Multiple anchors and the MOLE: Benefits for elicitation

Marianne H. Clausen (marianne.clausen@adelaide.edu.au)
University of Adelaide, SA 5005, Australia

Matthew B. Welsh (matthew.welsh@adelaide.edu.au)
University of Adelaide, SA 5005, Australia

Abstract

Anchoring is a well-known, robust effect causing estimates to be biased towards previously seen values – regardless of their relevance. Reducing anchoring bias is important for optimizing estimation. Herein, we tested the MOLE (More-Or-Less Elicitation) tool's ability to limit the impact of anchors on estimates. In a direct elicitation task, 62 participants' best estimates correlated with anchor values at 0.27 whereas, when using the MOLE, this relationship disappeared (*r* = .02). Results also showed, however, that expertise reduces the impact of anchoring (*r* = -0.46). We conclude that use of the MOLE assists in avoiding anchoring and that this will be most helpful in areas of high uncertainty.

Keywords: anchoring; elicitation; accuracy; expertise; decision making; repeated judgement, MOLE.

In the face of uncertainty, industry and government often rely on estimates made by experts to guide decisions – converting their beliefs into useable, numerical forms via a range of processes labelled 'elicitation' (Wolfson, 2001). Such estimates are useful, but also prone to systematic errors, known as biases (Kahneman, Slovic, & Tversky, 1982). One particularly robust bias arises from 'anchoring': the tendency for people to base estimates on numbers to which they have recently been exposed, regardless of their relevance (Tversky & Kahneman, 1974). Tversky and Kahneman (1974) famously demonstrated this by asking people to estimate the percentage of African countries in the United Nations, after judging whether it was higher or lower than a supposedly randomly selected value (of 10 or 65%). The group who saw the initial value of 10 provided a median response of 25, while those seeing 65 provided a median response of 45. This and numerous other studies have demonstrated the effect of initial values on subsequent estimates (for a recent review, see Furnham & Boo, 2011).

There are two well-known theories for mechanisms underlying anchoring. First, Kahneman and Tversky (1974) describe it as an *anchoring and adjustment* process: the anchor providing a starting value from which a person adjusts until they reach a value that they believe is reasonable. Therefore, the process results in people selecting amongst those values closest to the anchor - from the range they consider feasible. The second theory of anchoring is 'selective accessibility', or *priming*, where the initial approach is to consider whether or not the anchor itself is the true value (Mussweiler & Strack, 1999). If the anchor is determined not to be the true value, a person will search their memory for relevant clues as to what the true answer might be, but the starting point of this search is determined by the region in which the anchor falls. Both theories have support in the literature and both may contribute to the robust tendency for estimates to be skewed towards an initial value (see, e.g., Furnham & Boo, 2011).

Anchoring and Expertise

Given our reliance on expert opinion to reduce our uncertainty in most industries and fields of endeavor, the extent to which anchoring is reduced by particular expertise is, logically, of considerable importance; biases are likely to reduce the accuracy of expert estimates and thus erode the quality of decisions made based on their opinions.

While it might seem reasonable that people with greater expertise would produce more accurate estimates - with the effect of anchoring minimised as they rely on experience and knowledge in preference to the anchor - there is mixed evidence regarding anchoring and expertise. Some studies have found that people higher in knowledge (Wilson et al, 1996) or those more certain of their responses (Chapman & Johnson, 1994) were less affected by anchoring but others have demonstrated anchoring in subject matter experts (Mussweiler, Strack & Pfeiffer., 2000; Northcraft & Neale, 1987). That is, the common observation is that expertise may reduce but does not eliminate bias arising from anchoring.

More-or-Less Elicitation

More-or-Less Elicitation (MOLE) is a computerised elicitation tool, which produces a range for the values an unknown parameter might take, in a manner designed to decrease both overconfidence in those ranges (i.e., 'overprecision' as per Moore & Healy, 2008) and the effects of anchoring (Welsh & Begg, 2018; Welsh, Lee, & Begg, 2008, 2009). The basic, MOLE process (described in detail in the Method selection, below) presents users with two random values. The user indicates which of these they believe is closer to the true value, with the process being repeated numerous times. The MOLE incorporates four key insights to achieve reduced overconfidence and anchoring.

First, instead of making absolute judgments, the MOLE allows people to make relative judgements, which have long been known to be both easier and more accurate (see, e.g., Stroop, 1932; Miller, 1956). Instead of directly asking for estimates, respondents simply select which of two presented values they believe is more likely/closer to the truth.

Second, the MOLE uses repeated judgements, the average of which have been shown to be more accurate than single estimates (see, e.g., Herzog & Hertwig, 2009, Vul & Pashler, 2008). This allows multiple pairs of values to be presented for each value being estimated, enabling repeated judgements of the same value while preventing participants simply repeating answers, thus meeting the criteria for useful, repeated judgements from an individual (see, e.g., Herzog & Hertwig, 2009, Vul & Pashler, 2008).

Third, the MOLE retains portions of the possible range that users are uncertain of. Values are only excluded when the respondent's judgements indicate they are *certain* that values are not possible. This contrasts with more typical elicitation processes where the respondent decides which values to include – a process that is likely to result in them stopping their search for values once they move beyond those they are certain should be included and is thus likely to underestimate their true uncertainty (an explanation that echoes the explanation for anchoring invoked by Kahneman, 2011).

Fourth (and key herein), the process is expected to reduce the effect of any 'anchor' seen prior to the elicitation process – simply because the MOLE requires that the respondent consider many randomly selected values, thereby preventing them from focusing on that anchor – in line with the advice on best practice for avoiding anchoring (see, e.g., Kahneman, Lovallo & Sibony, 2011). Additionally, the values presented by the MOLE – being randomly selected – are likely to include values that would not otherwise have been considered by the user. This strategy, of considering alternative or contradictory information has been shown to reduce anchoring (Mussweiler et al., 2000; Russo & Schoemaker, 1992). When considered in light of the two theories of anchoring introduced earlier, it appears that the MOLE should, therefore, overcome the effect of anchoring, regardless of the underlying mechanism. That is, the MOLE provides many different values from which a user could begin *adjusting,* and all of these values may *prime* the user to consider them as the true value. The presence of numerous potential anchors should, therefore, prevent bias caused by a focus on any singular value.

Previous studies have demonstrated the benefits of the MOLE in reducing overconfidence and improving estimates in perceptual, epistemic and forecasting tasks (Clausen, 2017; Welsh & Begg, 2018; Welsh et al., 2008, 2009). Welsh and Begg (2018) also demonstrated that the initial values presented by the MOLE itself do not act as anchors. However, the tool's efficacy in preventing anchoring on a specific, prior value has, as yet, not been directly tested.

Aims

The primary aim of this research is to expand upon previous studies (Welsh & Begg, 2018; Welsh et al., 2008, 2009), which have investigated the MOLE with regards to accuracy and overconfidence, by testing the MOLE's ability to overcome the effect of anchoring. Previous work (e.g., Welsh & Begg, 2018) has made the logical assumption that the MOLE overcomes the effect of anchors, yet this theory has not been directly tested. Secondly, the effect that anchoring has on the best estimates participants produce will be assessed - to confirm whether anchoring is having a detrimental effect. While a detrimental effect seems logical, anchors may provide clues and improve performance in the case of low subject matter knowledge. Finally, the relationship between participant knowledge and anchoring will be examined – adding to the literature on differential effects of anchoring resulting from subject matter expertise.

Method

Participants

The study recruited N = 62 participants (38 females and 24 males) aged between 18 and 65 years (M = 31.15, SD = 12.81). One participant was excluded from the study due to inappropriate responding. Participants were recruited within the University of Adelaide and from the wider population. Participation was, initially, encouraged by course credit (1st Year Psychology students, n = 14) or the opportunity to win a $100 gift card (n = 17). To speed recruitment, later participants were offered a $20 gift card (n = 31).

Materials

The study consisted of the various measures and elicitation tasks (described below) incorporated into a single program, created by the first author using Visual Basic for Applications (VBA) in Microsoft Excel.

Demographics Participants were asked to provide information about themselves, including age, gender and their engagement with Australian Rules Football (ARF). Engagement with ARF was measured using four questions, such as how often participants played ARF or watched ARF matches. Participants rated their frequency of engagement for each of four questions on a four-point Likert scale, with a response of one corresponding to 'rarely or never' and four corresponding to 'more than once a week'. Possible scores for engagement with ARF thus ranged between 4 and 16 with higher scores indicating greater engagement.

Individual Differences A number of measures were included in the study as possible covariates of performance on the task or anchoring susceptibility. Specifically: Need for Closure (Webster & Kruglanksi, 1994; linked to range width in elicitation tasks by Kaesler, Welsh & Semmler, 2016); the 2-item Openness and Conscientiousness measures from the Ten Item Personality Inventory (TIPI; Gosling, Rentfrow & Swann, 2003; Openness having been linked to anchoring susceptibility by McElroy & Dowd, 2007); and the full-scale Openness and Conscientiousness measures from the NEO-FFI (Costa & McCrae, 1992) for comparison. However, no significant results were obtained for these and, as a result, they are not discussed further herein.

Experimental Tasks The study utilised forecasting questions, specifically related to Australian Football League (AFL) match results (NB: Australian Rules Football – ARF - is the sport and AFL is the national league). AFL results were chosen as the forecasting measure because they provided a sufficiently large set of numerical values of similar difficulty to predict. Participants were asked to consider the total points that a team would score in a specific game occurring within the next two rounds of the competition. For example: "What will be the total number of points that the St Kilda Saints score when they play the North Melbourne Kangaroos on Sunday, the 20th of August?". Given the study was conducted in Adelaide, the Adelaide-based AFL teams were excluded to limit the impact of specialist knowledge. The study included two methods of eliciting both a best estimate and the range that the participant was confident that the true value would fall within – the MOLE and a direct elicitation task (described below). The study elicited responses from participants for five questions under each condition.

Anchoring Questions In each case, participants were initially presented with an anchoring question. The anchoring question asked if a specific team would score more/less than a given value (20, 90, 160, 230 or 300), which were linearly distributed over the MOLE's initial range of 0-300. For example, "Will the Collingwood Magpies score less than 160 total points when they play the Geelong Cats on Saturday, the 19th of August?", where '160' is the anchoring value.

Only after responding to this anchoring question would the participant be asked to answer the corresponding question: (e.g.) "What will be the total number of points that the Collingwood Magpies score when they play the Geelong Cats on Saturday, the 19th of August?" using one of the two methods described below.

MOLE At each of 10 iterations within the MOLE process, participants were presented with two values randomly selected from a pre-determined widest possible range (for this study, 0 to 300 points). Participants were then asked to indicate which they believed would be closer to the true value by adjusting a slider from the default centre position (indicating the participant believed the values presented were equally likely as per the user interface shown in Figure 1). Adjusting the slider to the extreme left indicated that the value displayed on the right was not deemed possible and vice versa, whereas positions between the centre and the end-points were mapped onto levels of confidence between 50% and 100% in the selected value being closer to the true value.

If a participant indicated maximum (100%) confidence in one of the values, the possible range was truncated at the unselected value. For example, given an initial range of 0 to 300, if the values 95 and 240 were presented and the participant selected 95 with maximum confidence, then this would result in the range being truncated at 240 and the options displayed in later iterations being drawn from the 0-240 range rather than 0-300. This follows, logically, from the descriptions within the GUI, where the ends of the confidence scale explicitly rule out the alternative value. (NB - previous studies using the MOLE truncated the range at the midpoint, rather than the unselected value. This study adopted a more conservative approach to range truncation in order to lessen the chance of errors resulting in unrecoverable, overly narrow ranges as had sometimes been observed in previous work; Welsh & Begg, 2018.)

Where a participant's confidence in their selected value was less than complete, however, the range remained unchanged as this was taken as evidence that the participant believed that the alternative value still had some chance of being closer to the true value.

The remaining range after ten iterations of the MOLE was recorded as the participant's 100% confidence range – that is, the range that they could be assumed to be 100% confident would contain the true value. At each iteration, the participant's estimate of the true value was calculated from their confidence and the difference between the values. For example, equal (50%) confidence in the two values 100 and 200 would produce a best estimate of 150 whereas a 90% confidence that 100 was closer to the true value than 200 produced an estimate of 110. A participant's overall best estimate was calculated by the MOLE simply as the average of these estimates from each iteration.

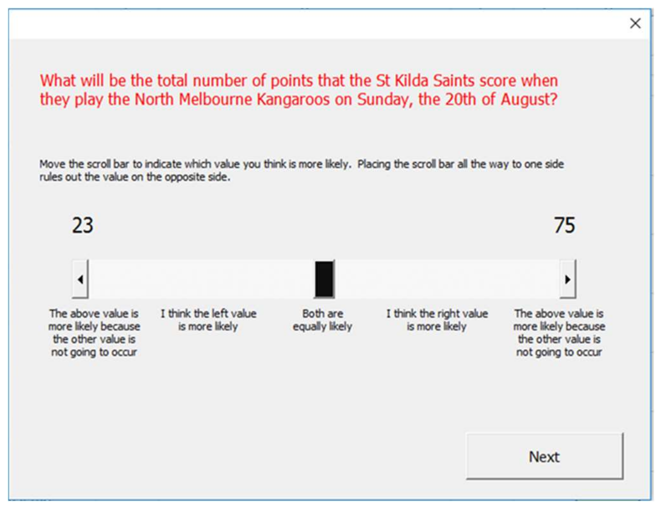

Figure 1: MOLE user interface

Direct Elicitation The direct elicitation task required participants enter a minimum value, a maximum value and a best estimate for the number of points they believed the specified team would score. The program checked that the minimum value was lower than the best estimate, which in turn had to be lower than the maximum; failure to meet these requirements resulted in the program displaying an error message, prompting the participant to revise their estimates. Otherwise, these direct estimates of the end-points of the range and the best estimate were simply recorded.

Procedure

After participants had registered their interest in the study, they were emailed a copy of the VBA program. The program included questions about the next two (weekly) rounds of AFL matches. Therefore, the specific questions included in the program changed on a weekly basis. Results were collected between the 5th of June 2017 and the 19th of

August 2017. The study used a within-participants design, with participants completing both elicitation tasks, allowing for comparison of performance across the tasks.

Participants were randomly assigned to complete either the MOLE or direct elicitation task first, with the intention of preventing order effects. While efforts were made to allocate participants equally to the conditions, the number of participants varied due to uptake and completion rates, with the end result that 33 participants completed the MOLE task first and 29 completed the direct elicitation task first.

The survey began with an information page, where participants indicated their agreement to participate. The survey progressed through, in order, demographic questions and individual difference measures, prior to beginning the elicitation tasks and concluded with instructions to save the file and return it to the researcher by email. Participants could only move forwards through the tasks, once a response was submitted it could not be altered.

Results

Dependent Measures

The dependent measures of the study were defined and or calculated for each participant within each elicitation condition as follows, with descriptive statistics presented in Table 1 alongside the independent measures age and Engagement (described in the Demographics section above).

Best Estimate. As described above, this was directly estimated by participants for each of the five questions in the Direct condition and calculated from their choices and confidence in the five MOLE questions.

Anchoring Score. The correlation between a participant's best estimates and the anchoring values across the five questions within each condition.

Accuracy Score. The correlation between a participant's best estimates and the true values across the five questions within each condition.

Error. The difference between a participant's best estimates and the actual value averaged across the five questions within each condition.

MOLE and Anchoring Given the typical anchoring effect, it was expected that there would be a positive correlation between anchor values and the best estimates participants produced via direct elicitation. Repeated measures correlation (r_{rm}) was used due to there being five data points for each participant. (Repeated measures correlation is equivalent to Pearson's correlation; however, it accommodates multiple data points per participant, increasing power, without violating the assumption of independence. It evaluates overall intra-individual relationships and can be calculated using a form of ANCOVA using the rmcorr package in R; see Bakdash & Marusich, 2017.) Analysis confirmed a small to medium, positive correlation, $r_{rm}(247) = .27$, $p < .001$, suggesting that higher anchor values did, in fact, result in higher best estimates.

Table 1: Descriptive statistics

Variable	Mean	Median	Min	Max	SD
Age	31.2	26.5	18	65	12.8
Engagement	7.6	8	4	16	2.9
Anchoring Score					
Direct	.14	.24	-.96	.99	.50
MOLE	.01	.01	-.96	.93	.51
Accuracy Score					
Direct	.02	-.10	-.92	.99	.54
MOLE	.17	.26	-.96	.94	.50
Average Error					
Direct	39.9	29.8	6.8	226.4	36.0
MOLE	46.3	36.9	8.1	120.4	26.6

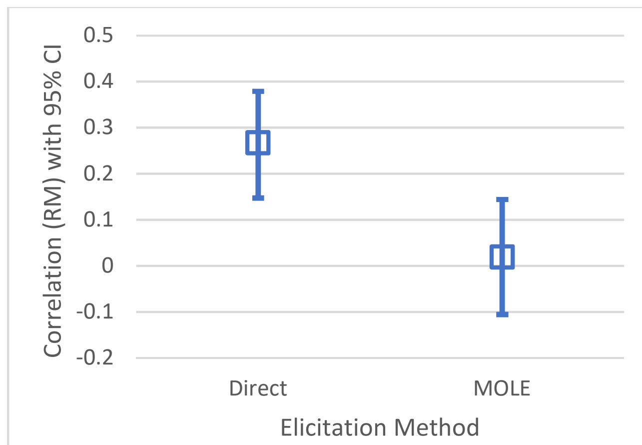

Figure 2: Correlations between anchors and best estimates by condition.

Additionally, it was hypothesised that anchoring would not be evident in the MOLE task. Repeated measures correlation analysis confirmed this with no correlation detected between MOLE best estimates and anchor values, $r_{rm}(247) = .02$, $p = .76$. The correlations and their confidence intervals are displayed in Figure 2, from which it can, from inspection of the CIs, be seen that two correlations also differ significantly from one another – providing further support for the linked hypotheses that anchors would affect best estimates obtained via Direct elicitation but not those generated by the MOLE.

Anchoring and Accuracy The extent to which anchors affected best estimates was assessed using two related measures: accuracy score and error. Firstly, within the direct elicitation condition, accuracy scores were expected to be negatively correlated with anchoring scores. Analysis confirmed a medium correlation of $r(60) = -.33$, $p = .01$. This suggests that participants who were more affected by anchors produced less accurate best estimates (when accuracy is defined as the correlation between best estimates and actual values, at the individual level). Secondly, error was expected to be positively correlated to anchoring score in direct elicitation. Analysis confirmed a large correlation of $r(60) = .44$, $p < .001$. This indicates that participants who were more affected by anchors produced best estimates that were further away from actual values. Overall, the results indicate that anchoring was associated with less accurate best estimates, regardless of whether accuracy was measured by the average error between best estimates and actual values, or the degree to which best estimates correlated with actual values. Finally, no significant correlations were seen between anchoring and accuracy in the MOLE condition, which is to be expected as anchoring was not evident in this condition.

Anchoring and Knowledge It was anticipated that greater knowledge (specifically, level of engagement with ARF) would be associated with less anchoring in the direct elicitation condition. Analysis confirmed a large negative correlation, $r(60) = -.46$, $p < .001$, as shown in the scatterplot below (Figure 3). That is, participants with greater knowledge of the subject were less affected by anchors.

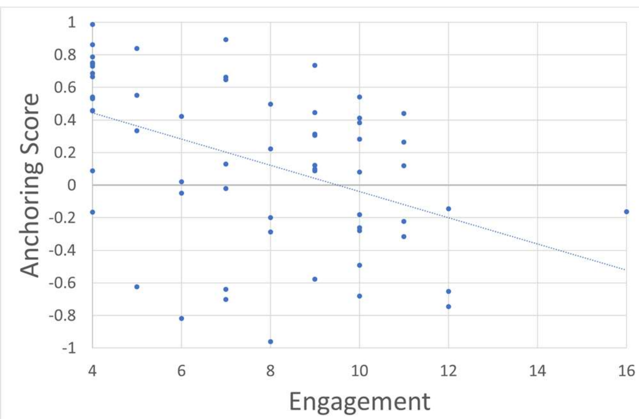

Figure 3: Scatterplot of anchoring score versus engagement with trendline.

Discussion

The results suggest More-or-Less Elicitation (MOLE) is successful in overcoming the effects of anchoring. In the direct elicitation condition, a typical anchoring effect was seen; that is, estimates correlating with the anchor values seen by participants. The same participants, however, showed no evidence of anchoring when using the MOLE, supporting the hypothesis that the large quantity of numbers presented to MOLE users prevents anchoring on any specific value.

In addition to detecting the presence of anchoring, the study demonstrated that anchoring was associated with less accurate estimates. This confirms the detrimental effect of anchoring on estimation.

The final component of the study was the investigation of the relationship between knowledge and anchoring. The positive correlation between accuracy and engagement and the large negative correlation between the degree to which participants engaged with the subject matter and their anchoring scores supports the common-sense expectation that those with greater expertise produce better estimates and are less likely to be led astray by anchors.

Caveats and Future research

The above conclusion is an encouraging finding for those depending on estimates made by experts, but it needs to be balanced with a key understanding. The processes of anchoring (adjustment and priming) both imply that a person's degree of uncertainty limits the effect that an anchor can have. Expert opinion is most valuable, however, where uncertainty is highest and typical elicitation processes may, thus, occur in situations more akin to lower levels of knowledge where anchoring remains a problem.

A related result requiring further comment is the negative slope of the trendline (see Figure 3), which supports the idea that knowledge reduces anchoring, but also implies that greater knowledge could be associated with negative anchoring – that is, more expert people reacting against the influence of the anchor. If averaged for higher engagement participants (Engagement ≥ 8), however, the anchoring values tend to be close to zero (Engagement ≥ 8, M = -.03, Min = -.96, Max = .73, SD = .43). Additionally, at the lowest level of engagement there is a high degree of anchoring (Engagement = 4, M = .58, Min = -.17, Max = .99, SD = .30). Thus, it seems likely that the study indicates only that a lack of knowledge is associated with anchoring, whereas greater knowledge is associated with less or no anchoring. Future research with a larger sample could, however, directly test the possibility of participant reactions against anchors.

Inspection of the scatterplot of engagement versus anchoring (see Figure 3) also indicates a cluster of participants at 4 for engagement. This reflects participants who responded "rarely or never" to all four engagement questions. In hindsight, this category would have been better divided into separate categories for "rarely" and "never" to obtain a better spread of data.

Table 1 indicates that average error is greater for the MOLE than for direct elicitation. This result seems to contradict the generally proposed benefits of the MOLE. However, while the MOLE estimates tend to be further away, they correlate more closely to actual values. This is likely due to the estimates being located within a much wider range of possible answers, with the wider ranges produced by the MOLE being the key mechanism in reducing overconfidence (Clausen, 2017). This highlights the benefit of calculating error in terms of both magnitude and trend. While this result is not detrimental to the current findings, future research could investigate improvements to the algorithm that calculates MOLE estimates.

Conclusions

Anchoring is a fundamental cause of bias in estimation and, as such, a central concern during the elicitation of expert opinion. The results support the idea that expertise can limit the impact of anchors on estimates - with more knowledgeable participants in the direct elicitation condition making estimates that were both more accurate and less affected by anchors. In contrast, estimates elicited using the MOLE showed no relationship to the anchoring values, regardless of participant expertise. That is, it seems the MOLE process of providing multiple, paired choices washes out the effect of any previously observed anchoring value in both knowledgeable and naïve estimators – in addition to reducing overconfidence as has been shown previously (Welsh & Begg, 2018). Given this, the MOLE seems to provide a better alternative to direct elicitation in situations where anchoring and overconfidence biases are of concern.

Acknowledgments

The data in this paper are a subset of those collected for MHC's 2017 Honours thesis in Psychology at the University of Adelaide. MBW's position is supported by Australian Research Council Grant LP160101460, which includes support from industry partners Santos and Woodside.

References

Bakdash, J. Z., & Marusich, L. R. (2017). Repeated measures correlation. *Frontiers in psychology, 8*, 456.

Chapman, G.B., Johnson, E.J., 1994. The limits of anchoring. *Journal of Behavioral Decision Making 7,* 223–242.

Clausen, M. H. (2017). Overconfidence and the MOLE: Investigating the role of anchoring and individual differences. Unpublished Psychology Honours thesis, University of Adelaide.

Costa, P. T., & MacCrae, R. R. (1992). *Revised NEO personality inventory (NEO PI-R) and NEO five-factor inventory (NEO-FFI): Professional manual*. Psychological Assessment Resources, Incorporated.

Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics, 40*(1), 35-42.

Gosling, S. D., Rentfrow, P. J., & Swann Jr, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in personality*, *37*(6), 504-528.

Herzog, S. M. and R. Hertwig (2009). The wisdom of many in one mind: improving individual judgments with dialectical bootstrapping. *Psychological Science 20*(2): 231-237.

Kaesler, M., Welsh, M. B., & Semmler, C. (2016). Predicting overprecision in range estimation. In A. Papafragou, Grodner, D., Mirman, D., & Trueswell, J.C. (Ed.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Kahneman, D. (2011). *Thinking, Fast and Slow*. New York, NY: Farrar, Straus, Giroux.

Kahneman, D., Lovallo, D., & Sibony, O. (2011). Before you make that big decision. *Harvard Business Review, 89*(6), 50-60.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgement under uncertainty: heuristics and bias*. Cambridge: Cambridge University Press.

McElroy, T., & Dowd, K. (2007). Susceptibility to anchoring effects: How openness-to-experience influences responses to anchoring cues. *Judgment and decision making, 2*(1), 48-53.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review 63*: 81-97.

Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review, 115*(2), 502-517.

Mussweiler, T., & Strack, F. (1999). Hypothesis-consistent testing and semantic priming in the anchoring paradigm: A selective accessibility model. *Journal of Experimental Social Psychology*, *35*(2), 136-164.

Mussweiler, T., Strack, F., & Pfeiffer, T. (2000). Overcoming the inevitable anchoring effect: considering the opposite compensates for selective accessibility. *Personality and Social Psychology Bulletin, 26*(9), 1142-1150.

Northcraft, G. B., & Neale, M. A. (1987). Experts, amateurs and real estate: an anchoring-and-adjustment perspective on property pricing decisions. *Organizational Behavior and Human Decision Processes, 39*, 84-97.

Russo, E. J., & Schoemaker, P. J. H. (1992). Managing Overconfidence. *Sloan Management Review, 33*, 7-17.

Stroop, J. R. (1932). Is the judgment of the group better than that of the average member of the group? *Journal of Experimental Psychology 15*(5): 550-562.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124-1131.

Vul, E. and H. Pashler (2008). Measuring the crowd within: probabilistic representations within individuals. *Psychological Science 19*(7): 645-647.

Webster, D. M., & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology, 67*(6), 1049.

Welsh, M. B. & Begg, S. (2018). More-Or-Less Elicitation (MOLE): reducing bias in range estimation and forecasting. *EURO Journal on Decision Processes*, https://doi.org/10.1007/s40070-018-0084-5.

Welsh, M. B., Lee, M. D., & Begg, S. H. (2008). More-or-Less Elicitation (MOLE): Testing a heuristic elicitation method. In V. Sloutsky, B. Love, & K. McRae (Eds.), *30th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Welsh, M. B., Lee, M. D., & Begg, S. H. (2009). Repeated judgments in elicitation tasks: efficacy of the MOLE method. In N. Taatgen, H. v. Rijn, L. Schomaker, & J. Nerbonne (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Wilson, T. D., Houston, C. E., Etling, K. M., & Brekke, N. (1996). A new look at anchoring effects: basic anchoring and its antecedents. *Journal of Experimental Psychology: General, 125*(4), 387-402.

Wolfson, L. J. (2001). Elicitation of probabilities and probability distributions. *International Encyclopedia of the Social Sciences*. E. Science, Elsevier Science: 4413-4417.