = never, 1 = sometimes, 2 = always Confidence (How confident are you to score this skill for this student?): L = a little, M = moderately, V = very	**Score**	**Confidence**
1. I ask questions to compare information about things I already know something about.			
2. I ask questions to help me learn about complex or complicated issues.			
3. I ask questions about things I'm interested in.			
4. I ask questions to understand more about things I notice in the wider world.			
5. I ask questions to understand why and how things happen.			
6. I ask questions to understand and analyse very complicated issues and abstract ideas or theories.			

Finding and organising information	Score: (When the student applies this skill, how often does the student demonstrate this level of skill?) 0 = never, 1 = sometimes, 2 = always Confidence (How confident are you to score this skill for this student?): L = a little, M = moderately, V = very	**Score**	**Confidence**
1. When I find and organise information I identify the parts I need and sort them in order of importance.			
2. When I find and organise information I collect and sort things that are the same or similar.			
3. When I find and organise information I compare it against criteria I have developed to make sure that the information is reliable.			
4. When I find and organise information I use more than one source and choose the relevant parts to use.			
5. When I find and organise information I carefully select and analyse it against criteria I have been given.			
6. When I find and organise information I put it into categories, like 'facts' and 'opinions'.			

Student:

Connecting information	Score: (When the student applies this skill, how often does the student demonstrate this level of skill?) 0 = never, 1 = sometimes, 2 = always Confidence (How confident are you to score this skill for this student?): L = a little, M = moderately, V = very	**Score**	**Confidence**
1. I make connections between information I learned in different places (e.g. at home and at school).			
2. I make connections between information I learned at different times (e.g. lessons last year or last week, and a lesson today).			
3. I make connections between information I learned in different places or times and use this to help me understand things.			
4. I make connections between information I learned in different subjects, places and times. I can explain the connections and why they improve my ideas.			
5. I make connections between information I learned in different subjects. I can explain the connections and why they are useful.			
6. I make connections between information I learned in different subjects, and use this to help me understand things in new ways.			

Creating new ideas	Score: (When the student applies this skill, how often does the student demonstrate this level of skill?) 0 = never, 1 = sometimes, 2 = always Confidence (How confident are you to score this skill for this student?): L = a little, M = moderately, V = very	**Score**	**Confidence**
1. I create new ideas by copying things I'm familiar with.			
2. I create new ideas by seeing how my ideas are related to existing ideas, and changing them to make them more imaginative.			
3. I create new ideas by coming up with more than one idea and then choosing the best.			
4. I create new ideas by combining ideas from different places, and using them in new and different ways.			
5. I create new ideas by coming up with complex solutions that combine ideas and are imaginative and unusual.			
6. I create new ideas by adding to things I already know.			

Student:

Evaluating a plan or idea	Score: (When the student applies this skill, how often does the student demonstrate this level of skill?) 0 = never, 1 = sometimes, 2 = always Confidence (How confident are you to score this skill for this student?): L = a little, M = moderately, V = very	**Score**	**Confidence**
1. When I come up with a plan or an idea I am able to tell you what decisions and choices I had to make.			
2. When I come up with a plan or an idea I am able to change my plan if the situation changes or if I find new information.			
3. When I come up with a plan or an idea I am able to compare my plan to someone else's, and explain which will be the best plan.			
4. When I come up with a plan or an idea I am able to explain the decisions and choices I had to make, and why I made them.			
5. When I come up with a plan or an idea I am able to discuss how reliable my information is and tell you where there might be mistakes in my reasoning.			
6. When I come up with a plan or an idea I am able to see where more information is needed, or where parts of the plan may need more thought.			

Appendix H Study Two Tests of Normality

Table H-1. *Study Two. Tests of normality for critical and creative thinking presentation format - student responses*

Sub-element	Year level statement	Presentation Format 1 *n* = 184					Presentation Format 2 *n* = 223				
		Missing cases	Skewness	Std. Error	Kurtosis	Std. Error	Missing cases	Skewness	Std. Error	Kurtosis	Std. Error
Asking Questions	F	1	-3.990	.180	15.278	.357	0	-3.468	.163	11.535	.324
	Y2	0	-0.470	.179	-1.147	.356	1	-0.536	.163	-0.757	.325
	Y4	0	-2.210	.179	3.416	.356	1	-2.093	.163	3.558	.325
	Y6	0	-1.189	.179	-0.001	.356	1	-0.901	.163	-0.490	.325
	Y8	0	-2.037	.179	2.947	.356	0	-2.142	.163	3.653	.324
	Y10	0	-0.432	.179	-1.148	.356	2	-0.490	.164	-1.030	.326
Finding and Organising Information	F	0	-0.957	.179	-0.632	.356	1	-0.881	.163	-0.657	.325
	Y2	0	-0.216	.179	-1.500	.356	0	-0.277	.163	-1.262	.324
	Y4	1	-0.443	.180	-1.437	.357	1	-0.220	.163	-1.639	.325
	Y6	0	-1.741	.179	1.522	.356	2	-1.508	.164	0.955	.326
	Y8	0	-0.542	.179	-1.228	.356	2	-0.504	.164	-0.851	.326
	Y10	0	-0.601	.179	-1.155	.356	1	-0.691	.163	-0.957	.325

Table H-1. (continued)

Sub-element	Year level statement	Presentation Format 1 *n* = 184					Presentation Format 2 *n* = 223				
		Missing cases	Skewness	Std. Error	Kurtosis	Std. Error	Missing cases	Skewness	Std. Error	Kurtosis	Std. Error
Connecting Information	F	1	-2.327	.180	4.133	.357	0	-2.392	.163	5.085	.324
	Y2	3	-0.756	.181	-1.032	.359	0	-1.022	.163	-0.132	.324
	Y4	0	-1.339	.179	0.320	.356	1	-1.513	.163	1.081	.325
	Y6	0	-1.839	.179	2.160	.356	0	-2.062	.163	3.424	.324
	Y8	0	-0.504	.179	-1.116	.356	0	-0.567	.163	-0.916	.324
	Y10	0	-0.464	.179	-1.247	.356	0	-0.337	.163	-1.135	.324
Creating New Ideas	F	0	-1.220	.179	0.095	.356	0	-1.237	.163	0.458	.324
	Y2	0	-1.509	.179	0.898	.356	2	-1.786	.164	2.193	.326
	Y4	0	-1.483	.179	0.746	.356	0	-0.971	.163	-0.644	.324
	Y6	1	-0.885	.180	-0.678	.357	0	-0.516	.163	-1.105	.324
	Y8	0	-0.733	.179	-1.014	.356	0	-0.729	.163	-0.747	.324
	Y10	0	-0.083	.179	-1.272	.356	0	0.000	.163	-1.520	.324
Evaluating a Plan or Idea	F	0	-0.801	.179	-0.849	.356	0	-0.839	.163	-0.770	.324
	Y2	0	-0.827	.179	-0.742	.356	1	-0.525	.163	-1.304	.325
	Y4	0	-0.919	.179	-0.737	.356	0	-0.894	.163	-0.769	.324
	Y6	0	-2.011	.179	2.917	.356	0	-1.433	.163	0.896	.324
	Y8	1	-1.143	.180	0.005	.357	1	-1.029	.163	-0.173	.325
	Y10	0	-0.504	.179	-1.116	.356	0	-0.412	.163	-1.280	.324

Table H-2. *Study Two. Tests of normality for critical and creative thinking presentation format - teacher responses*

Sub-element	Year level statement	Presentation Format 1 *n* = 184					Presentation Format 2 *n* = 223				
		Missing cases	Skewness	Std. Error	Kurtosis	Std. Error	Missing cases	Skewness	Std. Error	Kurtosis	Std. Error
Asking Questions	F	0	-2.462	.179	4.834	.356	0	-2.557	.163	5.066	.324
	Y2	0	-2.104	.179	2.938	.356	0	-1.685	.163	1.242	.325
	Y4	0	-1.164	.179	-0.376	.356	0	-1.323	.163	0.248	.324
	Y6	0	-0.965	.179	-0.749	.356	0	-1.431	.163	0.592	.324
	Y8	0	-0.620	.179	-1.378	.356	0	-1.179	.163	-0.157	.324
	Y10	0	-0.117	.179	-1.711	.356	0	-0.240	.163	-1.524	.324
Finding and Organising Information	F	0	-1.707	.179	1.718	.356	0	-1.649	.163	1.757	.324
	Y2	0	-1.200	.179	0.157	.356	0	-1.221	.163	0.171	.324
	Y4	0	-0.794	.179	-0.723	.356	0	-0.631	.163	-1.014	.324
	Y6	0	-0.569	.179	-1.146	.356	0	-0.585	.163	-1.343	.324
	Y8	0	-0.435	.179	-1.298	.356	0	-0.637	.163	-1.032	.324
	Y10	0	-0.289	.179	-1.344	.356	0	-0.378	.163	-1.394	.324

Table H-2. (continued)

Sub-element	Year level statement	Presentation Format 1 n = 184					Presentation Format 2 n = 223				
		Missing cases	Skewness	Std. Error	Kurtosis	Std. Error	Missing cases	Skewness	Std. Error	Kurtosis	Std. Error
Connecting Information	F	0	-2.665	.179	6.744	.356	0	-2.159	.163	3.957	.324
	Y2	0	-0.873	.179	-0.576	.356	0	-0.816	.163	-0.860	.324
	Y4	0	-1.625	.179	1.777	.356	0	-1.632	.163	1.393	.324
	Y6	0	-2.999	.179	8.792	.356	0	-1.996	.163	3.043	.324
	Y8	0	-0.665	.179	-0.914	.356	0	-0.499	.163	-1.129	.324
	Y10	0	-0.582	.179	-1.140	.356	0	-0.606	.163	-0.942	.324
Creating New Ideas	F	0	-4.328	.179	19.826	.356	0	-1.741	.163	2.195	.324
	Y2	0	-1.929	.179	2.417	.356	0	-1.170	.163	0.366	.324
	Y4	0	-1.013	.179	-0.196	.356	0	-0.657	.163	-0.847	.324
	Y6	0	-0.767	.179	-0.612	.356	0	-0.711	.163	-0.695	.324
	Y8	0	-1.472	.179	1.091	.356	0	-0.923	.163	-0.291	.324
	Y10	0	-0.020	.179	-1.473	.356	0	-0.105	.163	-1.019	.324
Evaluating a Plan or Idea	F	0	-2.462	.179	4.834	.356	0	-1.567	.163	1.239	.324
	Y2	0	-1.109	.179	-0.154	.356	0	-1.184	.163	0.260	.324
	Y4	0	-0.764	.179	-0.807	.356	0	-0.954	.163	-0.291	.324
	Y6	0	-1.061	.179	-0.474	.356	0	-1.197	.163	0.092	.324
	Y8	0	-0.308	.179	-1.424	.356	0	-0.380	.163	-1.337	.324
	Y10	0	0.050	.179	-1.457	.356	0	-0.062	.163	-1.333	.324

Appendix I Study Two Fleiss' Kappa Analysis

Table I-1. *Study Two. Results of preliminary analysis with Fleiss' Kappa comparing teacher and student responses*

Sub-element	Year level statement	Presentation Format 1		Presentation Format 2	
		K	*p*	*K*	*p*
Asking Questions	F	.019	.737	-.029	.575
	Y2	-.094	.089	-.023	.653
	Y4	-.012	.828	.009	.859
	Y6	.078	.154	.045	.383
	Y8	.001	.989	.035	.502
	Y10	.035	.503	.025	.599
Finding and Organising Information	F	-.083	.136	.087	.099
	Y2	.060	.267	.055	.270
	Y4	.045	.401	-.034	.486
	Y6	.026	.634	.060	.230
	Y8	.150	.005*	.046	.358
	Y10	.159	.003*	.041	.395
Connecting Information	F	.019	.755	.165	.004*
	Y2	.024	.668	.103	.042*
	Y4	.042	.474	.201	<.001*
	Y6	.020	.735	.161	.004*
	Y8	-.001	.979	.045	.366
	Y10	.119	.027*	.084	.088

Note. * $p < .050$

Table I-1. (continued)

Sub-element	Year level statement	Presentation Format 1		Presentation Format 2	
		K	*p*	*K*	*p*
Creating New Ideas	F	-.138	.018*	-.034	.560
	Y2	.043	.448	.023	.691
	Y4	-.018	.751	.017	.737
	Y6	.060	.285	.130	.010*
	Y8	.043	.439	.057	.274
	Y10	.065	.215	-.024	.608
Evaluating a Plan or Idea	F	-.001	.993	-.003	.946
	Y2	-.072	.197	.070	.164
	Y4	.035	.521	.070	.163
	Y6	.002	.968	.059	.267
	Y8	.095	.080	.055	.267
	Y10	.076	.148	.082	.084
	n =	183		224	

Note. * $p < .050$

Table I-2. *Study Two. Fleiss' Kappa results for individual teachers*

Teacher ID		1	2	3	4	5	6	7	8
n =		23	18	12	24	21	6	11	10
Presentation Format		2	2	1	2	1	2	2	2
Sub-element	Year level statement								
Asking Questions	F	-.045	.063	-.143	-.121	-.099	-.200	-.048	-.111
	Y2	-.158	-.185	-.274	-.097	.094	-.463	-.148	-.268
	Y4	-.103	-.103	-.062	-.155	.043	-.029	.254	-.235
	Y6	-.181	-.220	.273	-.027	-.050	.077	.120	-.266
	Y8	-.148	-.091	-.026	-.188	.125	.333	-.100	.375
	Y10	-.251	-.050	-.067	-.277	.152	-.067	.371	-.655*
Finding and Organising Information	F	.068	-.268	-.371	-.326	-.273	-.200	-.294	-.268
	Y2	.121	-.003	-.073	.303*	-.013	-.371	.022	-.235
	Y4	-.164	-.267	-.158	-.258	-.282	.467	-.048	-.516*
	Y6	-.383*	-.008	-.149	-.202	-.413*	-.500	.033	-.220
	Y8	-.016	-.031	.018	-.244	.125	.122	-.209	-.200
	Y10	-.261	.030	-.016	-.081	-.014	.077	.036	-.350
Connecting Information	F	-.241	-.029	.150	.365*	-.084	-.200	-.100	-.053
	Y2	.111	-.115	-.189	-.126	-.099	.250	-.128	.241
	Y4	.075	.390*	-.008	-.006	-.109	-.091	-.119	-.053
	Y6	.330	-.068	.273	-.012	-.167	-.091	-.119	-.053
	Y8	-.079	-.248	.018	-.269	-.079	-.067	-.269	-.468
	Y10	.404*	-.282	.061	-.145	-.048	.234	-.124	-.266

Note. # indicates the Kappa was not calculated due to perfect agreement between the judges. This is reported as having a significance of <.010 and a Kappa at or near 1.000 (almost perfect). [a]Some of the students in this class completed Presentation Format 1, some completed Presentation Format 2.
*$p \leq .050$.

Table I-2. (continued)

Teacher ID		9	10	11	12	13	14	15	16
n =		11	21	22	15	21	17	20	22
Presentation Format		2	2	1	[a]	2	2	1	1
Sub-element	Year level statement								
Asking Questions	F	-.119	-.105	-.285	-.034	-.024	-.141	-.026	.125
	Y2	-.058	-.170	.101	.012	-.200	-.093	-.087	.267
	Y4	.120	.311	-.303	-.228	-.077	.128	.337	-.006
	Y6	-.119	-.183	-.148	.044	.412*	.018	.064	.220
	Y8	.120	-.021	-.350	.302	.294	-.042	-.088	.391*
	Y10	-.269	-.446*	.090	.108	-.048	.003	-.018	.062
Finding and Organising Information	F	-.048	-.167	.110	-.145	.135	.157	.043	-.077
	Y2	-.285	-.238	.160	-.004	-.198	.026	.064	.294*
	Y4	-.124	-.485*	-.226	-.027	-.077	-.071	.093	.089
	Y6	.060	-.059	-.041	.439*	-.099	-.197	-.105	.297
	Y8	.033	-.385*	.104	.183	-.268	.013	-.019	.012
	Y10	-.048	-.310	.339*	.163	-.119	.168	.044	.141
Connecting Information	F	-.048	#*	-.106	-.034	.063	.101	-.053	.141
	Y2	-.039	-.215	-.050	-.214	-.019	-.328	-.065	.048
	Y4	-.048	-.077	.043	-.163	.417*	.122	-.111	.309*
	Y6	.154	-.024	-.035	-.111	-.111	.124	.134	.006
	Y8	-.128	.015	.036	-.292	-.184	-.149	-.337	.105
	Y10	-.248	-.112	.022	-.315	-.385*	.380*	.164	.244

Note. # indicates the Kappa was not calculated due to perfect agreement between the judges. This is reported as having a significance of <.010 and a Kappa at or near 1.000 (almost perfect). [a]Some of the students in this class completed Presentation Format 1, some completed Presentation Format 2.
*$p \leq .050$.

Table I-2. (continued)

Teacher ID		17	18	19	20	21	22	23	24
n =		15	17	22	17	11	6	22	23
Presentation Format		2	1	1	2	1	1	2	1
Sub-element	Year level statement								
Asking Questions	F	-.163	-.046	-.023	-.030	#*	#*	-.222	.193
	Y2	-.213	-.331	-.520*	-.172	-.189	-.333	.317*	-.204
	Y4	-.013	.315	-.137	-.030	-.222	-.200	-.211	-.074
	Y6	.196	-.105	-.238	-.236	-.089	.586	.053	.164
	Y8	-.195	-.204	-.093	-.097	.076	-.143	-.081	-.220
	Y10	.080	-.360*	.245	.079	-.107	-.364	-.077	-.211
Finding and Organising Information	F	-.271	-.236	-.069	-.046	.154	-.463	.318*	-.079
	Y2	.059	-.413*	.005	-.380	.137	-.171	-.054	-.068
	Y4	-.434*	-.496*	-.148	-.430*	.038	.182	-.095	-.195
	Y6	-.180	-.105	-.092	.325	.254	-.333	.013	-.145
	Y8	-.172	.029	-.185	-.133	.083	.385	.276	.319
	Y10	-.190	.023	.121	-.184	.296	-.029	.132	-.159
Connecting Information	F	-.163	-.062	-.056	-.062	-.307	-.091	.302	-.193
	Y2	.271	-.177	-.217	.217	.518*	-.241	.112	-.143
	Y4	.227	-.259	-.137	-.185	.224	-.333	.139	-.225
	Y6	.441	-.067	-.056	-.062	-.119	-.091	-.002	-.223
	Y8	.451*	-.149	-.130	-.118	.545*	-.333	-.107	-.211
	Y10	.078	-.137	-.060	.139	.542*	.077	-.049	.170

Note. # indicates the Kappa was not calculated due to perfect agreement between the judges. This is reported as having a significance of <.010 and a Kappa at or near 1.000 (almost perfect). [a]Some of the students in this class completed Presentation Format 1, some completed Presentation Format 2.
*$p \leq .050$

Table I-2. (continued)

Teacher ID		1	2	3	4	5	6	7	8
n =		23	18	12	24	21	6	11	10
Presentation Format		2	2	1	2	1	2	2	2
Sub-element	Variable								
Creating New Ideas	F	-.127	-.111	-.241	-.002	-.215	-.091	-.158	-.266
	Y2	-.094	-.062	-.157	-.231	.074	-.029	-.048	-.111
	Y4	.055	-.210	.106	-.098	-.403*	-.333	-.124	-.235
	Y6	-.054	-.448*	.381	.027	-.283	.077	.132	-.266
	Y8	-.094	.122	-.006	.106	-.040	.077	-.479	-.266
	Y10	-.183	-.231	-.143	-.202	-.183	-.636*	-.039	-.176
Evaluating a Plan or Idea	F	-.278	.017	.216	.206	.026	-.200	.033	-.194
	Y2	.084	-.313	-.231	.078	-.290	.314	.009	-.194
	Y4	.065	-.143	.195	-.267	-.242	.077	.413	.104
	Y6	-.057	-.003	.262	.072	-.322	.122	-.119	-.333
	Y8	.084	.198	.022	-.269	.221	.314	-.236	-.429
	Y10	.203	-.268	.445*	-.067	.222	.122	-.039	-.194

Note. # indicates the Kappa was not calculated due to perfect agreement between the judges. This is reported as having a significance of <.010 and a Kappa at or near 1.000 (almost perfect). [a]Some of the students in this class completed Presentation Format 1, some completed Presentation Format 2.
*$p \leq .050$

Table I-2. (continued)

Teacher ID		9	10	11	12	13	14	15	16
n =		11	21	22	15	21	17	20	22
Presentation Format		2	2	1	a	2	2	1	1
Sub-element	Variable								
Creating New Ideas	F	-.148	-.266	-.119	-.304	-.303	.673*	-.132	-.088
	Y2	-.128	-.135	.062	-.250	-.153	-.141	-.117	.166
	Y4	.518*	.033	-.048	-.053	-.050	-.214	-.081	-.220
	Y6	-.222	-.167	-.097	-.228	.108	.259	-.162	.349*
	Y8	-.248	.055	.104	-.078	.065	-.119	-.181	.053
	Y10	-.048	-.478*	.133	-.210	.011	-.140	-.128	.235
Evaluating a Plan or Idea	F	-.128	-.032	.033	-.111	-.019	-.137	.087	.283
	Y2	-.031	-.132	-.158	-.257	-.203	-.037	-.187	-.030
	Y4	.129	-.050	.254	.084	-.161	.081	-.194	.042
	Y6	.365	-.059	-.398*	-.208	.045	-.259	.116	.088
	Y8	.060	-.188	-.135	-.111	.098	-.240	.104	-.026
	Y10	-.128	-.312	-.126	-.148	.100	.103	.072	.101

Note. # indicates the Kappa was not calculated due to perfect agreement between the judges. This is reported as having a significance of <.010 and a Kappa at or near 1.000 (almost perfect). [a]Some of the students in this class completed Presentation Format 1, some completed Presentation Format 2.
*$p \leq .050$

Table I-2. (continued)

Teacher ID		17	18	19	20	21	22	23	24
n =		15	17	22	17	11	6	22	23
Presentation Format		2	1	1	2	1	1	2	1
Sub-element	Variable								
Creating New Ideas	F	-.111	-.437*	-.204	-.126	-.073	-.091	-.521*	-.012
	Y2	-.121	-.220	-.128	-.010	-.209	-.143	-.481*	-.154
	Y4	-.089	-.259	-.137	-.179	.241	-.171	-.446*	-.033
	Y6	-.162	-.123	-.204	.022	.276	.172	-.238	-.131
	Y8	-.281	-.187	.100	-.221	.358	-.371	-.423*	.010
	Y10	-.126	-.348*	-.049	.181	.180	.467	-.291	.095
Evaluating a Plan or Idea	F	-.067	-.447*	-.187	-.159	.248	-.241	-.530*	-.133
	Y2	-.103	-.285	.040	.051	.224	-.091	-.409*	-.220
	Y4	.087	-.122	-.012	-.280	.417	-.029	-.425*	-.293
	Y6	-.086	-.163	.558*	-.133	.577*	-.143	-.430*	-.225
	Y8	.300	-.149	.261	-.169	.057	-.067	-.401*	-.176
	Y10	.049	-.443*	-.039	-.422	.038	.122	-.360*	.060

Note. # indicates the Kappa was not calculated due to perfect agreement between the judges. This is reported as having a significance of <.010 and a Kappa at or near 1.000 (almost perfect). [a]Some of the students in this class completed Presentation Format 1, some completed Presentation Format 2.
*$p \leq .050$

Table I-3. *Study Two. Teacher variables: number of years teaching*

Sub-element	Variable	First Year		2-4 Years		5 or more years	
		K	*p*	*K*	*p*	*K*	*p*
Asking Questions	F	-.034	.894	-.110	.301	.008	.844
	Y2	.012	.952	.046	.656	-.077	.065
	Y4	-.228	.272	-.091	.417	.023	.578
	Y6	.044	.831	-.013	.897	.071	.088
	Y8	.302	.170	-.090	.455	.020	.622
	Y10	.108	.574	.049	.626	.017	.664
Finding and Organising Information	F	-.145	.471	.123	.245	-.008	.841
	Y2	-.004	.983	.183	.074	.033	.406
	Y4	-.027	.898	.115	.260	-.026	.512
	Y6	.439	.041	.029	.784	.018	.661
	Y8	.183	.354	.070	.489	.090	.024
	Y10	.163	.394	.319	.002	.052	.181
Connecting Information	F	-.034	.894	-.119	.327	.129	.005
	Y2	-.214	.279	.128	.233	.059	.146
	Y4	-.163	.460	.092	.437	.138	.001

Table I-3. (continued)

Sub-element	Variable	First Year		2-4 Years		5 or more years	
		K	*p*	*K*	*p*	*K*	*p*
Connecting Information cont.	Y6	-.111	.667	.077	.526	.104	.019
	Y8	-.292	.258	.107	.293	.014	.735
	Y10	-.315	.100	.286	.005	.079	.048
Creating New Ideas	F	-.304	.239	-.110	.301	-.064	.159
	Y2	-.250	.333	-.004	.968	.047	.286
	Y4	-.053	.794	.104	.316	-.026	.529
	Y6	-.228	.272	.088	.397	.082	.044
	Y8	-.078	.732	.108	.302	.035	.401
	Y10	-.210	.286	.108	.268	-.004	.909
Evaluating a Plan or Idea	F	-.111	.667	.120	.260	-.022	.596
	Y2	-.257	.258	-.054	.603	-.003	.939
	Y4	.084	.676	.215	.037	.024	.554
	Y6	-.208	.354	.013	.908	.038	.371
	Y8	-.111	.667	.071	.492	.050	.210
	Y10	-.148	.524	.059	.548	.074	.055
	n =	15		53		339	

Table I-4. *Study Two. Teacher variables: number of years teaching the current year level*

Sub-element	Variable	First Year		First time last year		Third year or more	
		K	*p*	*K*	*p*	*K*	*p*
Asking Questions	F	-.110	.240	-.023	.877	.009	.848
	Y2	.061	.500	-.331	.003	-.042	.350
	Y4	-.069	.458	-.055	.636	.017	.700
	Y6	.036	.686	-.107	.360	.077	.082
	Y8	-.013	.894	.019	.870	.019	.676
	Y10	.060	.500	.213	.049	-.021	.616
Finding and Organising Information	F	.041	.657	-.028	.825	-.001	.982
	Y2	.049	.585	.077	.476	.042	.333
	Y4	.032	.729	.025	.818	-.029	.496
	Y6	.081	.381	.042	.729	.017	.685
	Y8	.097	.290	-.003	.979	.096	.024
	Y10	.225	.015	.232	.031	.034	.413
Connecting Information	F	-.011	.909	-.119	.363	.141	.004
	Y2	-.042	.664	.055	.631	.070	.106
	Y4	.002	.983	-.004	.972	.154	<.001

Table I-4. (continued)

Sub-element	Variable	First Year		First time last year		Third year or more	
		K	*p*	*K*	*p*	*K*	*p*
Connecting Information cont.	Y6	.130	.191	-.086	.476	.102	.034
	Y8	-.062	.506	.047	.672	.015	.723
	Y10	.037	.681	.088	.423	.101	.019
Creating New Ideas	F	-.132	.175	-.152	.233	-.055	.263
	Y2	-.032	.738	-.111	.370	.045	.336
	Y4	.173	.062	.061	.593	-.055	.208
	Y6	-.133	.163	.047	.672	.095	.030
	Y8	-.049	.602	.049	.675	.049	.272
	Y10	.031	.722	.077	.469	-.016	.709
Evaluating a Plan or Idea	F	.021	.828	-.017	.883	-.013	.771
	Y2	-.133	.163	.127	.259	-.010	.823
	Y4	.133	.146	.213	.056	.003	.952
	Y6	-.044	.636	.432	<.001	-.002	.970
	Y8	.003	.972	.202	.066	.029	.492
	Y10	.026	.772	.021	.848	.076	.068
	n =	68		44		295	

Table I-5. *Study Two. Teacher variables: Has previously used the AC C&CrT continuum to assess students*

Sub-element	Variable	No		Unsure		Yes	
		K	*p*	*K*	*p*	*K*	*p*
Asking Questions	F	-.066	.289	-.076	.618	.016	.757
	Y2	-.045	.455	-.004	.976	-.072	.159
	Y4	.030	.619	.025	.866	-.015	.772
	Y6	.026	.667	.076	.600	.079	.117
	Y8	.034	.579	.111	.439	.004	.935
	Y10	-.016	.777	.039	.783	.062	.201
Finding and Organising Information	F	.085	.192	-.311	.028	.001	.983
	Y2	.073	.225	-.042	.773	.052	.281
	Y4	-.039	.503	-.096	.489	.020	.688
	Y6	.123	.043	-.381	.016	.027	.582
	Y8	.035	.569	.192	.210	.113	.020
	Y10	.101	.086	-.009	.955	.087	.068
Connecting Information	F	.223	<.001	-.084	.611	.045	.428
	Y2	.097	.113	-.095	.539	.055	.267
	Y4	.113	.084	-.157	.414	.158	.002

Table I-5. (continued)

Sub-element	Variable	No		Unsure		Yes	
		K	*p*	*K*	*p*	*K*	*p*
Connecting Information cont.	Y6	.104	.116	-.149	.439	.117	.033
	Y8	.031	.613	-.111	.433	.024	.620
	Y10	.113	.061	-.010	.945	.105	.030
Creating New Ideas	F	-.046	.486	-.184	.261	-.069	.216
	Y2	.013	.858	.038	.830	.044	.399
	Y4	.109	.077	-.337	.026	-.018	.713
	Y6	.126	.032	-.114	.479	.100	.050
	Y8	.104	.114	-.083	.586	.029	.560
	Y10	-.026	.654	-.019	.896	.027	.570
Evaluating a Plan or Idea	F	.034	.591	-.033	.821	-.026	.609
	Y2	.091	.133	-.243	.090	-.029	.559
	Y4	.060	.330	-.174	.211	.075	.131
	Y6	.189	.004	-.276	.066	-.027	.590
	Y8	.037	.543	.163	.281	.084	.086
	Y10	.016	.792	.214	.142	.080	.094
	n =	158		27		222	

Table I-6. *Study Two. Teacher variables: Confidence to use the AC C&CrT continuum to assess students*

Sub-element	Variable	Not confident		Somewhat confident		Reasonably confident	
		K	*p*	*K*	*p*	*K*	*p*
Asking Questions	F	-.063	.505	-.022	.659	.058	.537
	Y2	-.115	.147	-.047	.347	-.041	.660
	Y4	.035	.670	-.004	.940	-.010	.913
	Y6	.034	.668	.078	.105	.033	.707
	Y8	-.056	.490	.016	.742	.157	.085
	Y10	.060	.427	.012	.788	.028	.742
Finding and Organising Information	F	.024	.785	-.001	.979	.016	.858
	Y2	-.049	.545	.063	.177	.151	.078
	Y4	.034	.666	-.040	.381	.041	.627
	Y6	-.021	.801	.067	.153	.006	.942
	Y8	.090	.245	.075	.119	.140	.099
	Y10	.168	.033	.083	.073	.020	.814
Connecting Information	F	.158	.123	.097	.067	.044	.648
	Y2	-.027	.748	.034	.474	.193	.023
	Y4	-.014	.869	.114	.028	.269	.002

Table I-6. (continued)

Sub-element	Variable	Not confident		Somewhat confident		Reasonably confident	
		K	*p*	*K*	*p*	*K*	*p*
Connecting Information cont.	Y6	.045	.613	.069	.206	.204	.025
	Y8	-.099	.224	.050	.284	.036	.674
	Y10	.050	.547	.072	.120	.248	.003
Creating New Ideas	F	-.091	.339	-.062	.230	-.098	.317
	Y2	.064	.516	-.057	.269	.169	.057
	Y4	.029	.741	.026	.585	-.081	.356
	Y6	.082	.335	.068	.152	.217	.014
	Y8	.015	.863	.093	.055	-.044	.622
	Y10	-.033	.669	.020	.649	.019	.819
Evaluating a Plan or Idea	F	-.120	.153	.044	.367	-.003	.972
	Y2	-.081	.324	-.030	.528	.104	.221
	Y4	-.022	.786	.036	.445	.182	.034
	Y6	.106	.253	-.020	.676	.060	.507
	Y8	.086	.309	.054	.250	.082	.335
	Y10	-.040	.610	.075	.098	.193	.021
	n =	87		248		71	

Table I-7. *Study Two. Student variables: Year level of students*

Sub-element	Variable	Year 4		Year 5		Year 6		Year 7	
		K	*p*	*K*	*p*	*K*	*p*	*K*	*p*
Asking Questions	F	.035	.642	-.023	.875	-.07	.268	.001	.994
	Y2	.083	.257	-.007	.964	-.19	.002	-.033	.642
	Y4	-.031	.675	-.098	.44	.087	.155	-.054	.465
	Y6	.081	.262	.098	.444	.005	.936	.032	.662
	Y8	-.026	.721	-.023	.861	.039	.515	.035	.641
	Y10	.003	.963	-.233	.055	-.019	.745	.118	.094
Finding and Organising Information	F	.105	.146	-.141	.366	-.013	.838	-.100	.176
	Y2	.037	.608	.295	.036	-.074	.209	.100	.162
	Y4	.074	.295	-.029	.811	-.102	.074	.042	.553
	Y6	.035	.621	-.082	.494	-.096	.122	.126	.085
	Y8	.208	.003	-.072	.582	.004	.949	.100	.157
	Y10	.055	.422	-.076	.528	.091	.129	.153	.029
Connecting Information	F	.080	.331	.281	.030	-.004	.953	-.029	.723
	Y2	.113	.109	-.075	.531	-.011	.862	.044	.546
	Y4	.097	.193	.147	.246	.032	.624	.164	.035
	Y6	.132	.086	-.010	.943	.048	.477	.099	.238

Table I-7. (continued)

Sub-element	Variable	Year 4		Year 5		Year 6		Year 7	
		K	*p*	*K*	*p*	*K*	*p*	*K*	*p*
Connecting Information cont.	Y8	-.039	.587	-.021	.867	-.093	.127	.093	.186
	Y10	.216	.002	.001	.996	.006	.920	.022	.760
Creating New Ideas	F	-.032	.69	.009	.947	-.117	0.088	-.114	.140
	Y2	.073	.33	-.121	.389	-.090	0.193	0	.999
	Y4	-.019	.804	-.096	.440	-.046	0.455	.064	.376
	Y6	.129	.076	.094	.432	-.007	0.911	.001	.991
	Y8	.032	.667	.037	.795	-.022	0.72	.087	.232
	Y10	.073	.294	-.048	.691	-.100	0.08	.065	.350
Evaluating a Plan or Idea	F	-.017	.814	.130	.319	-.101	.096	.057	.440
	Y2	.026	.711	.002	.988	-.068	.257	-.088	.234
	Y4	.001	.994	-.271	.045	.037	.531	.181	.013
	Y6	.003	.972	.112	.406	0	.997	-.013	.861
	Y8	-.040	.565	-.151	.221	.049	.422	.194	.008
	Y10	.092	.188	.001	.992	.042	.462	.024	.728
	n =	108		35		157		106	

Table I-8. *Study Two. Student variables: Sex of students*

Sub-element	Variable	Female		Male	
		K	*p*	*K*	*p*
Asking Questions	F	-.038	.472	.029	.627
	Y2	-.076	.134	-.030	.594
	Y4	.008	.873	-.009	.876
	Y6	.053	.289	.072	.201
	Y8	-.014	.769	.075	.192
	Y10	.023	.626	.038	.496
Finding and Organising Information	F	.024	.633	-.019	.751
	Y2	.036	.464	.081	.141
	Y4	.020	.679	-.035	.522
	Y6	.122	.014	-.063	.263
	Y8	.161	<.001	.015	.787
	Y10	.100	.035	.084	.124
Connecting Information	F	.102	.066	.108	.088
	Y2	.090	.065	.030	.603
	Y4	.132	.011	.114	.062

Table I-8. (continued)

Sub-element	Variable	Female		Male	
		K	*p*	*K*	*p*
Connecting Information cont.	Y6	.118	.033	.079	.195
	Y8	.045	.360	-.010	.861
	Y10	.120	.013	.076	.169
Creating New Ideas	F	-.113	.042	.002	.980
	Y2	.052	.329	.012	.852
	Y4	.011	.832	-.003	.951
	Y6	.152	.002	.017	.768
	Y8	.074	.147	.029	.621
	Y10	.049	.295	-.030	.578
Evaluating a Plan or Idea	F	-.018	.717	.028	.624
	Y2	.012	.813	-.015	.791
	Y4	-.039	.424	.186	<.001
	Y6	.074	.151	-.013	.823
	Y8	.118	.015	.011	.843
	Y10	.066	.160	.089	.160
	n =	233		174	

Table I-9. *Study Two. Student variables: Learning complexities – Presentation Format 1*

Sub-element	Variable	No condition		Autism Spectrum, Auditory Processing, Dyslexia		EALD		Hearing, Intellectual, One Plan	
		K	*p*	*K*	*p*	*K*	*p*	*K*	*p*
Asking Questions	F	-.054	.389	.542	.007	-.073	.756		
	Y2	-.105	.081	-.213	.258	.154	.513		
	Y4	.006	.920	-.184	.367	-.158	.495		
	Y6	.079	.188	.078	.670	-.069	.754		
	Y8	.004	.944	-.045	.807	-.128	.572		
	Y10	.044	.438	-.014	.941	-.093	.662		
Finding and Organising Information	F	-.077	.206	-.149	.436	-.236	.355		
	Y2	.015	.803	.192	.296	.371	.102		
	Y4	.077	.186	-.229	.215	.006	.976		
	Y6	-.020	.738	.175	.345	.313	.162		
	Y8	.173	.003	-.229	.215	.114	.607		
	Y10	.170	.003	-.115	.532	.313	.162		
Connecting Information	F	.016	.810	-.277	.189	.608	.055		
	Y2	.022	.709	.108	.575	-.141	.539		
	Y4	.053	.410	-.034	.871	.033	.887		
	Y6	.032	.624	-.184	.367	.132	.569		
	n =	156		15		11		Too few complete subjects	

Table I-9. (continued)

Sub-element	Variable	No condition		Autism Spectrum, Auditory Processing, Dyslexia		EALD		Hearing, Intellectual, One Plan	
		K	*p*	*K*	*p*	*K*	*p*	*K*	*p*
Connecting Information cont.	Y8	-.019	.755	.181	.326	-.141	.539		
	Y10	.102	.077	.261	.166	-.008	.975		
Creating New Ideas	F	-.128	.042	-.200	.439	-.209	.369		
	Y2	.050	.419	-.111	.578	-.039	.863		
	Y4	.012	.852	-.149	.436	-.316	.163		
	Y6	.061	.324	.007	.969	.126	.564		
	Y8	.061	.317	-.136	.491	.064	.781		
	Y10	.056	.324	.062	.740	.180	.399		
Evaluating a Plan or Idea	F	.047	.436	-.495	.010	-.173	.487		
	Y2	-.057	.351	-.229	.215	-.141	.539		
	Y4	.007	.900	-.045	.810	.426	.074		
	Y6	.003	.958	-.034	.857	-.173	.487		
	Y8	.081	.172	.175	.345	-.132	.588		
	Y10	.108	.057	-.313	.088	.132	.554		
	n =	156		15		11		Too few complete subjects	

Table I-10. *Study Two. Student variables: Learning complexities - Presentation Format 2*

Sub-element	Variable	No condition		Autism Spectrum, Auditory Processing, Dyslexia		EALD		Hearing, Intellectual, One Plan	
		K	*p*	*K*	*p*	*K*	*p*	*K*	*p*
Asking Questions	F	-.022	.695	-.067	.767	-.125	.708	-.120	.681
	Y2	-.014	.800	-.123	.583	-.326	.187	.176	.534
	Y4	-.004	.944	.422	.078	.217	.437	-.429	.171
	Y6	.071	.199	-.437	.050	-.014	.956	-.077	.839
	Y8	.060	.281	-.143	.520	-.169	.552	-.321	.311
	Y10	.036	.485	.092	.657	-.178	.453	-.312	.246
Finding and Organising Information	F	.085	.130	.018	.938	.390	.154	-.302	.325
	Y2	.085	.110	-.207	.324	-.248	.305	-.273	.391
	Y4	-.051	.326	-.016	.938	-.038	.873	-.200	.503
	Y6	.070	.193	-.110	.602	.191	.440	-.292	.275
	Y8	.091	.093	-.297	.151	-.416	.151	-.273	.391
	Y10	.049	.344	-.257	.210	.053	.832	.000	1.000
Connecting Information	F	.164	.009	.195	.406	.234	.378	-.273	.471
	Y2	.100	.064	-.207	.337	.217	.406	.208	.512
	Y4	.209	<.001	.067	.749	.379	.144	-.400	.290
	Y6	.180	.002	.040	.874	-.200	.549	.067	.860
	n =	195		12		9		7	

Table I-10. (continued)

Sub-element	Variable	No condition		Autism Spectrum, Auditory Processing, Dyslexia		EALD		Hearing, Intellectual, One Plan	
		K	*p*	*K*	*p*	*K*	*p*	*K*	*p*
Connecting Information cont.	Y8	.063	.234	-.168	.418	-.125	.607	-.321	.311
	Y10	.119	.024	.061	.770	-.455	.067	-.321	.311
Creating New Ideas	F	-.056	.363	.619	.032	-.304	.277	-.077	.839
	Y2	.001	.984	.161	.530	.357	.284	-.200	.503
	Y4	.014	.791	.018	.938	-.009	.969	.023	.940
	Y6	.116	.030	.298	.164	.432	.101	-.585	.065
	Y8	.058	.298	.138	.537	-.169	.552	-.186	.516
	Y10	-.041	.417	.092	.661	.135	.576	-.273	.338
Evaluating a Plan or Idea	F	.025	.656	-.073	.729	-.248	.305	-.373	.189
	Y2	.112	.037	-.270	.191	-.137	.580	-.377	.169
	Y4	.074	.173	.082	.696	.053	.832	.000	1.000
	Y6	.088	.120	-.249	.237	.100	.697	-.321	.311
	Y8	.056	.302	-.257	.210	.393	.112	-.228	.409
	Y10	.054	.298	.045	.832	.327	.167	.138	.605
	n =	195		12		9		7	

Table I-11. *Study Two. Results of analysis with Fleiss' Kappa following removal of outliers*

Sub-element	Variable	Presentation Format 1		Presentation Format 2	
		K	*p*	*K*	*p*
Asking Questions	F	-.019	.766	-.030	.570
	Y2	-.138	.023	-.025	.626
	Y4	-.021	.739	.012	.819
	Y6	.060	.327	.047	.356
	Y8	-.058	.336	.038	.467
	Y10	.010	.867	.027	.577
Finding and Organising Information	F	-.105	.093	.085	.105
	Y2	.005	.928	.056	.263
	Y4	-.014	.822	-.039	.413
	Y6	-.055	.371	.058	.244
	Y8	.130	.031	.050	.323
	Y10	.118	.050	.039	.428
Connecting Information	F	-.036	.614	.165	.004
	Y2	-.080	.209	.101	.047
	Y4	-.104	.130	.200	<.001
	Y6	-.012	.867	.161	.004
	Y8	-.086	.165	.042	.401
	Y10	.062	.294	.088	.076

Table I-11. (continued)

Sub-element	Variable	Presentation Format 1		Presentation Format 2	
		K	*p*	*K*	*p*
Creating New Ideas	F	-.158	.017	-.035	.544
	Y2	-.065	.325	.021	.710
	Y4	-.028	.663	.020	.688
	Y6	-.031	.629	.126	.012
	Y8	.006	.919	.051	.324
	Y10	.023	.691	-.023	.632
Evaluating a Plan or Idea	F	-.070	.262	-.001	.985
	Y2	-.133	.036	.067	.178
	Y4	-.025	.680	.074	.143
	Y6	-.080	.201	.057	.279
	Y8	.084	.174	.058	.248
	Y10	.065	.264	.078	.100
	n =	151		222	

Appendix J Study Three Supporting Documents

Appendix J-1. *Study Three. Teacher Survey*

Study Two Teacher Survey

Record ID	_____
Which year level/s are you teaching this year?	◯ Year 1,2 and/or 3 ◯ Year 4 ◯ Year 5 ◯ Year 6 ◯ Year 7
Do you teach full time in the class?	◯ Yes ◯ No
How many students do you have in your class this year?	◯ Less than 25 students ◯ 25 or more students
Thinking about the students in your class, please estimate (to the best of your knowledge) their level of experience in class in 2020 in terms of Problem-based Learning or formal Inquiry Learning (e.g. International Baccalaureat or Kath Murdoch pedagogy)	◯ No experience ◯ A little bit of experience ◯ Some experience ◯ Quite a lot of experience ◯ Lots of experience
Are you planning to explicitly teach critical thinking skills or creative thinking skills to your students in Term 1 or 2 this year?	◯ Yes ◯ No
Please provide your First and Last name	_____
Which school are you teaching in?	_____
Please provide your best contact details (phone or email you actually check and respond to!)	_____

Note. This study was re-classified as Study Three following the completion of research.

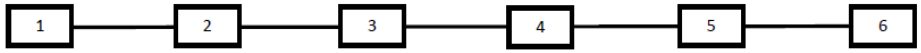
Appendix J-2. *Study Three. Critical and Creative Thinking Self-Report Measure*

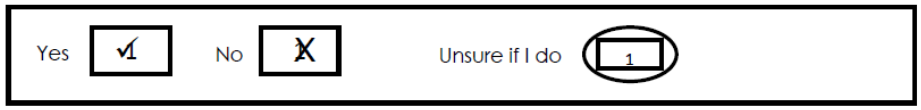
Class: _____ Name: _____

Critical and Creative Thinking

Example

Catching and throwing balls

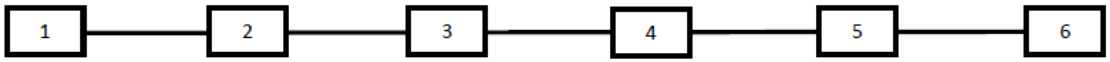




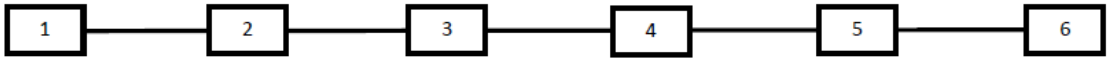
1

Critical and Creative Thinking

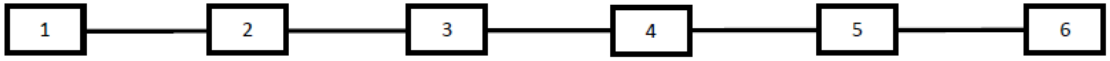
1. Asking Questions



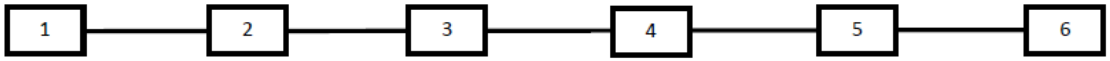
2. Finding and organising information



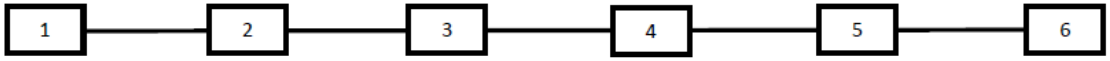
3. Connecting information

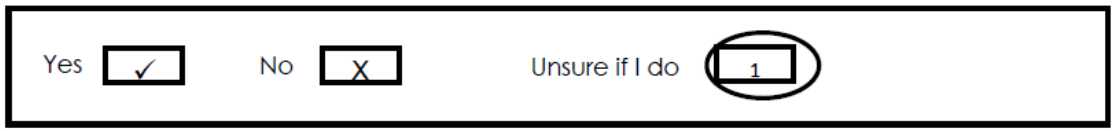


4. Creating new ideas



5. Evaluating a plan or idea





2

Motivation and Engagement

	Disagree Strongly	Disagree	Neither Agree nor Disagree	Agree	Agree Strongly
1	○	○	○	○	○
2	○	○	○	○	○
3	○	○	○	○	○
4	○	○	○	○	○
5	○	○	○	○	○
6	○	○	○	○	○
7	○	○	○	○	○
8	○	○	○	○	○
9	○	○	○	○	○
10	○	○	○	○	○
11	○	○	○	○	○

	Disagree Strongly	Disagree	Neither Agree nor Disagree	Agree	Agree Strongly
12	O	O	O	O	O
13	O	O	O	O	O
14	O	O	O	O	O
15	O	O	O	O	O
16	O	O	O	O	O
17	O	O	O	O	O
18	O	O	O	O	O
19	O	O	O	O	O
20	O	O	O	O	O
21	O	O	O	O	O
22	O	O	O	O	O

4

	Disagree Strongly	Disagree	Neither Agree nor Disagree	Agree	Agree Strongly
23	○	○	○	○	○
24	○	○	○	○	○
25	○	○	○	○	○
26	○	○	○	○	○
27	○	○	○	○	○
28	○	○	○	○	○
29	○	○	○	○	○
30	○	○	○	○	○
31	○	○	○	○	○
32	○	○	○	○	○
33	○	○	○	○	○

5

	Disagree Strongly	Disagree	Neither Agree nor Disagree	Agree	Agree Strongly
34	○	○	○	○	○
35	○	○	○	○	○
36	○	○	○	○	○
37	○	○	○	○	○
38	○	○	○	○	○
39	○	○	○	○	○
40	○	○	○	○	○
41	○	○	○	○	○
42	○	○	○	○	○
43	○	○	○	○	○
44	○	○	○	○	○

6

Table J-3. *Study Three. Motivation and engagement statements*

Higher Order Factors	First Order Factors	Statement code	Statement
Positive Motivation	Self Belief	SB1	If I try hard, I believe I can do my schoolwork well
		SB2	If I don't give up, I believe I can do schoolwork that is hard
		SB3	If I have enough time, I believe I can do well in my schoolwork
		SB4	If I try hard enough, I believe I can do all my schoolwork
	Valuing	V1	I'm able to use some of the things I learn at school in other parts of my life
		V2	Learning at school is important
		V3	What I learn at school will be useful in the future
		V4	It's important to understand what I'm taught at school
	Learning Focus	LF1	I feel very happy with myself when I really understand what I'm taught at school
		LF2	I feel very happy with myself when I do well at school by working hard
		LF3	I feel very happy with myself when what I learn at school shows me how something works
		LF4	I feel very happy with myself when I learn new things at school

Table J-3. (continued)

Higher Order Factors	First Order Factors	Statement code	Statement
Positive Engagement	Planning	PLN1	I usually have a plan for how to do my homework when I start it
		PLN2	Before I start a project, I plan out how I am going to do it
		PLN3	I have a plan for how to do my homework or projects when I start them
		PLN4	I usually stick to a homework plan
	Task Management	TM1	I usually do my homework in places where I can concentrate
		TM2	When I do homework, I get organized so I can do it well
		TM3	When I do homework, I try to find a place where I can do it well
		TM4	When I do homework, I usually do it where I can concentrate best
	Persistence	P1	If I can't understand my schoolwork, I keep trying until I do
		P2	If my homework is difficult, I keep working at it trying to figure it out
		P3	When I'm taught something that doesn't make sense, I spend time to try to understand it
		P4	I'll keep working at difficult schoolwork until I've figured it out

Table J-3. (continued)

Higher Order Factors	First Order Factors	Statement code	Statement
Negative Motivation	Anxiety	ANX1	When I have a project to do, I worry about it a lot
		ANX2	I worry about getting bad marks in tests and projects
		ANX3	When I do tests I don't feel very good
		ANX4	I worry about school and schoolwork
	Avoidance	AV1	The main reason I try at school is because I don't want people to think that I'm dumb
		AV2	The main reason I try at school is because I don't want people to think bad things about me
		AV3	The main reason I try at school is because I don't want to disappoint my parents
		AV4	The main reason I try at school is because I don't want my teacher to think bad things about me
	Uncertain Control	UC1	When I don't do well at school I don't know how to avoid that happening next time
		UC2	When I get a good mark I often don't know how I'm going to get that mark again
		UC3	When I get a bad mark I don't know how to avoid that happening next time
		UC4	I don't know how to get good marks at school

Table J-3. (continued)

Higher Order Factors	First Order Factors	Statement code	Statement
Negative Engagement	Self Sabotage	SS1	Sometimes I don't try hard at school so I can have a reason if I don't do well
		SS2	I sometimes don't work very hard at school so I can have a reason if I don't do well
		SS3	I sometimes waste time the night before a test so I can have a reason if I don't do well
		SS4	I sometimes leave homework until the last moment so I can have a reason if I don't do so well
	Disengagement	D1	Each week I'm trying less and less at school
		D2	I don't really care about school anymore
		D3	I'm not involved in things like class activities and class discussion at school
		D4	I've given up being interested in school

Appendix J-4. *Study Three. Teacher post-study questionnaire*

Dear (teacher),

Congratulations on having successfully completed a term of Problem Based Learning with your students.

Can you please respond to the following questions to help me understand more about how you engaged with the PBL process, and about how your students responded to this style of teaching.

1. Was this term fairly typical from the students point of view (apart from the PBL) or were there additional unusual activities, e.g. swimming lessons, excursions, school camp, sports day, etc?

2. How closely do you think you followed the PBL process as it was set out for you in the training session/s? If there were any changes you made to the process or how you managed the students, please describe them.

3. How did your students share information during the PBL research process?

4. Is there anything you would do differently with a PBL unit in the future?

5. Do you have any observations or comments about how your students engaged with this style of learning that you can share with me?

6. Do you anticipate continuing using PBL to teach with your students?

7. Do you have any other feedback?

Thank you so much for your considered responses!

Appendix J-5. *Study Three. Participation information and consent form*

Teacher Participant Information Study Two

University of South Australia

The impact of problem context on Critical and Creative Thinking

Researcher: Sue Gaardboe
Supervisor: Prof David Cropley
Unit: UniSA STEM
Human Ethics Protocol No: 202895

I am carrying out research for the purpose of my PhD, to find out whether using an industry partner and their genuine, messy, unsolved problem in STEM Problem-Based Learning, has any effect on helping primary school students develop critical and creative thinking skills.

For this study (Study Two) I am looking for teachers who have:

- a year 4, 5, 6 or 7 class in 2021
- an interest in teaching a 10 week STEM Problem-based Learning unit to their class in Term 3 2021 (no previous experience necessary)
- a class who have little or no previous experience of instruction through Inquiry or Problem-based Learning.

Study Two will be examining the effects of using an industry partner and a genuine industry problem on the development of critical and creative thinking skills in students.

You have been contacted as a potential collaborator because you have either worked with me in the past or have been recommended to me by a colleague.

This Participant Information Sheet tells you about the research project. It explains the processes involved with taking part. Knowing what is involved will help you decide if you want to take part in the research.

Please read this information carefully. Ask questions about anything that you don't understand or want to know more about.

Participation in this research is voluntary. If you don't wish to take part, you don't have to. Your decision to take part or not to take part (or to take part and then withdraw), will not affect your relationship with your employer, the researcher or the University of South Australia, now or in the future.

If you decide you want to take part in the research project and fit the required criteria, you will be asked to sign the consent section. By signing it you are telling us that you:

- Understand what you have read
- Consent to take part in the research project
- Consent to the use of your information as described.

This Participant Information Sheet is for you to keep.

Aim of the Project

This project aims to investigate the effect of the industry partner and their problem, on student critical and creative thinking, when used in a STEM Problem-Based Learning (PBL) scenario in a primary school setting.

Very little research has been done on the Critical and Creative Thinking Skills in the Australian Curriculum, and there is no current research on how much of an effect using an industry partner has on the development of students' skills. In fact, there is very little research done on primary school students and their learning as a whole, so this research is important as it helps fill this gap in world-wide, evidence-based research.

You may be aware that the use of industry partners and their problems are currently being promoted as an effective way to teach STEM in schools. This research will help us understand more about the effect this has on student learning, and whether the investment of the time and effort it takes to engage an industry partner in our students' education is worthwhile.

Procedures and activities

Teachers will be randomly assigned to either a control group or an intervention group.

Teachers in a Control Group:

1. The researcher and teacher will agree on suitable lesson times in Week 1 and Week 10 in Term 3, and around Week 4 in Term 4 2021 to run the assessment task.

2. All students will be provided with a consent form to complete prior to the assessment task being run, which is to be distributed and collected by the classroom teacher.
3. The researcher will present the assessment task to the students during one lesson (approximately 50 minutes). At conclusion of the lesson the researcher will collect the completed assessments for analysis.
4. Standard student information (name, age, gender, year group, NEP/One Plan/ILP, year level for English instruction) will be required for each student. All student data will be made anonymous once it has been collated with the student information.

Teachers in an Intervention Group:

9. Teachers eligible for Study Two will be provided with free training (1 ½ days) in STEM PBL during the April school holidays or during Term 2. This will include opportunities to plan the teaching and identify curriculum connections.
10. The researcher and teacher will agree on suitable lesson times in Week 1 and Week 10 in Term 3, and around Week 4 in Term 4 2021 to run the assessment task.
11. Teachers and their class will be randomly assigned to a test group.
12. All students will be provided with a consent form to complete prior to the assessment tasks being run, which is to be distributed and collected by the classroom teacher.
13. The STEM PBL unit will be run by the teacher in Term 3 2021.
14. After school support sessions for the teacher will be provided by the researcher as required.
15. The researcher will present the assessment tasks to the students during two lessons (approximately 50-minute lessons) in week one Term 3, in week 10 Term 3, and in week 6 term 4 (negotiable). At conclusion of the lessons the researcher will collect the completed assessments for analysis.
16. Standard student information (name, age, year group NEP/One Plan, ILP, year level for English instruction) will be required for each student. All student data will be made anonymous by allocation of a coded ID.
17. The researcher will provide the teacher with a copy of student responses for use in school reporting.

This research will examine the impact of an industry partner and their problem on student critical and creative thinking and on student motivation and engagement. Each scenario (industry partner + industry problem, industry partner + hypothetical problem, industry problem only, hypothetical problem only) will be tested with a minimum of 39 students. All testing will be carried out in primary schools in South Australia.

The benefits to you, of being involved in this research, will be a) to provide you with free training in STEM PBL, b) to support you in developing a greater understanding of the Critical and Creative Thinking skills in the General Capabilities, and c) to provide you with data on

your students which you may find helpful in your reporting to parents for 2021. In the longer term, the benefits of this research will be in helping us develop effective methods of teaching and assessing student skills.

It is not anticipated that there are any risks to participation in this study beyond those encountered in everyday life.

If you do consent to participate, you may withdraw at any time. Should you withdraw from the project, incomplete student data that has been collected will be excluded from the research, however it may not be possible to withdraw data from the study results if these have already had identifying details removed. If you decide to withdraw from the project, please notify a member of the research team before you withdraw. A member of the research team will inform you if there are any special requirements linked to withdrawing.

What will happen to the information?

By signing the consent form you consent to taking part in the research project. Any information obtained in connection with this research project that can identify you or your students will remain confidential. All student information will be de-identified immediately after it is collated. De-identified information collected as part of the study will be stored as digital data on a secure UniSA server or hard copy data kept in a locked facility. This data will be stored for 7 years on completion of the research (standard University of South Australia policy.

The information collected will only be used for the purpose of this research project. It will only be disclosed with participants' written consent, except as required by law.

How do I find out about the research?

A summary of the project results is expected to be available in July 2023, from Sue Gaardboe susan.gaardboe@mymail.unisa.edu.au. Individual class results will be provided to the teachers two weeks after the data collection.

Who has reviewed the research project?

The ethical aspects of this research project have been approved by the Human Research Ethics Committee (HREC) of the University of South Australia as required by the Australian government research requirements, specified in the National Statement on Ethical Conduct

in Human Research (2007 - updated 2018). This statement has been developed to protect the interests of people who agree to participate in human research studies.

Research contact person

Name	Mrs Sue Gaardboe
Position	PhD candidate
Telephone	0405149933
Email	susan.gaardboe@mymail.unisa.edu.au

If you have any complaints about any aspect of the project, the way it is being conducted or any questions about being a research participant in general, please contact:

Reviewing HREC approving this research and HREC Executive Officer details

Reviewing HREC name	University of South Australia Human Research Ethics Committee
HREC Executive Officer	Ms Vicki Allen
Telephone	+618 8302 6330
Email	humanethics@unisa.edu.au

University of South Australia

Human Research Ethics Committee
Teacher Consent Form

This project has been approved by the University of South Australia's Human Research Ethics Committee. If you have any ethical concerns about the project or questions about your rights as a participant please contact the Executive Officer of this Committee, Tel: +618 8302 6330; Email: humanethics@unisa.edu.au.

SECTION 1: CONTACT AND PROJECT DETAILS	
Researcher's Full Name	Susan Gaardboe
Contact Details	0405149933
Supervisor's Name *(students only)*	Prof David Cropley
Contact Details	83025344
Project Number	202895
Project Title	Study Two: Impact of Problem Context on Student Learning

SECTION 2: PARTICIPANT CERTIFICATION

In signing this form, I confirm that:

- I have read the Participant Information Sheet, and the nature and the purpose of the research project has been explained to me. I understand and agree to take part.
- I understand the nature of my involvement in the project.
- I understand that I may not directly benefit from taking part in the project.
- I understand that I can withdraw from the project at any stage and that this will not affect my status now or in the future.
- I confirm that I am over 18 years of age.
- I understand that while information gained during the project may be published, I will not be identified and personal results will remain confidential, unless required by law.

Participant's Signature	*Printed Name*	*Date*

SECTION 3: RESEARCHER CERTIFICATION

I have explained the study to the participant and consider that he/she understand what is involved.

Researcher Signature	*Printed Name*	*Date*

University of South Australia

The impact of problem context on Critical and Creative Thinking

Researcher: Sue Gaardboe
Supervisor: Prof David Cropley
Unit: UniSA STEM
Human Ethics Protocol No: 202895

I am carrying out research for the purpose of my PhD, to find out whether using an industry partner and their genuine, messy, unsolved problem in STEM Problem-Based Learning (STEM PBL), has any effect on helping primary school students develop critical and creative thinking skills. For the purpose of this research, STEM PBL refers to teaching and learning which incorporates knowledge from multiple disciplines such as Science, Technology, Engineering and Maths, to enable problems to be solved.

This study will be examining the effects of using an industry partner and a genuine industry problem on the development of critical and creative thinking skills in students.

This Participant Information Sheet tells you about the research project. It explains the processes involved with taking part. Knowing what is involved will help you decide if you want to take part in the research.

Please read this information carefully. Ask questions about anything that you don't understand or want to know more about.

If you decide you want to take part in the research project you will be asked to sign the consent section. By signing it you are telling us that you:

- Understand what you have read
- Consent to take part in the research project
- Consent to the use of your information as described.

This Participant Information Sheet is for you to keep.

Aim of the Project

This project aims to investigate the effect of the industry partner and their problem, on student critical and creative thinking, when used in a STEM Problem-Based Learning (PBL) scenario in a primary school setting.

Very little research has been done on the Critical and Creative Thinking Skills in the Australian Curriculum, and there is no current research on how much of an effect using an industry partner has on the development of students' skills. In fact, there is very little research done on primary school students and their learning as a whole, so this research is important as it helps fill this gap in world-wide research.

Procedures and activities

1. All students will be provided with a consent form to complete prior to the assessment tasks being run.
2. Teachers and their class will be randomly assigned to a control or test group.
3. The researcher will present two multiple choice assessment tasks to the students during two lessons (approximately 50-minute lessons) in approximately week one Term 3, in week 10 Term 3, and in week 6 term 4 (negotiable). At conclusion of the lessons the researcher will collect the completed assessments for analysis.
4. School statistical data relating to each student (i.e. - age, gender, year level, special considerations (NEP/ILP), English language instruction level) will be provided to the researcher for each student taking part in the research.

This research will examine the impact of an industry partner and their problem on student critical and creative thinking and on student motivation and engagement. Each scenario (industry partner + industry problem, industry partner + hypothetical problem, industry problem only, hypothetical problem only, control group) will be tested with a minimum of 39 students. All testing will be carried out in primary schools in South Australia.

The benefits to you, of being involved in this research, will be to support you in developing a greater understanding of the Critical and Creative Thinking skills through your child's engagement in the activities. In the longer term, the benefits of this research will be in helping us develop effective methods of teaching and assessing student skills.

It is not anticipated that there are any risks to participation in this study beyond those encountered in everyday life.

Participation in this research is voluntary. If you don't wish to take part, you don't have to. Your decision to have your child take part or not to take part (or to take part and then withdraw), will not affect your/your child's relationship with the school, your child's teacher, the researcher or the University of South Australia, now or in the future.

It is expected that all students will be required by the school to complete the STEM PBL unit and assessments as part of their standard class work. The students' individually identifiable results will be made available to their respective class teacher for the purposes of internal assessments/ reporting to parents - but results of the testing will only be used for the purposes of this research with parents' explicit consent. No individuals will be identified in the external reporting of results. The school may choose to provide alternate activities for students not taking part in the assessment activities. Your decision not to allow the results of the assessment activities to be used in the research will in no way affect your child's opportunity to engage in the PBL learning tasks run by the school. If you do consent to participate you may withdraw at any time. Data will be excluded from participants who wish to withdraw, however it may not be possible to withdraw your data from the study results if these have already had your identifying details removed.

What will happen to the information?

By signing the consent form you consent to the research team collecting and using information from you for the research project. Any information obtained in connection with this research project that can identify you or your child will remain confidential. All student information will be de-identified immediately after it is collated. De-identified information collected as part of the study will be stored as digital data on a secure UniSA server or hard copy data kept in a locked facility. This data will be stored for 7 years on completion of the research (standard University of South Australia policy).

The information collected will only be used for the purpose of this research project. It will only be disclosed with your written consent, except as required by law.

How do I find out about the research?

A summary of the project results is expected to be available in July 2023, from Sue Gaardboe susan.gaardboe@mymail.unisa.edu.au. Individual class results will be provided for teacher use after the data collection is concluded.

Who has reviewed the research project?

The ethical aspects of this research project have been approved by the Human Research Ethics Committee (HREC) of the University of South Australia as required by the Australian government research requirements, specified in the National Statement on Ethical Conduct in Human Research (2007 - updated 2018). This statement has been developed to protect the interests of people who agree to participate in human research studies.

Research contact person

Name	Mrs Sue Gaardboe
Position	PhD candidate
Telephone	0405149933
Email	susan.gaardboe@mymail.unisa.edu.au

If you have any complaints about any aspect of the project, the way it is being conducted or any questions about being a research participant in general, please contact:

Reviewing HREC approving this research and HREC Executive Officer details

Reviewing HREC name	University of South Australia Human Research Ethics Committee
HREC Executive Officer	Ms Vicki Allen
Telephone	+618 8302 6330
Email	humanethics@unisa.edu.au

This project has been approved by the University of South Australia's Human Research Ethics Committee. If you have any ethical concerns about the project or questions about your rights as a participant please contact the Executive Officer of this Committee, Tel: +618 8302 6330; Email: humanethics@unisa.edu.au.

SECTION 1: CONTACT AND PROJECT DETAILS	
Researcher's Full Name	Susan Gaardboe
Contact Details	0405149933
Supervisor's Name *(students only)*	Prof David Cropley
Contact Details	83025344
Project Number	202895
Project Title	Study Two: Impact of Problem Context on Student Learning

SECTION 2: PARTICIPANT CERTIFICATION

In signing this form, I confirm that:

- I have read the Participant Information Sheet, and the nature and the purpose of the research project has been explained to me. I understand and agree to my child taking part.
- I understand the nature of my child's involvement in the project.
- I understand that my child may not directly benefit from taking part in the project.
- I understand that I can withdraw my child from the project (assessment tasks) at any stage and that this will not affect my child's status now or in the future.
- I confirm that I am over 18 years of age.
- I understand that while information gained during the project may be published, my child will not be identified and their personal results will remain confidential, unless required by law.

Parent's Signature	*Printed Name*	*Date*
Participant's Signature	*Printed Name*	*Date*

SECTION 3: RESEARCHER CERTIFICATION

I have explained the study to the participant and consider that he/she understand what is involved.

Researcher Signature	*Printed Name*	*Date*

Appendix J-6. *Study Three. STEM Problem-Based Learning teacher training programme and resources*

Day 1		
Morning Session 1	What is STEM PBL? Why STEM PBL?	• What is STEM • What are the forms and differences between styles of PBL • Where do these styles of teaching fit in the Australian Curriculum/ DfE (STEM Learning strategy) • Assessing STEM learning
	Critical and Creative Thinking (C&CrTh)	• Where does it fit in the Australian Curriculum • Assessing C&CrTh (it's not about the solution, all about the process)
	The research project	• Significance of the research – its origins, purpose and results so far
Morning Session 2	The problem	• Identifying a problem • The role of industry partners • Use of commercial resources
Afternoon Session 1	Setting up a PBL Unit	• Programming for a 10 week unit • The Engineering Design Process • Useful resources (incl use of technologies, DfE resources, mentors, etc)
Afternoon Session 2	The research project	• Clarifying the level of support to be provided by researcher and expectations by participants

Day 2		
Morning Session 1	Planning in action – Personalised unit development	• Planning with TfEL/ Australian Curriculum • Mapping out the classroom program for Term 3 2021 (weekly activities and expectations)
Morning Session 2	Planning in action	• Step 1 – mindmapping the problem • Developing a rubric/ success criteria • Brainstorming mentors/industry connections/etc.

References

DfE web page, STEM school-industry partnerships, including videos of my previous work in this area: https://edi.sa.edu.au/educating/curriculum-strategies/stem-education/industry-partnershipsDepartment for Education and Child Development. (2016). *STEM Learning Strategy for DECD Preschool to Year 12.* Department for Education and Child Development South Australia.

Education Council. (2015). *National STEM School Education Strategy: A comprehensive plan for science, technology, engineering and mathematics education in Australia*. Education Council. http://www.educationcouncil.edu.au/site/DefaultSite/filesystem/documents/National%20STEM%20School%20Education%20Strategy.pdf

Rossi, D., Crisp, T., Andrew, J., Pearse, S. (2019). STEM School-industry partnerships. Department for Education. https://edi.sa.edu.au/library/document-library/educational-resources/stem-learning/industry_partnership_guide.pdf

Sample Lesson Outline for 10-week STEM PBL unit*

Week	Engineering Design Process stage	Activity	Notes	Lessons/ week
1	Define Identify	• Introduce the problem • Mind mapping (whole class) • Explaining the PBL Engineering Design process to the students	• industry partner presents to students (where used) • Students define the question to be investigated • Engineering Design diagram used to illustrate the process	1 - 2
2	Research Plan	Students work in research group – topic 1 (3-5 students/group)	• Teacher to determine research groups based on results of mind mapping. Students initially placed in group of personal interest where possible • Instigate method of information sharing between research groups (e.g. google docs, Teams, etc. All information to be shared) • Commence research	1 - 2
3		Students work in research group – topic 1	Research	1 - 2
4		Students work in research group – topic 2	Students move to second research topic (rotation of groups determined by mind mapping results)	1 - 2
5		Students work in research group – topic 2		1 - 2
6	Create	Students form solution groups (preferably 3-5 solution groups, including members of each topic group)	• Students each choose preferred solution from those identified and re-group into solution groups. • Produce prototype, presentation etc. incorporating research carried out on all research topics	2-3
7	Critique	Preliminary presentation of solutions (5 mins max/ group)	• Present to industry partner, member of Leadership team, etc • Feedback given to students on further information required or specific issues with the proposed solution	1
8	Redesign	Redesign and finalise model and presentation	Students re-design/ incorporate required changes, complete prototype or presentation	1 - 2
9				1 - 2
10	Share	Final presentations	Presentations to Industry partner/ panel/ other students etc.	1

*Each teacher must create their own Unit Plan which will fit with their school and teaching plan for the term. This sample Unit Plan incorporates the essential STEM PBL components, but does not specify the curriculum lessons (science, technologies, maths) that will be scheduled by the teacher to support the student work through the STEM PBL unit. These are planned to relate to the topic of the PBL investigation, but are taught through the regular planned Science, Maths and Technology lessons.

Web resources for Teachers

- Promoting a PBL Mindset: The "Dimmer Switch" Approach (Kamps, 2021)
- Searching for an answer: The critical role of new literacies while reading on the Internet (Henry, 2006)
- How to Read a Web Address (November, 2012)
- Problem-based Learning: As authentic as it gets (Stepien & Gallagher, 1993)
- Project Design Rubric for Gold Standard PBL (Buck Institute for Education PBL Works, nd)
- Problem Based Learning: A classroom Journey – film produced for the Department for Education following the journey of teachers and students learning PBL. (Fenton & Enlightening Films, 2019) https://vimeo.com/318145983

References

Buck Institute for Education PBL Works. (nd). Project Design Rubric. *Gold Standard PBL: Essential Project Design Elements.* Retrieved from https://my.pblworks.org/resource/document/project_design_rubric

Fenton, H., & Enlightening Films. (2019). Problem Based Learning: A classroom journey. *STEM Works Leaders Support*.

Henry, L. A. (2006). SEARCHing for an answer: The critical role of new literacies while reading on the Internet. *The reading teacher, 59*(7), 614-627. Retrieved from https://www.researchgate.net/publication/289355049_SEARCHing_for_an_Answer_The_Critical_Role_of_New_Literacies_While_Reading_on_the_Internet

Kamps, K. (2021). Promoting a PBL Mindset: The "Dimmer Switch" Approach. Retrieved from https://www.pblworks.org/blog/promoting-pbl-mindset-dimmer-switch-approach

November, A. (2012). How to Read a Web Address. Retrieved from https://novemberlearning.com/literacy_resource/4-how-to-read-a-web-address/

Stepien, W., & Gallagher, S. (1993). Problem-based learning: As authentic as it gets. *Educational Leadership, 50*, 25-25. Retrieved from https://www.ascd.org/el/articles/problem-based-learning-as-authentic-as-it-gets

Examples of Questions for Research

The question you develop with the students will obviously depend on the topic, but it's important to keep it open enough to enable multiple areas of research, but tight enough so the students can keep going back to it to remind themselves of what their solution needs to demonstrate.

'How can…' or 'What could…' type questions encourage multiple options.

You will need to move between the two activities of Defining the Problem and Identifying the Issues, to ensure that the final question for investigation encompasses all the areas the students identify. Remember the final question is developed by the students (with your assistance) but may NOT match the question you had developed when initially considering the problem.

Some examples of successful questions we have used with students are:

'How can we best transfer water to bypass the construction site and deliver the necessary flow?'

'How can we help Australian Vermiculture convince Australian farmers to use their products?'

'How can we improve the skills, performance and wellbeing of current female players whilst encouraging and promoting cross-coders and new athletes to the AFC?'

'How can we help Fulton Hogan safely and sustainably stop the erosion on lagoon walls 4 and 5 while keeping the Bolivar Treatment Plant running?'

'How can we increase the number of birds that come to the Magill Smart School without negatively impacting on the environment?'

How could you improve on this question: 'What makes a nature play space engaging?'?

(With thanks to Christie Evans, Brighton Primary School for permission to share questions)

Mentors in Class

You've been given this to read because you have agreed to work with the students as a Mentor.

What is a student mentor?

A student mentor is anyone who offers guidance to students. You are not here as a teacher or expected to have answers and you don't need to know much about their project area.

What is my role?

Your role is to help **guide** the students to identify what information they need to know (without actually telling them). You may also be able to help them work out how find the information, and even what to do with the information to achieve the goal or target set for them. It is about imparting your experience of learning and problem solving, without giving them the answer. The most important way you can help them is to ask questions that get the students to think! The students need to make all the decisions. Even if you think you know a better way to do something, or have what you think is a great idea to solve a problem, it's important that you try not to lead the students – they need to come up with the ideas themselves, and retain ownership of their learning.

What might this look like?

You might sit with a student or group of students who are searching for information on the internet.
Begin by having them explain what they are searching for. Ask questions until they can articulate precisely what information they need – this will help them to search more effectively. You might ask them to explain how they have chosen their search terms, and encourage them to think about whether they are the most efficient terms they could be using. You might also help them to understand what they are looking at on the internet, particularly if it is technical information, or very complex language.

The table over the page gives examples of some questions you might ask in different situations.

Most importantly, DO NOT TRY TO LEAD THE STUDENTS IN A PARTICULAR DIRECTION.

Your role is to concentrate on helping students find their own direction, and to encourage them to move forward if they have stalled.

You are welcome to answer any specific questions the students ask you. If you are an expert in an area, your knowledge may be helpful to the students, but please let them lead the way in asking questions, rather than trying to help them find 'the right answer'.

In this activity, it is most important that students are empowered to discover information rather than believing there is a "correct" answer.

Sue Gaardboe, UniSA 2.5.22

	Question examples	Purpose of question
Information search on the internet or in printed literature	Can you explain to me what information you are trying to find out?	Help students narrow down their search and to conduct more effective searches
	What search terms are you using?	
	Is there a more specific term/ different wording/ another way to phrase the search that would get you a better result?	
	How do you know which of the search results are reliable websites?	
For students who are not sure what to do next	What have you already done?	Help direct their focus to the next step
	What was the goal or target you're aiming to achieve?	
	What do you think the next step/s should be? (it sometimes help to draw a graphical representation)	
	If you don't know, how could you find out?	
General questions	Why do you think that?	Clarify thinking and identify misconceptions
	Why do you think that would work?	
	What more would you need to know?	Focus back on the main question/ issue
	Is that going to satisfy the success criteria/ answer the main question? (How, why, why not?)	

Sue Gaardboe, UniSA 2.5.22

Appendix J-7. *Study Three. PBL problems provided for the intervention*

Study Three Theory Problem

Task card: Grow 'fast food' to sell at market. From: Lioncrest. (2023). Macmillan STEM Investigations. https://lioncrest.com.au/our-range/stem-science/macmillan-stem-investigations-detail

The following text was provided to the teachers on a colourful card including images of food and a market stall.

Grow 'fast food' to sell at market

Your school is hosting a farmers' market. People usually go to farmers' markets to buy fresh, local produce. You want to make some money quickly, so you decide to sell some homegrown food at the market. You'll need a steady supply of fresh produce – and you'll need to grow it fast!

Your challenge is to design and create a greenhouse that will help you grow fast-growing plants to sell at market. Think about ways to make the project sustainable so you always have produce to sell; for example, you could use seeds from your plants to keep growing more plants. Also, keep in mind who might buy your produce. Could the way it is grown help sell your produce? As part of your STEM investigation, you will learn how the growth and survival of living things is affected by the physical conditions of their environment.

Key Understandings

This challenge will help you understand:

- Plants need water and sunlight to grow.
- Plants transform light energy into chemical energy that they can use to grow.
- Some plants grow more quickly than others.
- A plant's environment can affect how it grows and survives.
- A greenhouse can help plants grow faster.
- Sustainable gardening includes harvesting seeds from plants to grow new plants.

Key Skills

This challenge will help you:

- Make careful observations
- Consider different points of view
- Develop research skills
- Create environments to grow fast-growing edible plants

Farmers' markets are a great source of local produce. Can you find out which fresh foods people like to buy there?

Key Actions

- During this challenge, you will perform the following tasks:
- Find out about the best conditions for growing plants.
- Explore ways to create ideal growing conditions.
- Create a greenhouse for growing edible plants.
- Learn about automated watering systems.
- Find out about fast-growing edible plants.
- Grow edible plants to sell.

Key vocabulary

- conditions
- food choices
- food production
- greenhouse
- nutrient
- sustainable

Big Picture

In some parts of the Netherlands, 80 per cent of farming land is covered with greenhouses. The greenhouses help this tiny, cool-climate country produce huge amounts of food.

Design thinking stages

As part of your challenge, you need to follow the five-step design thinking process.

Empathise: Gather information about the types of fresh foods people like or expect to buy at a farmers' market, consider cost and identify reasons people do or don't buy food from markets.

Define: Work out the main issues with growing food fast.

Ideate: Come up with some creative ideas for growing food quickly in a homemade greenhouse.

Prototype: Choose one or two of your ideas and create them. Spend time making changes and trying new ways to create your idea.

Test: Check which ideas work best. Choose a final design to share.

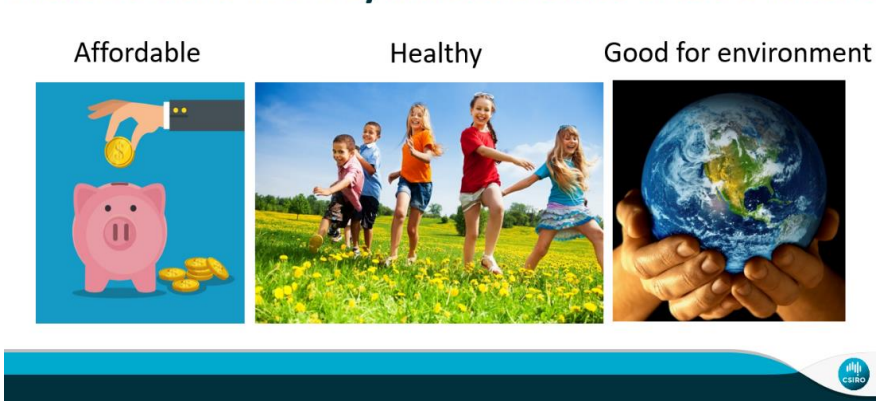
Study Three Industry Partner Problem

Selected images from the industry partner presentation to the students. From: Belobrajdic, D. (2022) Can we utilise nutrients from the oceans to produce healthy foods that taste good [PowerPoint slides].









Opportunity: Microalgae and macroalgae



>40% of global photosynthesis is contributed by algae

world's fastest-growing plants and can double their biomass daily

only 10 per cent of some 40,000 species isolated and cultured

Opportunity: Microalgae and macroalgae



The big ?
Produce high quality proteins that can be used in foods that taste great



How we are tackling the big ?



- Algae Culture Collection — •CSIRO, Hobart
- Identify best extraction & processing methods — •CSIRO, Werribee •CSIRO, Adelaide
- Evaluate protein quality — •CSIRO Adelaide •Use lab models of human digestion •Conduct clinical trials
- Optimise growth conditions of target algae — UTS
- Scale up production



LC-mass spectrometer





Appendix K Study Three Data Analysis

Table K-1. *Study Three. Tests of normality for critical and creative thinking assessment, control group: Asking Questions and Finding and Organising Information*

Treat-ment group	Time Point	Year level statement	Asking Questions					Finding and Organising Information				
			Missing cases	Skew	Std. Error	Kurtosis	Std. Error	Missing cases	Skew	Std. Error	Kurtosis	Std. Error
Control *N* = 64	1	F	2 (3.1%)	-5.427	.304	28.365	.599	2 (3.1%)	-0.733	.304	-1.004	.599
		Y2	2 (3.1%)	-0.585	.304	-1.023	.599	2 (3.1%)	-1.089	.304	-0.375	.599
		Y4	2 (3.1%)	-2.690	.304	6.858	.599	3 (4.7%)	-0.830	.306	-1.193	.604
		Y6	2 (3.1%)	-1.201	.304	-0.168	.599	3 (4.7%)	-1.435	.306	0.144	.604
		Y8	2 (3.1%)	-3.157	.304	8.232	.599	2 (3.1%)	-0.606	.304	-1.372	.599
		Y10	3 (4.7%)	-0.580	.306	-1.403	.604	3 (4.7%)	-0.098	.306	-1.734	.604
	2	F	10 (15.6%)	-3.192	.325	9.839	.639	10 (15.6%)	-0.849	.325	-0.550	.639
		Y2	10 (15.6%)	-1.698	.325	1.970	.639	11 (17.2%)	-1.458	.327	1.044	.644
		Y4	10 (15.6%)	-2.254	.325	4.506	.639	10 (15.6%)	-1.698	.325	1.970	.639
		Y6	10 (15.6%)	-1.381	.325	0.699	.639	10 (15.6%)	-1.729	.325	1.381	.639
		Y8	10 (15.6%)	-2.892	.325	6.607	.639	10 (15.6%)	-0.971	.325	-0.383	.639
		Y10	10 (15.6%)	-0.886	.325	-0.607	.639	10 (15.6%)	-0.570	.325	-1.325	.639

Table K-1. (continued)

Treat-ment group	Time Point	Year level statement	Asking Questions					Finding and Organising Information				
			Missing cases	Skew	Std. Error	Kurtosis	Std. Error	Missing cases	Skew	Std. Error	Kurtosis	Std. Error
Control *N* = 64	3	F	4 (6.3%)	-3.408	.309	12.091	.608	4 (6.3%)	-1.094	.309	-0.395	.608
		Y2	4 (6.3%)	-1.148	.309	0.106	.608	4 (6.3%)	-1.055	.309	-0.548	.608
		Y4	4 (6.3%)	-3.254	.309	9.719	.608	4 (6.3%)	-1.749	.309	1.273	.608
		Y6	4 (6.3%)	-1.046	.309	-0.288	.608	4 (6.3%)	-1.716	.309	1.440	.608
		Y8	4 (6.3%)	-4.384	.309	20.376	.608	4 (6.3%)	-1.148	.309	0.106	.608
		Y10	4 (6.3%)	-0.958	.309	-0.236	.608	4 (6.3%)	-1.068	.309	-0.149	.608

Table K-2. *Study Three. Tests of normality for critical and creative thinking assessment, control group: Connecting Information and Creating New Ideas*

Treatment group	Time Point	Year level statement	Connecting Information					Creating New Ideas				
			Missing cases	Skew	Std. Error	Kurtosis	Std. Error	Missing cases	Skew	Std. Error	Kurtosis	Std. Error
Control *N* = 64	1	F	2 (3.1%)	-2.023	.304	2.674	.599	2 (3.1%)	-1.602	.304	1.130	.599
		Y2	2 (3.1%)	-1.311	.306	0.475	.604	2 (3.1%)	-1.382	.304	0.220	.599
		Y4	2 (3.1%)	-2.023	.304	2.674	.599	2 (3.1%)	-1.513	.304	0.539	.599
		Y6	2 (3.1%)	-2.933	.304	7.398	.599	2 (3.1%)	-1.775	.304	1.969	.599
		Y8	2 (3.1%)	-1.220	.306	0.157	.604	2 (3.1%)	-1.068	.304	-0.322	.599
		Y10	3 (4.7%)	-0.720	.306	-1.105	.604	2 (3.1%)	-0.095	.304	-1.646	.599
	2	F	10 (15.6%)	-1.921	.325	2.581	.639	10 (15.6%)	-2.343	.325	5.137	.639
		Y2	10 (15.6%)	-3.192	.325	9.839	.639	10 (15.6%)	-2.603	.325	5.902	.639
		Y4	10 (15.6%)	-2.105	.325	3.060	.639	10 (15.6%)	-2.662	.325	6.634	.639
		Y6	10 (15.6%)	-0.736	.325	-1.049	.639	10 (15.6%)	-1.698	.325	1.970	.639
		Y8	10 (15.6%)	-1.217	.325	0.308	.639	10 (15.6%)	-1.199	.325	0.144	.639
		Y10	10 (15.6%)	-1.433	.325	0.693	.639	10 (15.6%)	-0.652	.325	-1.198	.639

Table K-2. (continued)

Treat-ment group	Time Point	Year level statement	Connecting Information					Creating New Ideas				
			Missing cases	Skew	Std. Error	Kurtosis	Std. Error	Missing cases	Skew	Std. Error	Kurtosis	Std. Error
Control *N* = 64	3	F	4 (6.3%)	-3.254	.309	9.719	.608	4 (6.3%)	-1.474	.309	0.750	.608
		Y2	5 (7.8%)	-1.629	.311	1.031	.613	4 (6.3%)	-2.155	.309	3.357	.608
		Y4	4 (6.3%)	-2.871	.309	7.023	.608	4 (6.3%)	-2.077	.309	2.843	.608
		Y6	4 (6.3%)	-2.686	.309	6.059	.608	4 (6.3%)	-1.261	.309	0.234	.608
		Y8	4 (6.3%)	-1.502	.309	0.700	.608	4 (6.3%)	-1.194	.309	0.093	.608
		Y10	4 (6.3%)	-1.191	.309	-0.121	.608	4 (6.3%)	-0.372	.309	-1.603	.608

Table K-3. *Study Three. Tests of normality for critical and creative thinking assessment, control group: Evaluating a Plan or Idea*

Treat-ment group	Time Point	Year level statement	Evaluating a Plan or Idea				
			Missing cases	Skew	Std. Error	Kurtosis	Std. Error
Control *N* = 64	1	F	2 (3.1%)	-1.908	.304	2.625	.599
		Y2	2 (3.1%)	-1.013	.304	-0.412	.599
		Y4	2 (3.1%)	-1.563	.304	1.154	.599
		Y6	2 (3.1%)	-2.040	.304	3.352	.599
		Y8	2 (3.1%)	-0.932	.304	-0.619	.599
		Y10	2 (3.1%)	-1.244	.304	0.222	.599
	2	F	10 (15.6%)	-1.058	.325	.088	.639
		Y2	10 (15.6%)	-1.653	.325	1.946	.639
		Y4	10 (15.6%)	-0.989	.325	-0.658	.639
		Y6	10 (15.6%)	-1.927	.325	3.021	.639
		Y8	10 (15.6%)	-0.971	.325	-0.591	.639
		Y10	10 (15.6%)	-1.377	.325	0.953	.639
	3	F	4 (6.3%)	-1.804	.309	1.718	.608
		Y2	4 (6.3%)	-1.649	.309	1.689	.608
		Y4	4 (6.3%)	-1.405	.309	0.146	.608
		Y6	4 (6.3%)	-2.434	.309	5.441	.608
		Y8	4 (6.3%)	-0.933	.309	-0.646	.608
		Y10	4 (6.3%)	-1.513	.309	0.986	.608

Table K-4. *Study Three. Tests of normality for critical and creative thinking assessment, theory group: Asking Questions and Finding and Organising Information*

Treat-ment group	Time Point	Year level statement	Asking Questions					Finding and Organising Information				
			Missing cases	Skew	Std. Error	Kurtosis	Std. Error	Missing cases	Skew	Std. Error	Kurtosis	Std. Error
Theory *N* = 106	1	F	5 (4.7%)	-2.519	.240	4.914	.476	5 (4.7%)	-1.270	.240	0.219	.476
		Y2	5 (4.7%)	-0.568	.240	-1.085	.476	5 (4.7%)	-0.932	.240	-0.661	.476
		Y4	5 (4.7%)	-3.241	.240	9.486	.476	5 (4.7%)	-1.050	.240	-0.630	.476
		Y6	5 (4.7%)	-1.616	.240	1.428	.476	5 (4.7%)	-1.206	.240	-0.097	.476
		Y8	5 (4.7%)	-3.850	.240	14.570	.476	5 (4.7%)	-0.619	.240	-1.061	.476
		Y10	5 (4.7%)	-0.931	.240	-0.453	.476	5 (4.7%)	-0.077	.240	-1.650	.476
	2	F	13 (12.3%)	-4.386	.250	19.644	.495	13 (12.3%)	-1.339	.250	0.684	.495
		Y2	13 (12.3%)	-1.278	.250	0.387	.495	13 (12.3%)	-0.710	.250	-0.758	.495
		Y4	13 (12.3%)	-2.773	.250	5.812	.495	13 (12.3%)	-1.894	.250	2.063	.495
		Y6	14 (13.2%)	-1.227	.251	-0.010	.498	13 (12.3%)	-1.804	.250	2.085	.495
		Y8	13 (12.3%)	-4.800	.250	23.440	.495	13 (12.3%)	-1.067	.250	-0.239	.495
		Y10	14 (13.2%)	-0.800	.251	-0.624	.498	13 (12.3%)	-0.707	.250	-1.075	.495

Table K-4. (continued)

Treatment group	Time Point	Year level statement	Asking Questions					Finding and Organising Information				
			Missing cases	Skew	Std. Error	Kurtosis	Std. Error	Missing cases	Skew	Std. Error	Kurtosis	Std. Error
Theory *N* = 106	3	F	17 (16.0%)	-5.428	.255	31.656	.506	17 (16.0%)	-1.069	.255	-0.136	.506
		Y2	17 (16.0%)	-1.377	.255	0.762	.506	17 (16.0%)	-0.541	.255	-1.311	.506
		Y4	17 (16.0%)	-2.514	.255	5.555	.506	17 (16.0%)	-1.107	.257	-0.400	.508
		Y6	17 (16.0%)	-0.939	.255	-0.411	.506	17 (16.0%)	-1.670	.255	1.610	.506
		Y8	17 (16.0%)	-3.161	.255	9.973	.506	18 (17.0%)	-0.808	.257	-0.765	.508
		Y10	18 (17.0%)	-1.156	.257	0.232	.508	18 (17.0%)	-0.717	.255	-1.057	.506

Table K-5. *Study Three. Tests of normality for critical and creative thinking assessment, theory group: Connecting Information and Creating New Ideas*

Treat-ment group	Time Point	Year level statement	Connecting Information					Creating New Ideas				
			Missing cases	Skew	Std. Error	Kurtosis	Std. Error	Missing cases	Skew	Std. Error	Kurtosis	Std. Error
Theory *N* = 106	1	F	5 (4.7%)	-3.241	.240	9.486	.476	5 (4.7%)	-1.936	.240	2.619	.476
		Y2	5 (4.7%)	-1.820	.240	1.992	.476	5 (4.7%)	-2.591	.240	5.832	.476
		Y4	5 (4.7%)	-1.805	.240	2.146	.476	5 (4.7%)	-2.189	.240	3.481	.476
		Y6	5 (4.7%)	-2.392	.240	3.796	.476	5 (4.7%)	-1.702	.240	1.427	.476
		Y8	5 (4.7%)	-1.117	.240	-0.091	.476	5 (4.7%)	-1.315	.240	0.661	.476
		Y10	5 (4.7%)	-1.464	.240	1.061	.476	5 (4.7%)	-0.240	.240	-1.428	.476
	2	F	13 (12.3%)	-2.401	.250	4.633	.495	13 (12.3%)	-2.739	.250	7.043	.495
		Y2	14 (13.2%)	-2.580	.251	6.147	.498	13 (12.3%)	-3.008	.250	8.316	.495
		Y4	13 (12.3%)	-2.350	.250	4.945	.495	13 (12.3%)	-2.214	.250	3.807	.495
		Y6	13 (12.3%)	-1.286	.250	0.481	.495	13 (12.3%)	-1.309	.250	0.385	.495
		Y8	13 (12.3%)	-0.921	.250	-0.695	.495	14 (13.2%)	-1.171	.251	0.149	.498
		Y10	14 (13.2%)	-2.046	.251	3.247	.498	14 (13.2%)	-0.489	.251	-1.294	.498

Table K-5. (continued)

Treat-ment group	Time Point	Year level statement	Connecting Information					Creating New Ideas				
			Missing cases	Skew	Std. Error	Kurtosis	Std. Error	Missing cases	Skew	Std. Error	Kurtosis	Std. Error
Theory *N* = 106	3	F	17 (16.0%)	-2.992	.255	7.848	.506	17 (16.0%)	-2.659	.255	6.565	.506
		Y2	17 (16.0%)	-2.055	.255	3.116	.506	17 (16.0%)	-1.980	.255	2.553	.506
		Y4	17 (16.0%)	-1.684	.255	1.866	.506	17 (16.0%)	-2.021	.255	2.839	.506
		Y6	17 (16.0%)	-1.831	.255	2.389	.506	17 (16.0%)	-1.173	.255	0.204	.506
		Y8	17 (16.0%)	-1.406	.255	0.506	.506	18 (17.0%)	-1.214	.257	-0.139	.508
		Y10	18 (17.0%)	-1.225	.257	0.161	.508	17 (16.0%)	-0.498	.255	-1.225	.506

Table K-6. *Study Three. Tests of normality for critical and creative thinking assessment, theory group: Evaluating a Plan or Idea*

Treat-ment group	Time Point	Year level statement	Evaluating a Plan or Idea				
			Missing cases	Skew	Std. Error	Kurtosis	Std. Error
Theory *N* = 106	1	F	5 (4.7%)	-1.599	.240	1.291	.476
		Y2	6 (5.7%)	-2.338	.241	4.413	.478
		Y4	6 (5.7%)	-1.438	.241	0.537	.478
		Y6	5 (4.7%)	-2.591	.240	5.832	.476
		Y8	6 (5.7%)	-1.120	.241	-0.268	.478
		Y10	6 (5.7%)	-1.804	.241	1.930	.478
	2	F	14 (13.2%)	-1.950	.251	2.730	.498
		Y2	14 (13.2%)	-1.826	.251	2.040	.498
		Y4	13 (12.3%)	-1.616	.250	1.656	.495
		Y6	13 (12.3%)	-2.720	.250	6.678	.495
		Y8	13 (12.3%)	-0.810	.250	-0.774	.495
		Y10	14 (13.2%)	-1.462	.251	0.384	.498
	3	F	17 (16.0%)	-1.521	.255	0.718	.506
		Y2	17 (16.0%)	-1.621	.255	1.259	.506
		Y4	17 (16.0%)	-1.369	.255	0.658	.506
		Y6	17 (16.0%)	-3.880	.255	15.963	.506
		Y8	18 (17.0%)	-1.061	.257	-0.291	.508
		Y10	17 (16.0%)	-1.136	.255	-0.214	.506

Table K-7. *Study Three. Tests of normality for critical and creative thinking assessment, industry partner group: Asking Questions and Finding and Organising Information*

Treat-ment group	Time Point	Year level statement	Asking Questions					Finding and Organising Information				
			Missing cases	Skew	Std. Error	Kurtosis	Std. Error	Missing cases	Skew	Std. Error	Kurtosis	Std. Error
Industry Partner *N* = 87	1	F	6 (6.9%)	-3.673	.267	14.199	.529	9 (10.3%)	-0.256	.272	-1.325	.538
		Y2	6 (6.9%)	-0.509	.267	-0.721	.529	8 (9.2%)	-0.328	.271	-1.016	.535
		Y4	6 (6.9%)	-1.946	.267	2.480	.529	7 (8.0%)	-0.731	.269	-1.091	.532
		Y6	7 (8.0%)	-1.728	.269	1.842	.532	7 (8.0%)	-1.634	.269	1.426	.532
		Y8	6 (6.9%)	-2.795	.267	7.276	.529	6 (6.9%)	-0.820	.267	-0.693	.529
		Y10	6 (6.9%)	-0.508	.267	-1.263	.529	6 (6.9%)	-0.597	.267	-1.058	.529
	2	F	18 (20.7%)	-3.702	.289	14.435	.570	18 (20.7%)	-2.088	.289	3.605	.570
		Y2	18 (20.7%)	-0.320	.289	-1.300	.570	18 (20.7%)	-0.921	.289	-0.468	.570
		Y4	18 (20.7%)	-2.325	.289	4.448	.570	18 (20.7%)	-1.230	.289	0.072	.570
		Y6	18 (20.7%)	-1.108	.289	0.100	.570	18 (20.7%)	-2.030	.289	3.455	.570
		Y8	18 (20.7%)	-2.945	.289	7.602	.570	19 (21.8%)	-1.079	.291	0.117	.574
		Y10	18 (20.7%)	-0.447	.289	-0.882	.570	18 (20.7%)	-0.828	.289	-0.587	.570

Table K-7. (continued)

Treat-ment group	Time Point	Year level statement	Asking Questions					Finding and Organising Information				
			Missing cases	Skew	Std. Error	Kurtosis	Std. Error	Missing cases	Skew	Std. Error	Kurtosis	Std. Error
Industry Partner *N* = 87	3	F	10 (11.5%)	-4.860	.274	22.200	.541	9 (10.3%)	-1.029	.272	-0.573	.538
		Y2	10 (11.5%)	-0.542	.274	-0.799	.541	9 (10.3%)	-0.910	.272	-0.744	.538
		Y4	10 (11.5%)	-3.273	.274	10.150	.541	9 (10.3%)	-0.929	.276	-0.864	.545
		Y6	10 (11.5%)	-1.444	.274	1.142	.541	10 (11.5%)	-1.865	.274	2.148	.541
		Y8	10 (11.5%)	-3.510	.274	11.722	.541	9 (10.3%)	-1.125	.272	0.112	.538
		Y10	10 (11.5%)	-0.402	.274	-0.653	.541	9 (10.3%)	-0.777	.272	-0.701	.538

Table K-8. *Study Three. Tests of normality for critical and creative thinking assessment, industry partner group: Connecting Information and Creating New Ideas*

Treat-ment group	Time Point	Year level statement	Connecting Information					Creating New Ideas				
			Missing cases	Skew	Std. Error	Kurtosis	Std. Error	Missing cases	Skew	Std. Error	Kurtosis	Std. Error
Industry Partner *N* = 87	1	F	8 (9.2%)	-2.447	.271	5.372	.535	6 (6.9%)	-0.747	.267	-0.768	.529
		Y2	6 (6.9%)	-0.980	.267	-0.276	.529	6 (6.9%)	-1.852	.267	1.897	.529
		Y4	8 (9.2%)	-2.167	.271	3.645	.535	6 (6.9%)	-1.736	.267	1.421	.529
		Y6	7 (8.0%)	-1.912	.269	2.773	.532	6 (6.9%)	-0.729	.267	-0.969	.529
		Y8	7 (8.0%)	-0.709	.269	-0.996	.532	6 (6.9%)	-0.697	.267	-0.873	.529
		Y10	6 (6.9%)	-0.381	.267	-1.383	.529	6 (6.9%)	-0.373	.267	-1.327	.529
	2	F	18 (20.7%)	-3.033	.289	9.353	.570	18 (20.7%)	-2.088	.289	3.605	.570
		Y2	18 (20.7%)	-2.854	.289	7.483	.570	18 (20.7%)	-3.372	.289	9.649	.570
		Y4	18 (20.7%)	-1.771	.289	2.175	.570	18 (20.7%)	-1.877	.289	1.950	.570
		Y6	18 (20.7%)	-1.168	.289	0.223	.570	18 (20.7%)	-0.971	.289	-0.556	.570
		Y8	19 (21.8%)	-1.392	.293	0.847	.578	19 (21.8%)	-1.015	.291	-0.252	.574
		Y10	18 (20.7%)	-0.895	.291	-0.463	.574	18 (20.7%)	-0.238	.289	-0.752	.570

Table K-8. (continued)

Treat-ment group	Time Point	Year level statement	Connecting Information					Creating New Ideas				
			Missing cases	Skew	Std. Error	Kurtosis	Std. Error	Missing cases	Skew	Std. Error	Kurtosis	Std. Error
Industry Partner *N* = 87	3	F	9 (10.3%)	-2.679	.272	6.298	.538	9 (10.3%)	-1.165	.272	0.341	.538
		Y2	9 (10.3%)	-1.625	.272	1.609	.538	10 (11.5%)	-1.761	.274	2.060	.541
		Y4	9 (10.3%)	-1.954	.272	2.775	.538	10 (11.5%)	-1.169	.274	-0.277	.541
		Y6	9 (10.3%)	-3.000	.272	8.063	.538	10 (11.5%)	-1.023	.274	-0.460	.541
		Y8	10 (11.5%)	-1.269	.274	0.415	.541	9 (10.3%)	-0.397	.272	-1.057	.538
		Y10	9 (10.3%)	-1.481	.272	1.165	.538	9 (10.3%)	-0.323	.272	-1.471	.538

Table K-9. *Study Three. Tests of normality for critical and creative thinking assessment, industry partner group: Evaluating a Plan or Idea*

Treat-ment group	Time Point	Year level statement	Evaluating a Plan or Idea				
			Missing cases	Skew	Std. Error	Kurtosis	Std. Error
Industry Partner *N* = 87	1	F	7 (8.0%)	-1.372	.269	0.419	.532
		Y2	6 (6.9%)	-1.487	.267	0.465	.529
		Y4	6 (6.9%)	-0.700	.267	-1.209	.529
		Y6	7 (8.0%)	-2.772	.269	7.136	.532
		Y8	7 (8.0%)	-0.601	.269	-1.183	.532
		Y10	7 (8.0%)	-1.150	.269	-0.135	.532
	2	F	18 (20.7%)	-1.342	.289	0.521	.570
		Y2	18 (20.7%)	-1.858	.289	2.647	.570
		Y4	18 (20.7%)	-0.771	.289	-0.875	.570
		Y6	19 (21.8%)	-1.866	.291	2.705	.574
		Y8	18 (20.7%)	-0.760	.289	-0.392	.570
		Y10	19 (21.8%)	-1.178	.291	0.047	.574
	3	F	9 (10.3%)	-1.158	.272	-0.142	.538
		Y2	9 (10.3%)	-2.301	.272	4.016	.538
		Y4	9 (10.3%)	-0.750	.272	-1.042	.538
		Y6	9 (10.3%)	-2.066	.272	3.389	.538
		Y8	9 (10.3%)	-0.856	.272	-0.613	.538
		Y10	9 (10.3%)	-0.732	.272	-1.167	.538

Table K-10. *Study Three. Tests of normality for motivation and engagement scale: Total positive motivation and total positive engagement*

Treatment group	Time Point	Total Positive Motivation					Total Positive Engagement				
		Missing cases	Skewness	Std. Error	Kurtosis	Std. Error	Missing cases	Skewness	Std. Error	Kurtosis	Std. Error
Control *N* = 64	1	11 (17.2%)	-0.511	.327	-0.591	.644	7 (10.9%)	-0.443	.316	0.797	.623
	2	10 (15.6%)	-0.779	.325	-0.095	.639	12 (18.8%)	0.029	.330	-0.830	.650
	3	4 (6.3%)	-1.091	.309	1.192	.608	6 (9.4%)	-0.404	.314	0.253	.618
Theory *N* = 106	1	9 (8.5%)	-0.790	.245	0.611	.485	15 (14.2%)	-0.474	.253	0.550	.500
	2	18 (17.0%)	-0.525	.257	-0.659	.508	18 (17.0%)	-0.037	.257	-0.995	.508
	3	21 (19.8%)	-1.214	.261	2.095	.517	21 (19.8%)	-0.227	.261	-0.448	.517
Industry Partner *N* = 87	1	14 (16.1%)	-1.041	.281	0.983	.555	9 (10.3%)	-0.216	.272	0.084	.538
	2	22 (25.3%)	-1.345	.297	2.047	.586	21 (24.1%)	-0.538	.295	0.257	.582
	3	11 (12.6%)	-0.393	.276	-0.115	.545	12 (13.8%)	-0.239	.277	-0.346	.548

Table K-11. *Study Three. Tests of normality for motivation and engagement scale: Total negative motivation and total negative engagement*

Treatment group	Time Point	Total Negative Motivation					Total Negative Engagement				
		Missing cases	Skewness	Std. Error	Kurtosis	Std. Error	Missing cases	Skewness	Std. Error	Kurtosis	Std. Error
Control *N* = 64	1	6 (9.4%)	-0.220	.314	-0.452	.618	7 (10.9%)	1.318	.316	2.014	.623
	2	13 (20.3%)	-0.262	.333	-0.673	.656	13 (20.3%)	0.421	.333	-0.872	.656
	3	10 (15.6%)	-0.036	.325	-0.945	.639	4 (6.3%)	1.325	.309	2.283	.608
Theory *N* = 106	1	14 (13.2%)	-0.247	.251	-0.497	.498	10 (9.4%)	0.766	.246	0.068	.488
	2	16 (15.1%)	-0.072	.254	-0.759	.503	16 (15.1%)	0.512	.254	-0.878	.503
	3	22 (20.8%)	0.144	.263	-0.839	.520	22 (20.8%)	0.750	.263	0.470	.520
Industry Partner *N* = 87	1	11 (12.6%)	0.435	.276	-0.101	.545	10 (11.5%)	0.581	.274	-0.074	.541
	2	23 (26.4%)	0.442	.299	0.571	.590	24 (27.6%)	1.223	.302	1.888	.595
	3	13 (14.9%)	1.040	.279	1.795	.552	13 (14.9%)	1.426	.279	2.130	.552

Table K-12. *Study Three. Motivation and engagement scale analysis for internal consistency – time points compared, all schools combined*

Higher Order Factor	First Order Factor	Cronbach's alpha reported^	Time Point 1		Time Point 2		Time Point 3	
			Cronbach's alpha calculated	If item (*x*) deleted: *	Cronbach's alpha calculated	If item (*x*) deleted:*	Cronbach's alpha calculated	If item (*x*) deleted:*
Positive Motivation	Self-Belief	.81	.69		.78		.80	
	Valuing	.78	.68	(Q1): .69	.76		.74	(Q1): .75
	Learning focus	.84	.69		.76		.77	
Positive Engagement	Planning	.85	.78		.85		.85	
	Task management	.86	.80		.86		.86	
	Persistence	.79	.69		.74		.76	
Negative Motivation	Anxiety	.72	.75		.77		.80	
	Failure avoidance	.87	.85		.87		.86	
	Uncertain control	.78	.75		.78		.81	
Negative Engagement	Self-sabotage	.80	.80		.83		.82	
	Disengagement	.72	.66	(Q3): .73	.67	(Q3): .69	.75	(Q3): .75

Note. ^ Reported in Martin, A.J. (2022). *Motivation and Engagement Scale (MES) Test User Manual. Sydney, Australia: Lifelong Achievement Group.*
* Value given if deleting the item would raise the value of Cronbach's alpha. *N* = 257

Table K-13. *Study Three. Motivation and engagement scale analysis for internal consistency – schools compared, Time Point 1*

First Order Factor		School A		School B		School C		School D	
	n =	99		45		23		87	
	Cronbach's alpha reported ^	Cronbach's alpha calculated	If item (*x*) deleted: *	Cronbach's alpha calculated	If item (*x*) deleted: *	Cronbach's alpha calculated	If item (*x*) deleted: *	Cronbach's alpha calculated	If item (*x*) deleted: *
Self-Belief	.81	.68	(Q3): .70	.60		.88		.66	
Valuing	.78	.62		.57	(Q1): .78	.83	(Q1): .85	.74	
Learning focus	.84	.57	(Q1): .58	.57	(Q4): .66	.84		.70	
Planning	.85	.76		.72	(Q1): .75	.89		.80	
Task management	.86	.80	(Q2): .81	.80	(Q2): .88	.83	(Q1): .88	.80	
Persistence	.79	.66		.65	(Q2): .72	.69		.75	
Anxiety	.72	.74	(Q3): .76	.72		.68	(Q3): .78	.79	
Failure avoidance	.87	.83	(Q3): .84	.81		.91		.81	
Uncertain control	.78	.78		.72	(Q1): .73	.57	(Q1): .68	.75	
Self-sabotage	.80	.82		.74		.94		.66	(Q4): .74
Disengagement	.72	.72	(Q3): .74	.35	(Q3): .73	.78	(Q3): .90	.61	(Q3): .64

Note. ^ Reported in Martin, A.J. (2022). *Motivation and Engagement Scale (MES) Test User Manual.* Sydney, Australia: Lifelong Achievement Group
* Value given if deleting the item would raise the value of Cronbach's alpha.

Table K-14. *Study Three. Motivation and engagement scale analysis for internal consistency – schools compared, Time Point 2*

First Order Factor		School A		School B		School C		School D	
	n =	99		45		23		87	
	Cronbach's alpha reported ^	Cronbach's alpha calculated	If item (*x*) deleted: *	Cronbach's alpha calculated	If item (*x*) deleted: *	Cronbach's alpha calculated	If item (*x*) deleted: *	Cronbach's alpha calculated	If item (*x*) deleted: *
Self-Belief	.81	.72		.79		.90		.80	
Valuing	.78	.68		.78		.59	(Q3): .79	.86	
Learning focus	.84	.57		.78		.90		.78	
Planning	.85	.84		.83	(Q4): .88	.94		.84	
Task management	.86	.82		.82	(Q1): .84	.93		.86	(Q2): .87
Persistence	.79	.79		.74	(Q2): .77	.60	(Q1): .63	.70	(Q3): .79
Anxiety	.72	.78		.81		.62		.78	(Q3): .80
Failure avoidance	.87	.88		.87		.86	(Q3): .88	.81	
Uncertain control	.78	.80		.77	(Q4): .79	.68	(Q4): .74	.77	
Self-sabotage	.80	.81	(Q1): .84	.74	(Q4): .80	.84	(Q1): .88	.85	
Disengagement	.72	.64	(Q3): .74	.71	(Q1): .75	.83	(Q3): .87	.52	(Q1): .64

Note. ^ Reported in Martin, A.J. (2022). *Motivation and Engagement Scale (MES) Test User Manual.* Sydney, Australia: Lifelong Achievement Group
* Value given if deleting the item would raise the value of Cronbach's alpha.

Table K-15. *Study Three. Motivation and engagement scale analysis for internal consistency – schools compared, Time Point 3*

First Order Factor		School A		School B		School C		School D	
	n =	99		48		23		87	
	Cronbach's alpha reported ^	Cronbach's alpha calculated	If item (*x*) deleted: *	Cronbach's alpha calculated	If item (*x*) deleted: *	Cronbach's alpha calculated	If item (*x*) deleted: *	Cronbach's alpha calculated	If item (*x*) deleted: *
Self-Belief	.81	.84		.79		.80	(Q1): .88	.68	
Valuing	.78	.81	(Q1): .83	.69	(Q2): .71	.56	(Q3): .62	.68	
Learning focus	.84	.80		.80		.61		.70	(Q2): .73
Planning	.85	.85		.77	(Q2): .81	.90	(Q2): .91	.86	
Task management	.86	.85		.78		.97		.86	
Persistence	.79	.74	(Q1): .78	.71	(Q1): .75	.63	(Q1): .70	.83	
Anxiety	.72	.78		.80		.89		.76	
Failure avoidance	.87	.86	(Q3): .90	.74		.95		.83	(Q3): .84
Uncertain control	.78	.81		.65	(Q1): .72	.86	(Q1): .89	.80	(Q4): .81
Self-sabotage	.80	.85		.78		.77		.74	(Q3): .75
Disengagement	.72	.78		.74	(Q3): .78	.74	(Q3): .82	.73	

Note. ^ Reported in Martin, A.J. (2022). *Motivation and Engagement Scale (MES) Test User Manual.* Sydney, Australia: Lifelong Achievement Group
* Value given if deleting the item would raise the value of Cronbach's alpha.

Table K-16. *Study Three. Motivation and engagement scale analysis for internal consistency – EALD/ Non-EALD students compared – Time Point 1*

First Order Factor		EALD student		Non-EALD student	
	N =	86		168	
	Cronbach's alpha reported ^	Cronbach's alpha calculated	If item (*x*) deleted: *	Cronbach's alpha calculated	If item (*x*) deleted: *
Self-Belief	.81	.68	(Q3): .69	.70	
Valuing	.78	.73		.64	(Q1): .68
Learning focus	.84	.53	(Q2): .54	.74	
Planning	.85	.76		.80	
Task management	.86	.77	(Q2): .78	.82	
Persistence	.79	.67		.69	
Anxiety	.72	.75	(Q3): .79	.75	
Failure avoidance	.87	.85		.84	
Uncertain control	.78	.73		.75	
Self-sabotage	.80	.83		.79	
Disengagement	.72	.66	(Q3): .71	.65	(Q3): .75

Note. ^ Reported in Martin, A.J. (2022). *Motivation and Engagement Scale (MES) Test User Manual.* Sydney, Australia: Lifelong Achievement Group. * Value given if deleting the item would raise the value of Cronbach's alpha.

Table K-17. *Study Three. Motivation and engagement scale analysis for internal consistency – EALD/ Non-EALD students compared – Time Point 2*

First Order Factor		EALD student		Non-EALD student	
	N =	86		168	
	Cronbach's alpha reported ^	Cronbach's alpha calculated	If item (*x*) deleted: *	Cronbach's alpha calculated	If item (*x*) deleted: *
Self-Belief	.81	.74		.80	
Valuing	.78	.73		.78	
Learning focus	.84	.62	(Q2): .68	.79	
Planning	.85	.89		.82	
Task management	.86	.86		.85	
Persistence	.79	.78		.71	
Anxiety	.72	.78		.77	
Failure avoidance	.87	.87		.87	
Uncertain control	.78	.80		.76	
Self-sabotage	.80	.79	(Q1): .85	.85	
Disengagement	.72	.61	(Q3): .76	.69	(Q1): .72

Note. ^ Reported in Martin, A.J. (2022). *Motivation and Engagement Scale (MES) Test User Manual*. Sydney, Australia: Lifelong Achievement Group. * Value given if deleting the item would raise the value of Cronbach's alpha.

Table K-18. *Study Three. Motivation and engagement scale analysis for internal consistency – EALD/ Non-EALD students compared – Time Point 3*

First Order Factor		EALD student		Non-EALD student	
	N =	86		171	
	Cronbach's alpha reported ^	Cronbach's alpha calculated	If item (*x*) deleted: *	Cronbach's alpha calculated	If item (*x*) deleted: *
Self-Belief	.81	.82		.78	
Valuing	.78	.80		.69	
Learning focus	.84	.77		.76	
Planning	.85	.87		.82	
Task management	.86	.85		.86	
Persistence	.79	.66	(Q1): .69	.81	(Q2): .82
Anxiety	.72	.79		.78	
Failure avoidance	.87	.86	(Q3): .88	.85	
Uncertain control	.78	.81		.78	
Self-sabotage	.80	.82		.80	
Disengagement	.72	.80		.72	(Q3): .74

Note. ^ Reported in Martin, A.J. (2022). *Motivation and Engagement Scale (MES) Test User Manual.* Sydney, Australia: Lifelong Achievement Group. * Value given if deleting the item would raise the value of Cronbach's alpha

Table K-19. *Study Three. Cochran's Q analysis of critical and creative thinking – control group*

Sub-element	Year level statement	*p*	Time Point 1		Time Point 2		Time point 3	
			% 'Yes'	Missing cases	% 'Yes'	Missing cases	% 'Yes'	Missing cases
Asking Questions	R	.497	93.8	2	75.0	10	84.4	4
	Y2	.037 [a]	46.9	2	62.5	10	59.4	4
	Y4	.196	82.8	2	68.8	10	84.4	4
	Y6	.961	65.6	2	57.8	10	57.8	4
	Y8	.368	89.1	2	76.6	10	87.5	4
	Y10	.965	51.6	2	48.4	10	54.7	4
Finding and Organising Information	R	.080	53.1	2	46.9	10	60.9	4
	Y2	.891	62.5	2	57.8	11	60.9	4
	Y4	.304	60.9	3	62.5	10	75.0	4
	Y6	1.000	73.4	3	65.6	10	71.9	4
	Y8	.393	53.1	2	50.0	10	59.4	4
	Y10	.082	39.1	3	43.8	10	57.8	4
Connecting Information	R	.074	78.1	2	65.6	10	84.4	4
	Y2	.022 [a]	64.1	3	75.0	10	70.3	5
	Y4	.494	78.1	2	68.8	10	82.8	4
	Y6	<.001 [b c]	85.9	2	46.9	10	81.3	4
	Y8	.163	62.5	2	54.7	10	68.8	4
	Y10	.167	53.1	2	59.4	10	62.5	4

Note. On at least one case, the value of the weight variable was zero, negative, or missing. Such cases are invisible to statistical procedures and graphs which need positively weighted cases but remain on the file and are processed by non-statistical facilities such as LIST and SAVE.
[a] $p < .050$ TP 1 to TP 2. [b] $p < .001$ TP 1 to TP 2. [c] $p < .001$ TP 2 to TP 3.

Table K-19. (continued)

Sub-element	Year level statement	*p*	Time Point 1		Time Point 2		Time point 3	
			% 'Yes'	Missing cases	% 'Yes'	Missing cases	% 'Yes'	Missing cases
Creating New Ideas	R	.204	71.9	2	70.3	10	67.2	4
	Y2	.589	70.3	2	71.9	10	76.6	4
	Y4	.504	73.4	2	71.9	10	82.8	4
	Y6	.951	73.4	2	62.5	10	62.5	4
	Y8	.965	60.9	2	54.7	10	60.9	4
	Y10	.353	37.5	2	45.3	10	45.3	4
Evaluating a Plan	R	.055	75.0	2	51.6	10	73.4	4
	Y2	.115	59.4	2	62.5	10	68.8	4
	Y4	.229	70.3	2	53.1	10	70.3	4
	Y6	.504	76.6	2	65.6	10	78.1	4
	Y8	.902	57.8	2	51.6	10	56.3	4
	Y10	.519	64.1	2	57.8	10	67.2	4

N = 64

Note. On at least one case, the value of the weight variable was zero, negative, or missing. Such cases are invisible to statistical procedures and graphs which need positively weighted cases but remain on the file and are processed by non-statistical facilities such as LIST and SAVE.
[a] $p < .050$ TP 1 to TP 2. [b] $p < .001$ TP 1 to TP 2. [c] $p < .001$ TP 2 to TP 3.

Table K-20. *Study Three. Cochran's Q analysis of critical and creative thinking – PBL group*

Sub-element	Year level statement	*p*	Time point 1		Time point 2		Time point 3	
			% 'Yes'	Missing cases	% 'Yes'	Missing cases	% 'Yes'	Missing cases
Asking Questions	R	.030 [a]	83.4	11	77.7	31	82.4	27
	Y2	.007 [b]	44.0	11	46.6	31	49.7	27
	Y4	.542	80.3	11	73.1	31	74.6	27
	Y6	.109	69.4	12	54.4	32	54.9	27
	Y8	.764	84.5	11	77.2	31	77.2	27
	Y10	.429	50.8	11	40.9	32	45.6	28
Finding and Organising Information	R	<.001 [c]	50.8	14	61.1	31	54.4	26
	Y2	.726	47.2	13	45.6	31	47.7	26
	Y4	.028 [d]	58.0	12	62.2	31	54.9	29
	Y6	.128	65.8	12	65.8	31	65.3	27
	Y8	.084	49.7	11	51.8	32	50.8	27
	Y10	.019 [a]	40.9	11	46.1	31	47.2	26
Connecting Information	R	.490	81.3	13	72.0	31	75.6	26
	Y2	.011 [e]	65.3	11	71.5	32	66.3	26
	Y4	.717	73.1	13	66.8	31	65.8	26
	Y6	<.001 [c f]	78.2	12	54.9	31	71.0	26
	Y8	.237	55.4	12	52.8	33	59.1	27
	Y10	.119	55.4	11	60.6	33	58.5	27

Note. On at least one case, the value of the weight variable was zero, negative, or missing. Such cases are invisible to statistical procedures and graphs which need positively weighted cases but remain on the file and are processed by non-statistical facilities such as LIST and SAVE.
[a] $p < .050$ TP1 to TP3. [b] $p < .010$ TP1 to TP3. [c] $p < .001$ TP1 to TP2. [d] $p < .050$ TP1 to TP2. [e] $p < .010$ TP1 to TP2. [f] $p < .001$ TP2 to TP3.

Table K-20. (continued)

Sub-element	Year level statement	*p*	Time point 1		Time point 2		Time point 3	
			% 'Yes'	Missing cases	% 'Yes'	Missing cases	% 'Yes'	Missing cases
Creating New Ideas	R	.050 [d]	63.2	11	69.9	31	65.3	26
	Y2	.021 [g]	77.7	11	75.6	31	66.8	27
	Y4	.721	75.6	11	67.9	31	65.8	26
	Y6	.946	62.7	11	54.4	31	54.4	27
	Y8	.559	57.0	11	51.8	33	47.7	27
	Y10	.229	38.9	11	34.7	32	39.4	26
Evaluating a Plan	R	.646	67.4	12	61.7	32	60.6	26
	Y2	.787	74.6	12	64.2	32	67.9	26
	Y4	.983	61.1	12	54.9	31	54.4	26
	Y6	.832	80.3	12	69.4	32	74.6	26
	Y8	.729	54.9	13	46.1	31	51.3	27
	Y10	.011	66.8	13	58.5	33	53.4	26
N = 136								

Note. On at least one case, the value of the weight variable was zero, negative, or missing. Such cases are invisible to statistical procedures and graphs which need positively weighted cases, but remain on the file and are processed by non-statistical facilities such as LIST and SAVE.
[d] $p < .050$ TP1 to TP2. [g] $p < .050$ TP2 to TP3.

Table K-21. *Study Three. Cochran's Q analysis of critical and creative thinking – theory group*

Sub-element	Year level statement	*p*	% 'Yes' Time point 1	Missing cases	% 'Yes' Time point 2	Missing cases	% 'Yes' Time point 3	Missing cases
AQ	R	.121	82.1	5	82.1	13	80.2	17
	Y2	<.001 [a b]	46.2	5	58.5	13	57.5	17
	Y4	.465	85.8	5	79.2	13	70.8	17
	Y6	.073	69.8	5	58.5	14	49.1	17
	Y8	.257	87.7	5	83.0	13	74.5	17
	Y10	.414	55.7	5	47.2	14	52.8	18
FO	R	.159	64.2	5	59.4	13	51.9	17
	Y2	.248	57.5	5	45.3	13	42.5	17
	Y4	.025 [e]	48.1	5	54.7	13	46.2	18
	Y6	.272	64.2	5	67.0	13	62.3	17
	Y8	.099	48.1	5	54.7	13	46.2	18
	Y10	.015 [d]	36.8	5	48.1	13	46.2	17
CI	R	.128	85.8	5	73.6	13	74.5	17
	Y2	.875	73.6	5	73.6	14	67.0	17
	Y4	.279	72.6	5	72.6	13	62.3	17
	Y6	<.001	84.0	5	58.5 [c]	13	64.2	17
	Y8	.164	60.4	5	52.8	13	59.4	17
	Y10	.107	67.0	5	68.9	14	54.7	18
CR	R	.727	74.5	5	75.5	13	71.7	17
	Y2	.260	81.1	5	77.4	13	67.0	17
	Y4	.197	78.3	5	71.7	13	62.3	17
	Y6	.358	72.6	5	59.4	13	53.8	17
	Y8	.799	64.2	5	55.7	14	56.6	18
	Y10	.120	37.7	5	41.5	14	39.6	17
EP	R	.459	69.8	5	67.9	14	62.3	17
	Y2	.396	78.3	6	67.0	14	62.3	17
	Y4	.971	67.9	6	64.2	13	57.5	17
	Y6	.486	81.1	5	75.5	13	77.4	17
	Y8	.673	61.3	6	49.1	13	51.9	18
	Y10	.026 [d]	72.6	6	65.1	14	54.7	17

N = 106

Note. On at least one case, the value of the weight variable was zero, negative, or missing. Such cases are invisible to statistical procedures and graphs which need positively weighted cases, but remain on the file and are processed by non-statistical facilities such as LIST and SAVE.
[a] $p < .050$ TP1 to TP2. [b] $p < .001$ TP2 to TP3. [c] $p < .001$ TP1 to TP2. [d] < .050 TP1 to TP3. [e] $p = .056$ TP1 to TP2 and TP1 to TP3

Table K-22. *Study Three. Cochran's Q analysis of critical and creative thinking – industry partner group*

Sub-element	Year level statement	*p*	% 'Yes' Time point 1	Missing cases	% 'Yes' Time point 2	Missing cases	% 'Yes' Time point 3	Missing cases
AQ	R	.169	85.1	6	72.4	18	85.1	10
	Y2	.629	41.4	6	32.2	18	40.2	10
	Y4	.013 [a]	73.6	6	65.5	18	79.3	10
	Y6	.187	69.0	7	49.4	18	62.1	10
	Y8	.717	80.5	6	70.1	18	80.5	10
	Y10	.875	44.8	6	33.3	18	36.8	10
FO	R	<.001 [e f]	34.5	9	63.2	18	57.5	9
	Y2	.003 [b c]	34.5	8	46.0	18	54.0	9
	Y4	.368	51.7	7	52.9	18	55.2	11
	Y6	.453	67.8	7	64.4	18	69.0	10
	Y8	.598	51.7	6	48.3	19	56.3	9
	Y10	.618	46.0	6	43.7	18	48.3	9
CI	R	.779	75.9	8	70.1	18	77.0	9
	Y2	.001 [b]	55.2	6	69.0	18	65.5	9
	Y4	.504	73.6	8	59.8	18	70.1	9
	Y6	<.001 [c d]	71.3	7	50.6	18	79.3	9
	Y8	.026 [c]	49.4	7	52.9	20	58.6	10
	Y10	.002 [e]	41.4	6	50.6	19	63.2	9
CR	R	.015 [c]	49.4	6	63.2	18	57.5	9
	Y2	.055	73.6	6	73.6	18	66.7	10
	Y4	.707	72.4	6	63.2	18	70.1	9
	Y6	.572	50.6	6	48.3	18	55.2	10
	Y8	.107	48.3	6	47.1	19	36.8	9
	Y10	.679	40.2	6	26.4	18	39.1	9
EP	R	.368	64.4	7	54.0	18	58.6	9
	Y2	.738	70.1	6	60.9	18	74.7	9
	Y4	.852	52.9	6	43.7	18	50.6	9
	Y6	.549	79.3	7	62.1	19	71.3	9
	Y8	.784	47.1	7	42.5	18	50.6	9
	Y10	.296	59.8	7	50.6	19	51.7	9

N = 87

Note. On at least one case, the value of the weight variable was zero, negative, or missing. Such cases are invisible to statistical procedures and graphs which need positively weighted cases, but remain on the file and are processed by non-statistical facilities such as LIST and SAVE.
[a]$p < .050$ TP 1 to TP 3. [b]$p < .010$ TP 1 to TP 2. [c]$p < .050$ TP 1 to TP 2. [d]$p < .010$ TP 2 to TP 3. [e]$p < .010$ TP 1 to TP 3. [f]$p < .001$ TP 1 to TP 2.

Table K-23. *Study Three. Friedman test analysis of motivation and engagement survey – control group and PBL group*

Treatment Group	Higher Order Factor	*p*	$\chi^2(2)$	Median Time point 1	Missing cases	Median Time point 2	Missing cases	Median Time point 3	Missing cases
Control *N* = 64	Positive Motivation	.160	3.669	4.4167	11	4.5000	10	4.3750	4
	Positive Engagement	.018	7.988	3.6667	7	4.0000	12	3.8750	6
	Negative Motivation	.536	1.248	3.2500	6	3.1667	13	3.0417	10
	Negative Engagement	.184	3.386	1.8571	7	1.5714	11	1.5714	4
PBL *N* = 193	Positive Motivation	.002	12.823	4.4167	23	4.4167	40	4.2500	32
	Positive Engagement	.010	9.166	3.7500	24	3.7917	39	3.6667	33
	Negative Motivation	.001	13.624	2.7500	25	2.6667	39	2.4167	35
	Negative Engagement	.010	9.131	1.7143	20	1.6429	39	1.5714	35

Table K-24. *Study Three. Friedman test analysis of motivation and engagement survey – theory group and industry partner group*

Treatment Group	Higher Order Factor	*p*	$\chi^2(2)$	Median Time point 1	Missing cases	Median Time point 2	Missing cases	Median Time point 3	Missing cases
Theory *N* = 106	Positive Motivation	.016	8.255	4.3333	9	4.4167	18	4.2500	21
	Positive Engagement	.025	7.350	3.8333	15	3.5833	18	3.6667	21
	Negative Motivation	.008	9.595	2.9167	14	2.7917	16	2.7083	22
	Negative Engagement	.130	4.087	1.7143	10	1.7143	16	1.7857	22
Industry Partner *N* = 87	Positive Motivation	.081	5.026	4.4167	14	4.4167	22	4.2083	11
	Positive Engagement	.087	4.874	3.7083	9	3.8333	21	3.6667	12
	Negative Motivation	.104	4.519	2.5000	11	2.5000	23	2.3750	13
	Negative Engagement	.062	5.550	1.7143	10	1.5714	23	1.4286	13