Effects of instructional conditions and experience on student reflection: a video annotation study

Negin Mirriahi[a*], Srećko Joksimović[a], Dragan Gašević[b], Shane Dawson[a]

[a]*Teaching Innovation Unit, University of South Australia,*
160 Currie St, Adelaide, SA, 5000, Australia

[b]*Faculty of Education, Monash University,*
19 Ancora Imparo Way, Clayton, Victoria, 3800, Australia

**Contact. Negin Mirriahi: Email: negin.mirriahi@unisa.edu.au*

This paper reports on the findings of a study that investigated the effects of instructional conditions and prior experience on students' self-reflection. The study was conducted with the use of a video annotation tool that was used by undergraduate performing arts students to reflect on their video recorded performances. The study shows a consistent positive effect of previous experience with the video annotation tool for engagement with reflection. Graded instructional conditions with feedback had a positive effect on increasing higher order reflections particularly for students with prior experience with the video annotation tool for reflective purposes. The finding suggests that when including reflection in the curriculum, it is important to consider introducing it at a program or degree level rather than individual courses in order to provide opportunity for students to gain experience with reflection and any particular tool that is used (e.g. a video annotation tool). Further, reflective tasks should be scaffolded into the curriculum with ample opportunity for formative feedback and summative assessment in order to encourage higher order thinking and foster students' metacognitive awareness and monitoring for increased goal-setting and acknowledgement of the motive or effect of their observed behavior.

Keywords: instructional conditions; reflection; self-reflection; self-regulated learning; video annotation

Introduction

Reflection can be considered as a critical skill common to all disciplinary fields (Henderson, Napan, & Monteiro, 2004; Ryan & Ryan, 2013). In essence, the process of reflection is underpinned by self-regulated learning (SRL) principles (Panadero, Jonsson, & Botella, 2017; Pintrich, 2000; Zimmerman & Tsikalas, 2005) that serve to effectively promote higher order thinking, develop professional practice, facilitate learning from experience (Coulson & Harvey, 2013) and stimulate autonomy and confidence (Carey, Harrison, & Dwyer, 2017). Given the importance of reflection in learning, it is not surprising there has been much research targeting how best to develop reflective practice (Ryan & Ryan, 2013). When integrated into the

curriculum with sufficient scaffolds, students can develop the reflective skills needed for effective self-monitoring and self-regulating as life-long learners (Hawe & Dixon, 2016).

There have been many curriculum-based initiatives designed to foster reflective skills such as journal entries (Henderson et al., 2004; Lew & Schmidt, 2011). More recently, the use of video for reflection purposes has been increasingly used in a diversity of education contexts. For instance, video as a tool to aid reflection has been adopted in medical education (Hulsman & van der Vloodt, 2015), pre-service teacher education (Blomberg, Sherin, Renkl, Glogger, & Seidel, 2014), dance education (Leijen, Lam, Wildschut, Robert-Jan Simons, & Admiraal, 2009), and music education (Daniel, 2001). The context of the present study centres on student use of video and video annotations to promote self-reflection in an undergraduate music education program. Prior research on student use of reflective journals in music education has been shown to aid establishment of goals, prompt conversation between students and the lecturer, and help students take ownership of their learning (Carey et al., 2017). While Daniel (2001) investigated how students perceive the use of videos for self-assessing their performance, there are limited studies on the specificity (or type of) reflective statements students make and the impact of experience and the instructional condition on the level of reflection (e.g. higher order reflection in the form of goals or motives compared to lower order reflection in the form of observation of behaviour). This study aims to address this gap by examining the types of video annotations students make (reflective statements) when reflecting on their video-recorded performance.

Reflection as part of self-regulated learning

In their theoretical framing of self-regulated learning (SRL), Winne and Hadwin suggest that students are active agents in their learning process by choosing their study tactics (Winne, 2011; Winne & Hadwin, 1998). As a result, students create products of their learning such as quiz

responses (assessment score), or writing tasks (e.g., reflective statements). These products of learning are essentially evaluated against pre-conceived standards to determine the effectiveness of their chosen study tactics. The process of evaluation of study tactics and learning products is central to *metacognitive monitoring* and constitutes the most critical activity associated with the notion of *reflective learning*. Metacognitive monitoring is the source of *internal feedback* that drives future learning (Butler & Winne, 1995). Accordingly, reflective practice is recognized in many professions such as medicine and finance (Epstein, Siegel, & Silberman, 2008; Milligan, Fontana, Littlejohn, & Margaryan, 2015) as crucial for allowing practitioners "to evaluate their own strengths, weaknesses and learning needs in order to maintain a level of competence in accordance with the professional norm" (Hulsman & van der Vloodt, 2015, p. 356). Similarly, in the performing arts sector, students' reflective skills are critical for career prospects in their profession (Daniel, 2001).

The literature to date has documented student weaknesses in their ability to undertake effective reflective learning practice as part of SRL. These weaknesses are commonly associated with a poor choice of study tactics and inaccuracies in judging the effectiveness of their learning (Bjork, Dunlosky, & Kornell, 2013; Dunlosky & Lipko, 2007). Rather than assuming students will somehow develop effective reflection capabilities, the desirable skills need to be promoted as part of the formal curriculum (Driessen, Tartwijk, & Dornan, 2008). This notion is grounded in existing empirical work in educational psychology. For example, Azevedo and colleagues demonstrated that students assigned to the SRL condition without appropriate scaffolding and development tended to adopt ineffective learning strategies (Azevedo, Moos, Greene, Winters, & Cromley, 2008). More recently, Dunlosky and colleagues have provided a comprehensive overview of the types of effective study tactics and instructional recommendations that support

students in their learning process (Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013). Central to this line of work is the integration of instructional tasks that scaffold the development of students' reflective skills.

Instructional conditions and experience

The use of various technologies in undergraduate education, such as blogs and video annotation software have been noted as tools that can promote reflective learning (Hall & Davison, 2007; Hulsman, Harmsen, & Fabriek, 2009; Leijen et al., 2009). However, the simple availability of a tool in a learning environment, does not mean that the students will immediately use it, nor use it effectively (Clarebout, Elen, Collazo, Lust, & Jiang, 2013). Winne (2006) posits that students will adopt a new tool only if they i) are aware that the tool can be useful for their learning; ii) recognise that the tool can be applied to their task at hand; and iii) have sufficient skills to use the tool effectively. Instructional designs, therefore, need to scaffold the process of tool use in order to help students appreciate and adopt the new tool in their regular study practice (McGill & Klobas, 2009). Building on the work of Winne (2006), the authors identified that after students were introduced to a video annotation tool with graded tasks, they sustained both the volume and quality of use[1] of the tool in future courses regardless of the task being assessed. Hence, the scaffolding and design of the *instructional conditions* are considered as essential for promoting effective use of educational technology for learning purposes (Gašević et al., 2017).

[1] The volume of use is measured by the number of video annotations created and the amount of text produced in the video annotations.

Assessment of learning products

The quality of learning products can be considered as measure of the effectiveness of student learning. One way of assessing the quality of learning products (e.g. the reflective statements) associated with self-feedback or reflection is through the level of specificity. According to Goodman, Wood, and Hendrickx (2004), specificity is determined by the level of information offered in feedback messages. Much of the literature on specificity relates to peer feedback and coding schemes used to assess the levels of peer feedback provided which typically focus on the task level (e.g. correct or not), process level (e.g. provides strategies for improvement), self-regulation level (e.g. probes questions to guide the peer), and self level (e.g. positive reinforcement) (Gan & Hattie, 2014; Hattie & Timperley, 2007). Similarly, by noting specific information, students provide evidence about the extent to which they have internalized standards necessary to evaluate the effectiveness of their own learning - the core activity of SRL (Bjork et al., 2013; Winne, 2011; Winne & Hadwin, 1998). Moreover, the specificity of their goals is associated with their expertise development (Zimmerman, 2006). Therefore, reflection of one's own learning, as a form of self-feedback, needs to be specific to offer evidence of learning progression.

Research questions

The literature offers several models for assessment of reflection specificity. These models have been typically applied in studies that investigated the effects of instructional designs that included peer-evaluation as a way to increase the effectiveness of reflections (Hulsman et al., 2009). However, what is much less understood is the association of specificity of reflection with i) different instructional conditions under which a tool for reflective learning is used in different

courses, and ii) previous experience with the tool for reflective learning. Therefore, this paper looks at the following two research questions:

RQ1. What is the effect of the instructional conditions at the course level (assessed vs. non-assessed) on the frequency of specific types of student reflections?

RQ2. What is the effect of students' prior experience with a learning tool on the frequency of specific types of students' reflections in future courses when instructional conditions at the course level (assessed vs. non-assessed) are the same and when they differ?

Method

Study design

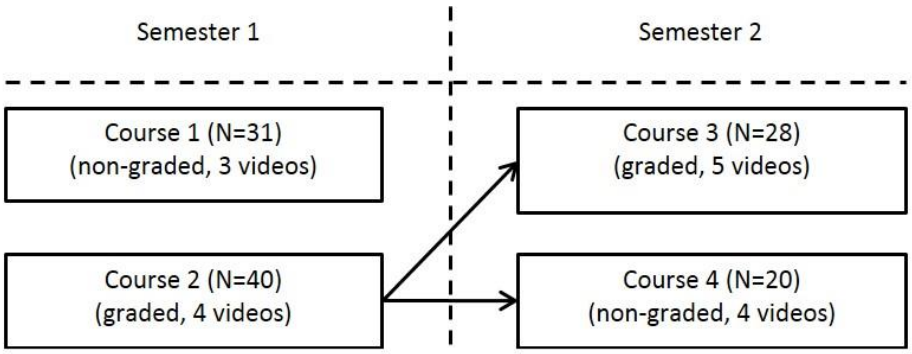
The study took the form of a natural experiment (Dunning, 2012) whereby the conditions of the experiment were outside the control of the investigators. The study was placed within the context of courses available at a higher education institution in North America. The researchers had no control over the experimental assignment of the study participants to either the assessed or non-assessed condition. Rather, the study was conducted in an ecologically valid setting in which the assignment to the experimental conditions was performed through the participants' enrolment in the courses used in the study. Analysis was conducted using trace data that was logged by the Collaborative Lecture Annotation System (CLAS), a web-based video annotation software used in the courses (Risko, Foulsham, Dawson, & Kingstone, 2013). The instructional conditions for the use of CLAS in each of the courses and students' experience with the tool determined the experimental conditions.

Materials

CLAS includes *two types of video annotation features*: i) *time-stamped annotations* – annotations that are associated to a particular part of a video; and ii) *general annotations* – comments that are not linked to any particular part of a video, but rather enable users to create a general annotation or summary of the entire video. Similar to contemporary video players, CLAS has functionalities for interacting with the video – e.g., pause playing, resume playing, rewind, and fast forward.

Sample

Four undergraduate courses in the performing arts discipline offered across two semesters within the 2012/2013 academic year were included in the study. In all four courses, students' performances were video recorded and made available for review and reflective annotation using CLAS. All students enrolled in Courses 1 through 4 were included in the study and students could enroll in any course, provided they met the course prerequisites. Hence, some students enrolled in multiple courses across the two semesters leading to some students having had experience with CLAS in a course prior to starting their second course. Figure 1 illustrates the relationship between the four courses used in the study and the instructional condition of each course. The notes include the number of students who took each course and the number of students who took multiple courses indicating prior experience with CLAS and reflective annotations.



Notes:

Students enrolled in course 1 but not enrolled in Course 2 (n=23)

Students enrolled in Course 2, but not enrolled in Course 3 (n=29)

Students enrolled in Courses 1 and 2 (n=8)

Students enrolled in Courses 2 and 4 (n=11)

Students enrolled in Course 2, but not enrolled in Course 1 (n=32)

Students enrolled in Course 3, but not enrolled in Course 2 (n=10)

Students enrolled in Course 2, but not enrolled in Course 3 (n=22)

Students enrolled in Course 4, but not enrolled in Course 2 (n=9)

Students enrolled in Course 2 and 3 (n=18)

Figure 1. Relationship between courses and the instructional conditions of each.

Content analysis

To assess the type of and frequency of reflective annotations made, quantitative content analysis (Rourke & Anderson, 2004) was undertaken using an adaptation of Hulsman et al.'s (2009) coding scheme. In their study, Hulsman and colleagues (2009) coded student reflections resulting in four reflection categories: *describing a key event (*consisting of (1) *observation of own behavior* and (2) *motive or effect of own behavior)* and *finding new solutions (*consisting of (3) *asking feedback for improvements* and (4) *indicating a goal).* In the present study, Hulsman et al's (2009) coding scheme was adapted to consist of three categories – observation, motive/effect, and goal. The category *asking feedback for improvements'* was not included as the reflective task was an independent activity in contrast to the study conducted by Hulsman and colleagues where there was an element of peer feedback. Coding instructions were clarified as shown in Table 1. The unit of analysis was a subset of a sentence as shown in the example below

representing two coded units of analysis (goal; observation):

Staying with the ensemble through preps and through releases are very important [goal] and I am not doing those anywhere near enough [observation].

A total of 971 general annotations were coded. To test the validity of the coding process, one coder coded a subsample of 70 annotations in which 268 units of analysis were identified. A second coder coded a subsample of 59 annotations in which 342 units of analysis were detected. Both coders received the same instructions and went through the same training process for coding. They both independently coded smaller subsamples of general annotations until they reached Cohen's Kappa inter-rater reliability higher than .75. Both coders were also asked to recode their datasets. The values of Cohen's Kappa for intra-rater reliability for the first coder was .98 and for the second one .96.

Table 1. Coding scheme used for content analysis.

Category	Definition	Example
Observation	The student indicates what they have observed or how they felt or feel about something. This observation may apply to their own behaviour or the behaviour of the other students. They do **not** indicate why the behaviour occurred.	'*I feel like that in this lab I felt more comfortable than before standing in front of this group than in previous labs.*' '*I still continue to have problems with making eye contact...*'
Motives/ Effects	The student indicates what they have observed but also explains **why** it occurred (in the past). It does **not** apply to when a student writes a goal and provides a reason for the goal.	'*Whenever I'm getting tense when I play trumpet I try and simplify what I'm doing.*' *'..., but being up there made me feel insecure and nervous, which led to my eyes dropping more frequently.'*
Goals	The student indicates what they will do next time or what they need to work on (in the future or from now on). This also includes if a student indicates an ideal standard behaviour that should be achieved (in the future).	'*I think for my next lab my goals shall be to try to make eye contact with everyone at least once.*' '*I think a good way to work on keeping the eyes focused would be to work with a partner.*'

Variables

Fixed effects

To address the research questions, two independent (and categorical binary) variables were identified as fixed effects – instructional condition and prior experience with CLAS. The instructional condition variable (used to address research question 1) represented whether the reflective learning experience with CLAS was graded and formative feedback provided. The prior experience variable (used to address research question 2) represented whether students enrolled in a prior course where reflective learning with CLAS was graded and formative feedback provided. Prior experience refers to students who completed Course 2 prior to enrolling in Courses 3 and 4.

Outcome variables

Outcome variables were counts of the three reflection types (observation, motive/effect, and goal) per student. Counts of the total number of general annotations were also used to investigate the effects of instructional conditions and experience on the total volume of the activity performed with the tool.

Random effects

While each of the variables was measured at the individual level, students were nested within the courses (e.g. some students enrolled in multiple courses). To account for the variability at the individual and course level, we identified *student within a course* as a random effect, a categorical variable representing variability. Generally, random effects are categorical variables representing variability amongst subjects

Statistical analysis

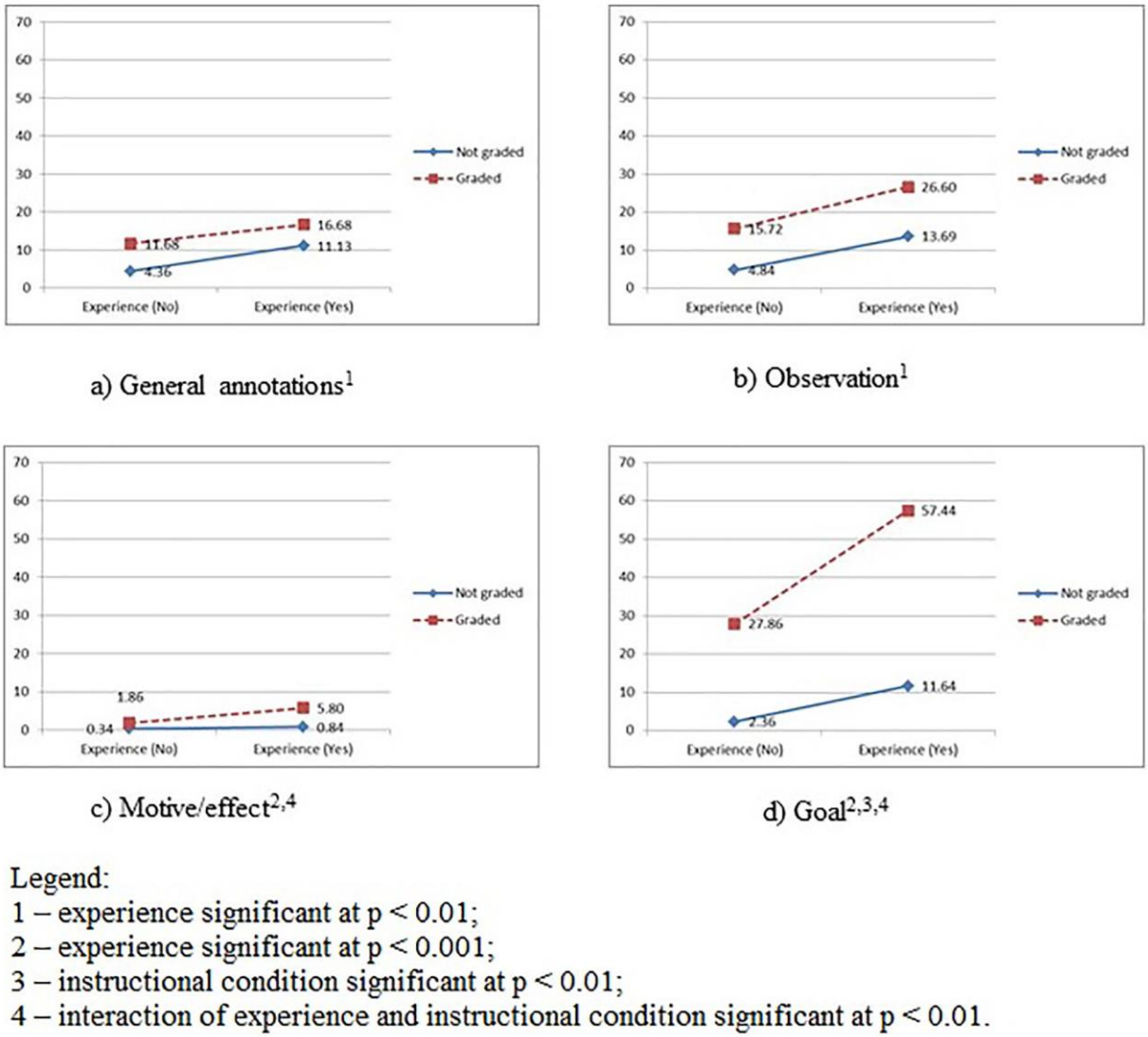
Due to the nested structure of the data (students within a course) and potential issue of correlated data due to grouping (Seltman, 2012), linear mixed models were used to address the research questions. The flexible and robust approach of mixed-effects modeling allows for the modeling of a wide set of correlation patterns, a suggested recommendation when studying similar datasets (Pinheiro & Bates, 2009; Seltman, 2012). By including a combination of fixed and random effects and after accounting for any extraneous random effects, mixed-effects models can assess the influence of the fixed effects on the dependent variables.

Mixed-effects models were used to examine the association between the two conditional factors (i.e., instructional design and previous experience) and the associated outcomes (i.e. counts of each level of student reflection and the total count of general annotations). In order to assess whether the independent variables predict the activity (count of reflections) above and beyond the random effects, two models were built for each of the dependent variables – *null model* and *final model* (see Table). The *null model* included the random effect only (*student within a course*). However, when the association between count of motive/effect reflections and the independent variables were analyzed, the model could not fit such a structure of random effects. Therefore, only in the single case, *student* was used as a random effect. On the other hand, the *final model* included *condition*, *experience*, and *interaction between condition and experience*, as fixed effects. Intraclass correlation coefficient (ICC) (Raudenbush & Bryk, 2002), second-order Akaike information criterion (AICc), and likelihood ratio test (Hastie, Tibshirani, & Friedman, 2011) were used to decide on the best fitting model (Table). An effect size (R^2) for each model as goodness-of-fit measure was also estimated, calculating the variance explained using the method suggested by Xu (2003).

Linear mixed-effects models were conducted using R v.3.0.1 software for statistical analysis with package lme4 (Bates, Maechler, Bolker, & Walker, 2015). The hypotheses specify the direction of the effect, however two-tailed tests were used for significance testing with an alpha level of .05.

Results

Figure illustrates the estimated mean values for the three reflective states (observation, motive/effect, and goals) and overall total general annotations. The results are determined after performing the mixed linear modeling based on prior experience with CLAS in a graded condition and current instructional conditions (graded and formative feedback vs. non-graded and no formative feedback).



Legend:
1 – experience significant at $p < 0.01$;
2 – experience significant at $p < 0.001$;
3 – instructional condition significant at $p < 0.01$;
4 – interaction of experience and instructional condition significant at $p < 0.01$.

Figure 2. Effects of instructional conditions and experience with video annotation on the counts of general annotations and types of reflections.

Overall, the estimated mean value of the counts of motive/effect reflections can be seen to be considerably lower than the other types of reflective states. The results also suggest that a student's prior experience with the tool and the interaction between experience and instructional condition also have a significant effect on the quantity of reflective statements made. The estimated mean values of goals increased significantly when students had prior experience with

the tool and the instructional condition was graded. The effect of interaction between the prior experience and the instructional condition is also significant. Finally, only prior experience with the tool had significance on the estimated mean values for the total count of general annotations and observations.

The likelihood ratio test for ***general annotation count*** models yielded a significantly better fit for the *final* model (i.e., the model that included fixed and random effects) than the *null* model. The model (see Figure 3d) shows a significant effect of previous experience on the total number of general annotations produced – $F(1, 25.150) = 8.723$, $p = .007$, while the effect of the instructional condition was marginally significant – $F(1, 2.048) = 12.214$, $p = .071$. These results can be interpreted as students who have prior experience with CLAS, produce a higher number of general annotations than those who have not used CLAS previously. Pairwise comparison of least square means test further confirmed this by showing significant difference between the two levels of experience: no experience vs. with experience – $t(25.10) = -2.95$, $p = .007$, 95% CI [-8.01, -1.43].

Comparing the *null* and the *final* models for ***observation count*** as a dependent variable, there was a better fit for the *final model* that accounted for fixed and random effects (see Table 2). However, the model (Figure 3b) revealed a significant effect of previous experience only – $F(1, 71.996) = 7.678$, $p = .007$. This result suggests that students with previous experience using CLAS produce more *observations* when they reflect on their performance. Further analysis of pairwise comparison of least square means revealed a significant difference between the two levels of previous experience: no experience vs. with experience – $t(11.40) = -5.51$, $p < .001$, 95% CI [-26.40, -11.40].

Table 2. Inferential statistics for the model fit assessment.

Model	χ^2	df	R^2 marg.	R^2 cond.	AICc	ICC student	ICC course
Count of general annotations							
null			-	.707	638.834	.286	.421
final	**14.220****	**3**	**.338**	**.846**	**624.837**	**.675**	**.092**
Observation count							
null			-	.691	781.502	.275	.416
final	**11.077***	**3**	**.226**	**.741**	**763.103**	**.422**	**.243**
Motive/Effect count							
null			-	≈ 0	524.464	-	≈ 0
final	**34.417*****	**3**	**.293**	**.360**	**492.927**	**-**	**.095**
Goal count							
null			-	.684	832.911	.104	.580
final	**42.866*****	**3**	**.656**	**.891**	**786.733**	**.657**	**.027**

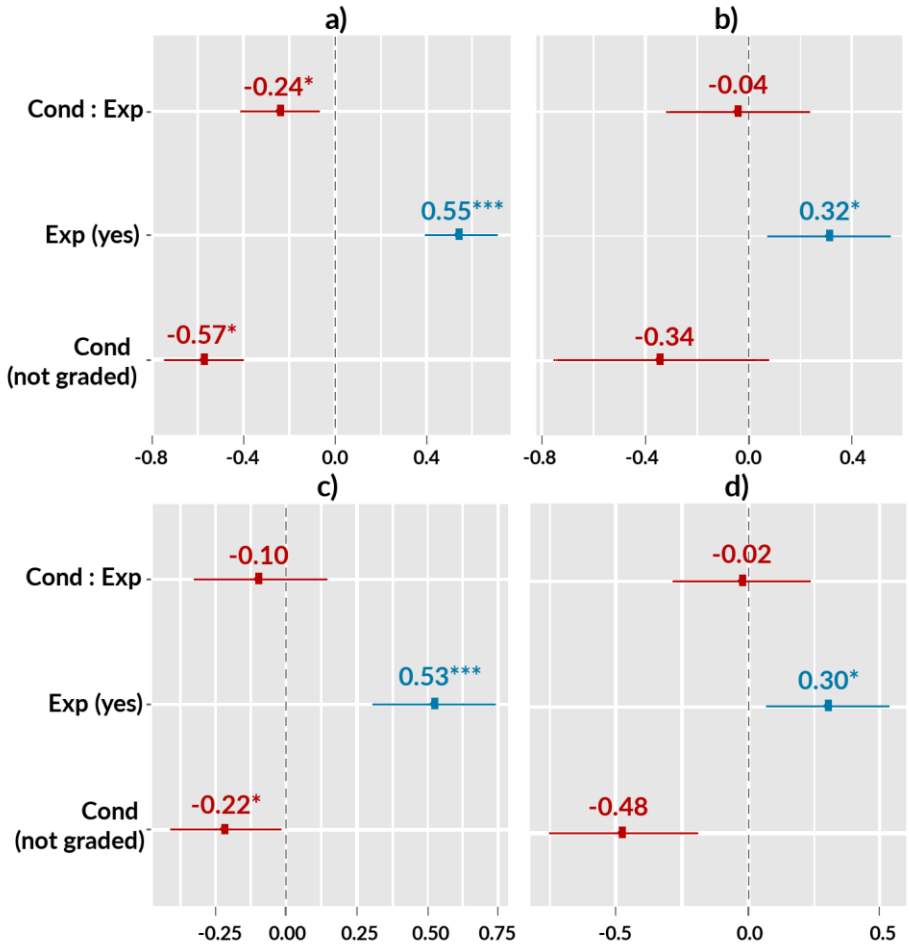


Figure 3. Effect size for fixed effects (instructional condition and experience) on the reflection types (*denotes significance).

The likelihood ratio test between the *null* and *final* model for the ***motive/effect count*** revealed a significantly better fit of the model that accounted for instructional conditions and prior experience in the tool use (Table 2). The model (Figure 3c) shows a significant effect of *condition*, $F(1, 87.161) = 9.697$, $p = .002$ and *experience*, $F(1, 88.338) = 27.408$, $p < .001$. These results indicate that students with previous experience with CLAS and enrolled in a course with the graded condition tend to produce more Motive/Effect reflections. The pairwise comparison of least square means further confirmed differences between the two levels of the previous experience: no experience vs. with experience – $t(88.30) = -5.24$, $p < .001$, 95% CI [-

4.72, -2.12] and between the two instructional conditions: graded vs. non graded – $t(87.20) = 3.11$, $p = .002$, 95% CI [0.73, 3.33].

The likelihood ratio test between the *null* and *final* model for the ***goal count*** as the outcome variable revealed a significantly better fit of the model that accounted for instructional conditions and prior experience in the tool use (Table 2). The model (Figure 3a) shows a significant effect of *condition*, $F(1, 2.141) = 89.884$, $p = .009$, *experience*, $F(1, 11.377) = 30.347$, $p < .001$, and *interaction between condition and experience*, $F(1, 11.377) = 6.829$, $p = .023$. This can be interpreted as students in a graded condition and with previous experience using CLAS tend to produce more goals in their reflections. Furthermore, pairwise comparison of least square means showed significant differences between the two instructional conditions: graded vs. non graded – $t(2.10) = 9.48$, $p = .009$, 95% CI [20.80, 51.60], as well as between the two levels of previous experience: no experience vs. with experience - $t(11.40) = -5.51$, $p < .001$, 95% CI [-26.40, -11.40].

Discussion

Research question 1 aimed to identify the effect of the instructional conditions (assessed vs. non-assessed) on the frequency of specific types of students' reflections. The findings of the study suggest that the instructional condition assigned to a course (e.g. graded reflections with formative feedback) influences the number of student reflections related to reflections of higher specificity levels – i.e., *motive/effect* and *goals*. As previous research has posited (Grossman, 2009; Henderson et al., 2004; Ryan, 2013), there are different levels of reflection which are used to describe the specific differences between reflective statements ranging from descriptive or observational accounts to higher order and transformative or goal-oriented. However, broadly, the different levels focus on two main elements of reflection: a way to make sense of an

experience and to reimagine or plan for the future (Ryan, 2013). It is the latter, often called ‘reconstructing’ or simply ‘goal-setting’ which is the most difficult for students to achieve as they need to demonstrate new ways of approaching an issue or possible actions to take (Ryan, 2013). This relates to the findings of this study (specifically, research question 1) where more goals were written by students in courses where their reflections were graded and feedback provided. Since goal-setting is a more challenging reflective task (Ryan, 2013), raising extrinsic motivation (e.g. grades) by making it transparent so that the instructor will read the annotations to provide formative feedback and grade quality can be seen to motivate students to place greater effort on the task. Likewise, the writing of motive/effect reflections uses higher order thinking than merely describing a particular behavior (e.g. observation). Again the findings show that when the instructional conditions prescribe that the instructor reads and grades the reflective comments students are more inclined to think more critically about their observed behavior and provide an explanation (e.g. motive) or note how their behavior may influence a result (e.g. effect).

Research question 2 related to how prior experience with a learning tool, obtained in the instructional condition with grading and feedback, affects the frequency of specific types of students’ reflections in future courses when instructional conditions are the same; and when they differ. The results of the study further emphasise that prior experience using video annotation software for graded reflection (and with formative feedback provided) has a positive effect on the number of reflective annotations made on all three specificity levels of reflection. This supports Winne's (2006) argument that the adoption of a learning tool is based on a combination of student awareness of the benefits of a learning tool and how it can be applied to a particular task along with skills to use the tool. It further supports the prior findings that scaffolding is

required when introducing a new learning tool (McGill & Klobas, 2009) and that students require experience using a new learning tool where they receive grades and feedback in order to sustain their use of the tool in the future when it is an optional activity or resource (Gašević et al., 2017). In addition to the need for experience for the adoption of a new learning tool, it is also critical for learning how to reflect. Learning to reflect is not a quick process and requires time and early inclusion in the curriculum to allow for initial scaffolding leading to greater learner agency (Coulson & Harvey, 2013). Pedagogical decisions pertaining to reflective tasks should consider where in a program or degree students have been introduced to reflective pratice, how it has been developed and what further support is needed to further expand their skills (Ryan & Ryan, 2013). Hence, as students gain more experience with tasks that promote reflection, they will be more likely to engage in reflective practice in the future.

Students with experience gained in the instructional condition with grading and feedback benefited the most in terms of higher order reflections (motive/effect and goals) when they are again requested to reflect in a graded condition with feedback. This is corroborated by the significant effect of the interaction term in the mixed models for the counts of motive-effects and goals. This is due to two main factors – familiarity with the study tactic and familiarity with the cues (i.e., task conditions) (O'Neil, 2014; Winne, 2006). That is, students with experience have a more elaborated schemata for reflection in using the video annotation tool (Winne & Hadwin, 1998). These findings emphasise the importance in recognising when students may not have the required experience or knowledge of the learning tactics and effective strategies for successful completion of a task and provide additional guidance when required.

The change in how students approach the reflective tasks in the two different instructional conditions reported in the present study coupled with their increasing experience

resonates with the literature on self-regulated learning. Butler and Winne (1995) stress that as students monitor their learning, they evaluate their tactics or strategies and adapt their approaches and effort invested in alignment with a set of standards (preconceived or established). In relation to students' reflective annotations, as students gain experience with the reflective tasks within the instructional condition of graded and formative feedback and as they progress to future courses where the instructional condition remains the same, their strategy leads to a greater number of higher order reflections. In courses where the instructional condition differs (e.g. no graded reflections and no formative feedback), students adapt their strategy (e.g. focus more on observational reflections). However, what remains to be resolved is how students can be encouraged to continue to tap into their higher order thinking and continue to write both observational reflections and the more challenging motive/effect and goal-setting types regardless of the instructional condition. As Butler and Winne (1995) posit, external feedback (e.g. from the instructor) can be helpful in guiding students to monitor and adapt their strategies rather than leaving students to rely on their own internal feedback which does not necessarily provide adequate or accurate information about the quality of their performance nor suggest more appropriate strategies. If formative feedback is offered to students despite the reflective task not being graded, it can help promote student understanding of the purpose of the learning task and act as a catalyst for self-regulated learning by affecting students' motivations, thinking, and actions and contribute to improved meta-cognitive self-monitoring and self-regulation (Hawe & Dixon, 2016).

Conclusion

The findings of this study show that prior experience and instructional conditions influence students' reflections. Experience has a significant effect on the number of observational

annotations made as well as the overall number of annotations while a combination of prior experience with reflection and summative assessment with feedback have a significant effect on more higher order reflections (e.g. goals and motive/effect). These findings are in line with Dunlosky and colleagues (2013) who stress the importance of instructional conditions (e.g. practice testing) to scaffold and develop SRL proficiency.

Hence, when incorporating reflection practices into the curriculum, it is important to consider introducing it at the program or degree level rather than individual disparate courses. A more holistic approach can provide further opportunity for students to gain experience with reflection processes using a particular supported medium (e.g. a video annotation tool). Further, reflective tasks should be scaffolded into the curriculum with ample opportunity for formative feedback and summative assessment (e.g. grades) in order to encourage higher order thinking and foster students' metacognitive awareness and monitoring for increased goal-setting and acknowledgement of the motive or effect of their observed behavior. Future longitudinal studies can expand on this research by investigating the effect that experience and instructional conditions over a greater amount of time can have on students' reflective writing and SRL skills when formative feedback and summative assessment is gradually reduced over the course of a three or four year degree program.

The study reported has some key limitations that future research can address. First, the study was conducted at one higher education institution and within a single discipline area. Future studies can investigate the type of reflective annotations students write in other disciplines and at other higher education institutions. Second, the study was largely exploratory with no control over the setting or instructional conditions such as instructor cuing effects that may have affected the findings. Future experimental or more controlled studies can be designed to include

instances where formative feedback is provided on non-graded reflections to identify, along with prior experience, whether instructor feedback, grades, instructor cueing effects, or a combination is most conducive for high order reflection.

Acknowledgements

The views expressed in this publication/activity do not necessarily reflect the views of the Australian Government Office for Learning and Teaching. We would like to thank Daniyal Liaqat, Dennis Alonzo, and Bardia Mohabbati for their assistance with coding the annotation data.

Disclosure statement
No potential conflict of interest was reported by the authors

Funding
The study was funded by the Natural Sciences and Engineering Research Council of Canada (Discovery Grant) and the Social Sciences and Humanities Research Council of Canada (Insight Grant). In addition, support for this publication/activity has been provided by the Australian Government Office for Learning and Teaching [grant number ID13-2901]

References

Azevedo, R., Moos, D. C., Greene, J. A., Winters, F. I., & Cromley, J. G. (2008). Why is externally-facilitated regulated learning more effective than self-regulated learning with hypermedia? *Educational Technology Research and Development*, *56*(1), 45–72. https://doi.org/10.1007/s11423-007-9067-0

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). *lme4: Linear mixed-effects models using Eigen and S4*. Retrieved from http://CRAN.R-project.org/package=lme4

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-Regulated Learning: Beliefs, Techniques, and Illusions. *Annual Review of Psychology*, *64*(1), 417–444. https://doi.org/10.1146/annurev-psych-113011-143823

Blomberg, G., Sherin, M. G., Renkl, A., Glogger, I., & Seidel, T. (2014). Understanding video as a tool for teacher education: investigating instructional strategies to promote reflection. *Instructional Science*, *42*(3), 443–463. https://doi.org/10.1007/s11251-013-9281-6

Butler, D. L., & Winne, P. H. (1995). Feedback and Self-Regulated Learning: A Theoretical Synthesis. *Review of Educational Research*, *65*(3), 245–281. https://doi.org/10.3102/00346543065003245

Carey, G., Harrison, S., & Dwyer, R. (2017). Encouraging reflective practice in conservatoire students: a pathway to autonomous learning? *Music Education Research*, *19*(1), 99–110. https://doi.org/10.1080/14613808.2016.1238060

Clarebout, G., Elen, J., Collazo, N. A. J., Lust, G., & Jiang, L. (2013). Metacognition and the Use of Tools. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 187–195). Springer New York. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4419-5546-3_13

Coulson, D., & Harvey, M. (2013). Scaffolding student reflection for experience-based learning: a framework. *Teaching in Higher Education*, *18*(4), 401–413. https://doi.org/10.1080/13562517.2012.752726

Daniel, R. (2001). Self-assessment in performance. *British Journal of Music Education*, *18*(03). https://doi.org/10.1017/S0265051701000316

Driessen, E., Tartwijk, J. van, & Dornan, T. (2008). The self critical doctor: helping students become more reflective. *BMJ*, *336*(7648), 827–830. https://doi.org/10.1136/bmj.39503.608032.AD

Dunlosky, J., & Lipko, A. R. (2007). Metacomprehension A Brief History and How to Improve Its Accuracy. *Current Directions in Psychological Science*, *16*(4), 228–232. https://doi.org/10.1111/j.1467-8721.2007.00509.x

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving Students' Learning With Effective Learning Techniques Promising Directions From Cognitive and Educational Psychology. *Psychological Science in the Public Interest*, *14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Dunning, T. (2012). *Natural Experiments in the Social Sciences: A Design-Based Approach*. Cambridge University Press, Cambridge, United Kingdom.

Epstein, R. M., Siegel, D. J., & Silberman, J. (2008). Self-monitoring in clinical practice: A challenge for medical educators. *Journal of Continuing Education in the Health Professions*, *28*(1), 5–13. https://doi.org/10.1002/chp.149

Gan, M. J. S., & Hattie, J. (2014). Prompting secondary students' use of criteria, feedback specificity and feedback levels during an investigative task. *Instructional Science*, *42*(6), 861–878. https://doi.org/10.1007/s11251-014-9319-4

Gašević, D., Mirriahi, N., Dawson, S., & Joksimović, S. (2017). Effects of instructional conditions and experience on the adoption of a learning tool. *Computers in Human Behavior*, *67*, 207–220. https://doi.org/10.1016/j.chb.2016.10.026

Goodman, J. S., Wood, R. E., & Hendrickx, M. (2004). Feedback Specificity, Exploration, and Learning. *Journal of Applied Psychology*, *89*(2), 248–262.

Grossman, R. (2009). Structures for Facilitating Student Reflection. *College Teaching*, *57*(1), 15–22. https://doi.org/10.3200/CTCH.57.1.15-22

Hall, H., & Davison, B. (2007). Social software as support in hybrid learning environments: The value of the blog as a tool for reflective learning and peer support. *Library & Information Science Research*, *29*(2), 163–187.

Hastie, T., Tibshirani, R., & Friedman, J. (2011). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition* (2nd ed. 2009. Corr. 7th printing 2013 edition). New York, NY: Springer.

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hawe, E., & Dixon, H. (2016). Assessment for learning: a catalyst for student self-regulation. *Assessment & Evaluation in Higher Education*, 1–12. https://doi.org/10.1080/02602938.2016.1236360

Henderson, K., Napan, K., & Monteiro, S. (2004). Encouraging reflective learning: An online challenge. In *Beyond the comfort zone: Proceedings of the 21st ASCILITE Conference* (pp. 357–364). Retrieved from http://www.ascilite.org.au/conferences/perth04/procs/henderson.html

Hulsman, R. L., Harmsen, A. B., & Fabriek, M. (2009). Reflective teaching of medical communication skills with DiViDU: Assessing the level of student reflection on recorded consultations with simulated patients. *Patient Education and Counseling*, *74*(2), 142–149. https://doi.org/10.1016/j.pec.2008.10.009

Hulsman, R. L., & van der Vloodt, J. (2015). Self-evaluation and peer-feedback of medical students' communication skills using a web-based video annotation system. Exploring content and specificity. *Patient Education and Counseling*, *98*(3), 356–363. https://doi.org/10.1016/j.pec.2014.11.007

Leijen, Ä., Lam, I., Wildschut, L., Robert-Jan Simons, P., & Admiraal, W. (2009). Streaming video to enhance students' reflection in dance education. *Computers & Education*, *52*(1), 169–176. https://doi.org/10.1016/j.compedu.2008.07.010

Lew, D. N. M., & Schmidt, H. G. (2011). Writing to learn: can reflection journals be used to promote self-reflection and learning? *Higher Education Research & Development*, *30*(4), 519–532. https://doi.org/10.1080/07294360.2010.512627

McGill, T. J., & Klobas, J. E. (2009). A task–technology fit view of learning management system impact. *Computers & Education*, *52*(2), 496–508. https://doi.org/10.1016/j.compedu.2008.10.002

Milligan, C., Fontana, R. P., Littlejohn, A., & Margaryan, A. (2015). Self-regulated learning behaviour in the finance industry. *Journal of Workplace Learning*, *27*(5), 387–402. https://doi.org/10.1108/JWL-02-2014-0011

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, *22*, 74–98. https://doi.org/10.1016/j.edurev.2017.08.004

Pinheiro, J., & Bates, D. (2009). *Mixed-Effects Models in S and S-PLUS* (1st ed. 2000. 2nd printing 2009 edition). New York: Springer.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. *Handbook of Self-Regulation*, *451*, 451–502.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods* (Second Edition edition). Thousand Oaks: SAGE Publications, Inc.

Risko, E. F., Foulsham, T., Dawson, S., & Kingstone, A. (2013). The Collaborative Lecture Annotation System (CLAS): A New TOOL for Distributed Learning. *IEEE Transactions on Learning Technologies*, *6*(1), 4–13. https://doi.org/10.1109/TLT.2012.15

Rourke, L., & Anderson, T. (2004). Validity in quantitative content analysis. *Educational Technology Research and Development*, *52*(1), 5–18. https://doi.org/10.1007/BF02504769

Ryan, M. (2013). The pedagogical balancing act: teaching reflection in higher education. *Teaching in Higher Education*, *18*(2), 144–155. https://doi.org/10.1080/13562517.2012.694104

Ryan, M., & Ryan, M. (2013). Theorising a model for teaching and assessing reflective learning in higher education. *Higher Education Research & Development*, *32*(2), 244–257. https://doi.org/10.1080/07294360.2012.661704

Seltman, H. J. (2012). *Experimental design and analysis*. Online. Retrieved from http://www.stat.cmu.edu/hseltman/309/Book/Book.pdf

Winne, P. H. (2006). How Software Technologies Can Improve Research on Learning and Bolster School Reform. *Educational Psychologist*, *41*(1), 5–17. https://doi.org/10.1207/s15326985ep4101_3

Winne, P. H. (2011). A cognitive and metacognitive analysis of self-regulated learning. In D. H. Schunk & B. Zimmerman (Eds.), *Handbook of self-regulation of learning and performance* (pp. 15–32). New Work and Oxon: Routledge.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Xu, R. (2003). Measuring explained variation in linear mixed effects models. *Statistics in Medicine*, *22*(22), 3527–3541. https://doi.org/10.1002/sim.1572

Zimmerman, B. J. (2006). Development and Adaptation of Expertise: The Role of Self-Regulatory Processes and Beliefs. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 705–722). New York, NY, US: Cambridge University Press.

Zimmerman, B. J., & Tsikalas, K. E. (2005). Can Computer-Based Learning Environments (CBLEs) Be Used as Self-Regulatory Tools to Enhance Learning? *Educational Psychologist*, *40*(4), 267–271. https://doi.org/10.1207/s15326985ep4004_8