Welsh, Matthew Brian; Begg, Stephen Hope; Bratvold, Reidar Brumer


© the authors

PERMISSIONS

correspondence from:
Business Mgr
Cognitive Science Society Inc. [cogsci@psy.utexas.edu]
University of Texas - Austin
Department of Psychology
108 E. Dean Keeton, Stop A8000
Austin

The copyright for articles and figures published in the Proceedings are held by the authors, not the Society

http://hdl.handle.net/2440/41191
Efficacy of Bias Awareness in Debiasing Oil and Gas Judgments

Matthew B. Welsh (matthew.welsh@adelaide.edu.au)
Australian School of Petroleum, University of Adelaide, North Terrace
Adelaide, SA 5005, Australia

Steve H. Begg (steve.begg@adelaide.edu.au)
Australian School of Petroleum, University of Adelaide, North Terrace
Adelaide, SA 5005, Australia

Reidar B. Bratvold (reidar.bratvold@uis.no)
Department of Petroleum Engineering, University of Stavanger
4036 Stavanger, Norway

Abstract

It is argued that biases such as anchoring and overconfidence contribute to a US$30 billion/year loss in the oil and gas industry (Goode, 2002). The most commonly used debiasing technique, within the industry, is awareness-style training, where participants have the biases and debiasing techniques described to them without specific training. Given such training is delivered by consultants, however, there is little available evidence of its efficacy and concern regarding a lack of up to date debiasing methods. We present a study designed to measure the benefit of such awareness-style training for the well-known anchoring and overconfidence biases, using a sample of petroleum engineering students. Results indicate that course attendance reduced participants’ overconfidence - calibration improving by 21% - but no benefit was observed for anchoring. It is argued that this difference results from the debiasing technique for anchoring requiring a greater degree of domain knowledge than the students possessed. A detailed analysis of the relationship between accuracy in, and susceptibility to, the anchoring question supported this – indicating that participants who simply relied on the anchor actually performed better than those who hazarded their own guesses. Potential benefits of debiasing and, specifically, the incorporation of up to date debiasing techniques, are discussed along with the need for further research.

Keywords: decision making; debiasing; anchoring; overconfidence; oil and gas; judgment.

Introduction

Previous research has demonstrated that most people are susceptible to cognitive biases when making decisions or judgments under conditions of uncertainty (Morgan & Henrion, 1990; Tversky & Kahneman, 1974). That is, people tend to give responses that are systematically biased compared to calculated, optimal solutions as a result of the simplified mechanisms that our brains use to make judgments.

Two of the most commonly discussed biases are: anchoring (Tversky & Kahneman, 1974), which describes the tendency people have to base estimates on any number at hand regardless of its relevance to the question at hand (Chapman & Johnson, 2002); and overconfidence (Lichtenstein, Fischhoff, & Phillips, 1982; Tversky & Kahneman, 1974), which describes the tendency of people to give too narrow bounds when asked to set a range that they are certain to some stated level of confidence that a value will fall within.

In an industry such as oil and gas, biased decision-making can be extremely expensive; industry observers argue that such biases are, in part at least, responsible for the large losses in the oil industry – one recent estimate being that unexpected outcomes during exploration cost the industry US$30 billion each year (Goode, 2002). With the cost of drilling a single, off-shore well regularly exceeding US$60 million, the estimates of technical parameters that feed into the models of oil fields used to predict oil reserves can have a huge impact on whether a company chooses to drill or walk away from a prospect. If these estimates are subject to cognitive biases – as most judgments under uncertainty are – the accuracy of these models are also compromised and poor decisions will result.

Biases in Oil and Gas Decisions

The question for the industry, of course, is: which biases are affecting oil and gas decisions and how can these be reduced? This question has not, of course, avoided investigation. Within a few years of Tversky and Kahneman’s (1974) groundbreaking work on biases, Capen (1976) introduced the concept of overconfidence to the industry – demonstrating its effect on oil and gas professionals. In 2002, however, Hawkins, Cunningham and Coopersmith (2002) noted that less than half of observed values fell inside the 80% confidence intervals commonly used for predicting uncertain parameters in the oil industry, which seems clearly indicative of overconfidence’s continued impact. Merkhofer (1987) has generalized the results of these type of studies to the “2/50” rule: When people are asked to provide estimates of quantities lying between their 1% and 99% confidence limits (that is, with only a 2% chance of being wrong), usually about 50% of the answers fall outside this range and, thus, they are overconfident in their predictions.

Many oil and gas professionals are also familiar with the anchoring bias, through the work of Rose (2001), but other biases are rarely mentioned (see Pieters, 2004 for a recent exception). This is regarded as particularly damaging due to the use of analogs in oil and gas decision-making. That is, given the scarcity of data on a new oil prospect, analogous
prospects are used to inform estimates of the parameters of interest. These analogue parameters, however, are widely regarded as acting as anchors on estimates (Rose, 2001).

Despite three decades of bias research, however, attempts at improving oil and gas decision-making have primarily focused on improving decision tools rather than looking at the impact of biases on industry decisions. Given that this approach has not demonstrated clear improvements in the industry’s economic outcomes (Bratvold, Begg, & Campbell, 2002), however, more attention is now being paid by the industry to cognitive biases and, in particular, methods for debiasing judgments.

**Debiasing Oil and Gas Decisions**

Debiasing refers to any technique designed to avoid or reduce susceptibility to bias (Larrick, 2004). As noted above, the oil and gas industry, while focused primarily on improving decision tools, has been aware of the existence of cognitive biases for a significant period. In that time, a number of groups have looked at the possibility of debiasing industry decisions.

Pete Rose, for example, has been offering training in decision-making since soon after Capen’s (1976) original introduction of the concept of biases to oil and gas and discusses what he calls the “prospector myth” (Rose, 2001), which incorporates anchoring and overconfidence. Campbell, Campbell and Campbell (2001), similarly discuss both anchoring and overconfidence.

The courses offered by such industry providers are, typically, short and intensive with the biases being demonstrated to the participants in general terms, using general questions. One concern for people interested in establishing the efficacy of such awareness-based debiasing, however, is that the groups offering this service are consultants whose primary goal is to confidentially aid their clients rather than publish their methods and results.

A second concern is that the techniques commonly described as being beneficial for debiasing tend not to be included in these courses. For example, it is widely accepted (for a review of studies, see: Morgan & Henrion, 1990) that overconfidence is best reduced using repeated feedback. That is, giving people sets of calibration questions and then showing them how often the true answers to the questions falls within their estimated ranges, before giving them further sets of questions. Such intensive training is generally not possible in the short industry courses.

By comparison with overconfidence, for which debiasing techniques are well-known, susceptibility to anchoring is seen to be extremely robust (Chapman & Johnson, 2002), with little literature on debiasing other than Wilson, Houston, Eting and Brekke (1996), who argue that - while knowledge of the anchoring effect is ineffective, high levels of knowledge about the subject matter enable anchoring to be avoided. Their data, however, do not necessarily support this statement. While the “high knowledge” groups they describe are less susceptible to the anchors they used, they are no more accurate in their estimates – indicating that they might be better described as high confidence rather than high knowledge, especially given that their level of knowledge was self-reported.

Similarly, Mussweiler, Strack and Pfeiffer’s (2000) contention that a “consider the opposite” strategy reduces the anchoring effect assumes that the true answer is already known to the elicitor (if not the elicitee) - which is certainly not the case in oil and gas where uncertainty about the true state of the world remains throughout the decision-making process. A variant that does not require prior knowledge of the answer is possible, however, where participants are told to adjust their estimates away from the anchor – the assumption being that the participant has some degree of knowledge and thus that their biased estimate lies between the anchor they saw and the hypothetical, unbiased estimate they otherwise would have made.

These observations, together, indicate another problem in anchoring research that must be addressed before it can be transferred to real world contexts such as oil and gas. Specifically, previous research has not clarified the relationship between accuracy and susceptibility in anchoring tasks. In real world environments, the goal of reducing bias is, of course, to produce more accurate estimates – not simply estimates that lie further from the anchoring value. The assumption, given an expert sample, would be that reducing the effect of the bias leads to greater accuracy but this relationship has yet to be shown.

**Research Questions**

Given the above, three research questions were formulated. The first two focus on the ability of awareness-style training in biases and debiasing techniques to reduce susceptibility to the two most talked about biases in the oil and gas context – anchoring and overconfidence. The final question asks whether the proposed, inverse relationship between accuracy and susceptibility in an anchoring task exists.

1. Does awareness of anchoring reduce susceptibility to anchoring?
2. Does awareness of overconfidence reduce susceptibility to overconfidence?
3. Are accuracy and susceptibility to anchoring inversely related?

**Methodology**

**Participants**

Data was collected from 51 petroleum engineering students, 37 males and 14 females, with a mean age of 22.6 (SD = 3.5). While not a sample of industry experts, all of the students were in their final year of study and thus possessed technical industry knowledge.

**Questionnaires**

The primary questionnaire used was prepared for teaching purposes and, as such, included a wide range of cognitive
biases, including anchoring and overconfidence. The questionnaire used to test for improvement following training, however, repeated only the anchoring and overconfidence questions.

The anchoring question asked participants to estimate the world’s proved oil reserves in billions of barrels. This question was preceded by a question asking whether the value was greater or less than either 573.9 or 1721.6 billion barrels (the low and high anchors, being one-half and one-and-one-half times the researched value of 1147.7, respectively). The ten overconfidence questions, by comparison all asked participants to set ranges that they were 80% confident that some oil industry value (e.g., the USA’s daily oil imports) fell within.

**Procedure**

Participants were handed the primary questionnaire and allowed 45 minutes to complete the questions. Immediately following this, they received 3 hours of lectures on the psychology of decision-making, cognitive biases and debiasing techniques – as part of a 5 day intensive short course on decision making in the oil and gas industry. This course was theoretical in nature – simple examples of the various biases and debiasing techniques were shown but no training, as such, was given.

A single method for debiasing anchoring was described – participants simply being told to adjust their estimates further away from the anchor in accordance with the hypothesis regarding the hypothetical, unbiased estimate described earlier.

However, three methods for debiasing overconfidence were described that could be used by a person to debias their own estimates: a simple warning that people tend to be overconfident and that they should, therefore, make their estimated ranges wider (Lichtenstein et al., 1982); consideration of counter-factuals, or reasons that the value might fall outside the initially estimated range (Hawkins et al., 2002); and the use of initial best guesses, which has been demonstrated to lead to wider ranges (Block & Harper, 1991). An interesting aside regarding this last method, however, is that the majority of industry courses assume the opposite – that asking for a best guess first reduces the width of estimated ranges, in accordance with Tversky and Kahneman’s (1974) original, untested hypothesis – and teach this to their students as fact.

While using more up-to-date recommendations, the training format used here otherwise parallels the training courses used in the oil and gas industry by consultants and, as such, offers a good indication of the efficacy of awareness as a debiasing method. After receiving the lectures, participants were given the follow-up questionnaire and allowed 30 minutes to complete it.

**Data Transformations**

In order to enable measurement of participants’ susceptibility to the two biases, their raw estimates for the anchoring task and their raw calibration score for the overconfidence task were transformed into susceptibility scores as follows:

\[
S_{\text{anch}} = - | E - A |
\]  

Where \( E \) is the participant’s estimate and \( A \) the anchoring value they saw.

\[
S_{\text{over}} = - | 8 - C |
\]  

Where \( C \) is the participant’s calibration score out of 10. Note that for both susceptibility values, \( S_{\text{anch}} \) and \( S_{\text{over}} \), higher scores indicate greater susceptibility to the relevant bias.

In addition to these susceptibility scores, accuracy scores were calculated for the anchoring responses – reflecting the distance that participants’ estimates lay from the true researched answer to the question. That is:

\[
A_{\text{anch}} = - | E - 1147.7 |
\]  

Where \( E \) is the participant’s estimate and 1147.7 is the true value. A higher \( A_{\text{anch}} \) score thus indicates greater accuracy.

**Results**

**Anchoring**

Forty-four participants provided valid responses to the anchoring question both pre- and post-training. Their responses were extremely variable and positively skewed, with many more high than low estimates. For this reason, the medians and interquartile ranges of participants’ estimates under the two anchoring conditions, both pre- and post-training, are shown in Figure 1, rather than means and standard deviations.

Looking at Figure 1, it is clear that participants showed a strong anchoring effect, with estimates made in the low anchor condition being significantly lower than those made in the high anchor condition both pre- and post-training. There also seems to be some evidence of improvement, with the median estimates of both groups falling noticeably closer to the true value post-training. A repeated measures ANOVA, however, conducted to test the strength of these effects, found that while anchor group was significant, \( F(3,84) = 11.95, p < .001 \), there was no significant effect of training or any interaction between training and anchor group. That is, no support was found for Hypothesis 1 – that awareness of anchoring would reduce susceptibility to the bias.

To separate out the effects of susceptibility to the bias and accuracy, finer-grained analyses examined the changes in susceptibility to anchoring and in accuracy (calculated from equations 1 and 3, above) across the training conditions separately. Figure 2 shows the mean susceptibility to anchoring and mean accuracy of participants in both anchoring conditions, pre- and post-training. Here one sees
a different pattern of results for the two anchoring conditions, with the participants seeing the low anchor becoming less susceptible and more accurate after training but the participants in the high anchor condition reacting in the opposite manner – becoming more susceptible and less accurate after training.

A repeated measures ANOVA undertaken to examine the susceptibility results indicated that the difference in susceptibility between the pre- and post-training sessions was not significant but confirmed that a significant interaction effect existed between training and anchor group, \( F(1,42)=4.25, p=.045 \). A second repeated measures ANOVA, examining the accuracy results, similarly found no significant main effect of training on accuracy and, in this case, the interaction effect between training and anchor group also failed to reach significance, \( F(1,42)=3.54, p=.067 \). Taken together, these results support the earlier observation that the training did not have either of the expected effects of reducing susceptibility or increasing accuracy – instead differentially affecting participants who had seen the high and low anchors.

As a final test of Hypothesis 3, correlations were calculated between participant’s accuracy and susceptibility scores. In the pre-training condition, the correlation was almost perfect, \( r(42)=0.98, p<.001 \), and it remained strong in the post-training data as well, \( r(42)=0.79, p<.001 \). The counterintuitive implication of this is that the more accurate participants were those who relied most heavily on the anchor they saw – thus indicating no support for Hypothesis 3.

**Overconfidence**

All 51 participants gave valid responses to the set of 10 calibration questions used for assessing overconfidence. The calibration scores of participants – that is, the number of hits out of 10 from the ranges set in response to the 10 calibration questions – are summarized in Figure 3, which shows the proportion of participants achieving each possible calibration score and compares these to the expected distribution for participants accurately setting 80% confidence intervals on each question.

The clearest effect in Figure 3 is the general trend of overconfidence – the majority of participants scoring below 5/10, compared to less than 1% of participants who would be expected to do so if accurately setting 80% confidence intervals. Goodness-of-fit tests confirmed that participants, pre- and post-training, differed significantly from the expected distribution of results, \( \chi^2(10)=3.85\times10^5 \) and \( 1.62\times10^5 \), \( p<.001 \) and \( p<.001 \), respectively.

The data in Figure 3 also, however, show an improvement in calibration following training; the mean calibration score prior to training being 2.6 (SD=2.3) and rising to 4.7 (SD=2.6) after training. A repeated measures t-test indicated that this change was significant, \( t(50)=3.71, p<.001 \), offering support for Hypothesis 2, which holds that awareness of overconfidence will reduce susceptibility to the bias.

**Discussion**

**Overview**

The results are mixed in terms of evaluating the efficacy of awareness-style training in debiasing, with support found for only one of the three hypotheses examined herein: Hypothesis 2, which argued that overconfidence would be reduced by making the participants aware of the effect and methods of avoiding or reducing it. This result and the failure to find support for the remaining hypotheses are discussed in greater detail below.
Hypothesis 1: Debiasing Anchoring
As noted above, no evidence was found that anchoring susceptibility was reduced or that participants’ accuracy improved after being made aware of the anchoring bias and having a single technique described to attempt to avoid it.

There seem to be at least two possible explanations for this. The first, of course, is that the debiasing method described to the participant’s simply does not work. Logically, however, this seems unlikely. If, as described above, a person makes an estimate that has been biased by their viewing of an anchor – then moving that estimate away from the anchor should reduce the bias and lead to more accurate estimates, assuming that the adjustment does not overshoot the true value.

What seems more likely is that the participants used herein did not have sufficient knowledge of the subject about which they were asked for such a technique to be useful to them. As noted above, the technique assumes a hypothetical unbiased estimate constructed from a person’s other knowledge that is biased by the presence of the anchor. In the absence of such domain knowledge that enables a person to determine which side of the anchor the true value should lie on, then, adjusting one’s estimate away from the anchor is, in effect, random. The discussion of this possibility is taken up again in the discussion of Hypothesis 3, below.

Hypothesis 2: Debiasing Overconfidence
In contrast to the failure to reduce anchoring, clear evidence was found in support of the hypothesis that awareness of overconfidence and debiasing techniques would enable people to reduce their level of overconfidence.

The participants, after training, had mean calibration scores of 4.7, up from 2.6. While even post-training the participants remained markedly overconfident, this improvement represents the number of surprises (values falling outside the predicted range) falling by 21%. It is also interesting to note that, post-training, the student participants described herein outperformed an industry sample who scored an average of 4.2 on the same questions (Welsh, Begg, & Bratvold, 2005).

Needless to say, an improvement even a fraction of the size observed here in oil and gas industry predictions could, potentially, save billions of dollars per annum.

Hypothesis 3: Accuracy and Susceptibility
The failure of the data to support Hypothesis 3, which predicted that accuracy would be inversely related to susceptibility on the anchoring task, is the most surprising but also perhaps the most interesting of the results.

The extremely strong positive correlations between the accuracy and susceptibility measures indicate, against expectations, that participants who rely more heavily on the anchor ended up making more accurate estimates. This is interesting as it indicates the adaptive value of the anchoring and adjustment heuristic itself. That is, this result shows that it can be a rational strategy to rely on the anchor.

Given the observation made above - that the students may not have had sufficient knowledge of the field in which they were being asked questions to enable them to debias their responses - it follows that they had no realistic constraints on what their answer to the anchoring question might be. By comparison, the anchoring values used in the experiment were chosen such that they would seem reasonable to people with a degree of industry knowledge; specifically, being set at one-half and one-and-one-half times the researched value. In the absence of alternate knowledge, therefore, the anchoring value at least constrained the estimates of participants who relied on them to the correct order of magnitude while participants who ignored the anchors were able to give wildly wrong estimates. That is, given no alternate knowledge, reliance on the anchor improved a participant’s odds of giving an estimate close to the true value.

Clearly, this effect will be restricted to a subset of environments where the anchors being used are at least somewhat informative. It is, however, worth noting that the standard greater/less than format used in anchoring questions reflects environments in which this would hold - as cut-offs used for such questions are usually set within the response range considered credible by the question’s writer.

Further Research
The results suggest a number of interesting research directions. Firstly, although previous research has indicated relatively little difference between industry and non-industry personnel in terms of bias susceptibility (Welsh et al., 2005), assessment of the benefit accrued by industry personnel sitting a training course like the one described here still needs to be undertaken. In particular, such work needs to take in to account the lower starting level of overconfidence that such personnel have been observed to have (Welsh et al., 2005) – an obvious concern being that the relative improvement might be less for people whose overconfidence is already lower and whose confidence in their own judgments is higher.

Research is also needed to establish the relationship between knowledge and susceptibility to anchoring more clearly. As noted above, previous studies (e.g., Wilson et al., 1996) have allowed self-reporting of a participant’s level of knowledge. This needs to be corrected by measuring participants’ knowledge of a domain and then assessing their susceptibility to anchoring within that area. Then participants with varying degrees of knowledge could be introduced to the debiasing technique described above for the anchoring bias and the hypothesized relationship between knowledge and anchoring could be tested.

Finally, the benefits of debiasing for a wider range of known cognitive biases needs to be considered. While biases beyond anchoring and overconfidence are rarely spoken of in oil and gas, there are strong arguments that effects such as framing and availability are impacting industry decisions (Kruger & Evans, 2004; Pieters, 2004)
Conclusions

The results presented above offer some hope for oil and gas companies that are banking on better economic outcomes resulting from investment in awareness-style training for their staff. Were attendance at such training courses to provide a decrease in overconfidence comparable to that observed herein on the technical parameter estimations made by industry personnel, then the benefits in terms of improved economic forecasting would easily be in the billions of dollars per annum across the industry. That said, a number of caveats remain.

Firstly, while overconfidence was shown reduced by the training, the training used herein made use of at least one debiasing technique that industry courses do not and which, in fact, they specifically recommend against (see, e.g., Campbell et al., 2001); the use of best guesses. It is also worth noting that, while improvement was observed, the participants were still markedly overconfident post-training.

The concern regarding the content of current industry courses must also extend to anchoring and other biases. While anchoring, at least, is commonly mentioned, there is little evidence that specific debiasing techniques for it are currently being taught or that any benefit is obtained from simple awareness-style training. Other biases have simply not yet been raised to the industry’s consciousness.

Finally, the fact that the same question set was used to test participants both pre- and post-training may also strike some as problematic. It should, however, be kept in mind that the participants received no specific feedback on their answers to any of the questions prior to completing the post-training questionnaire. The training used general examples only, which argues against any improvement in accuracy having occurred simply as a result of the test-retest format. Differences between the conditions can thus safely be argued to reflect the effect of training.

To conclude, research into biases and debiasing techniques has the potential to benefit the oil and gas industry to the tune of billions of dollars per year but a great deal of work remains to be done before this could be realized. The industry’s current, awareness-style training, while having a demonstrable effect, is outdated and incomplete and would see an immediate benefit from the incorporation of more recent research on biases.

Acknowledgements

This research was made possible through ExxonMobil and Santos’ support of the Centre for Improved Business Performance at the Australian School of Petroleum, University of Adelaide. We thank Dan Navarro and three anonymous reviewers for their comments.

References