Does Anything Predict Anchoring Bias?

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

Abstract

Anchoring – the tendency for recently seen numbers to affect estimates – is a robust bias affecting expert and novice judgements across many fields. An anchoring task, in which people (N=301) estimated the number of circles in 10 stimulus figures after comparison to an anchoring value, was conducted within a larger study including numerous intelligence, personality, decision style and attention measures. Individual anchoring susceptibility was calculated and compared to potential predictor variables. Two of eight broad ability measures (from Catell-Horn-Carroll intelligence theory) correlated weakly but significantly with anchoring (Gq = 0.16, Gf = 0.12). No decision style or attention measures correlated significantly with anchoring, nor did the Big 5 personality traits, directly. Indirectly, however, as the anchoring task continued and fatigue increased, people relied more on anchors and higher neuroticism may have increased this tendency. Overall, results suggest our ability to predict anchoring is poor and implications of this are discussed.

Keywords: anchoring bias; intelligence; personality; decision styles; attention.

Introduction

Anchoring-and-adjustment (hereafter ‘anchoring’) describes the tendency for people’s numerical estimates to be biased towards recently seen numbers – regardless of the relevance of those numbers to the task at hand (Tversky & Kahneman, 1974). This effect is robust, affecting judgements across a wide range of tasks and domains (see, e.g., Chapman & Johnson, 1994; Furnham & Boo, 2011) and resisting many efforts at debiasing (see, e.g., Mussweiler, Strack, & Pfeiffer, 2000; Wilson, Houston, Etling, & Brekke, 1996).

This effect has significant practical implications, affecting experts judgements in: pricing and negotiations (Northcraft & Neale, 1987); technical estimates under uncertainty (Welsh & Begg, 2016); and forecasting (Lawrence & O'Connor, 1992). Given this, there is significant utility in predicting susceptibility to anchoring – to enable identification of people in need of extra support to avoid the effect or selection of people more resistant to the bias. The literature, however, shows inconsistent results, perhaps because the application of individual differences research to the study of anchoring has been somewhat piecemeal.

Consider, for example, the Big 5 personality traits (Costa & McCrae, 1992). McElroy and Down (2007), starting with the hypothesis that openness-to-experience includes a greater willingness to take on outside influences, found evidence that people with higher openness were more susceptible to anchoring. Eroglu and Croxton (2010), however, linked anchoring with high conscientiousness and agreeableness and with low extraversion. Again, this has some face validity, in terms of more conscientious people perhaps being more inclined to duly consider the anchor and thus be affected by it, or more agreeable people being more willing to go along with it as a suggested answer, but this paper also specifically failed to replicate the finding relating openness to anchoring.

Given anchoring can include both a priming effect, altering the information called from memory, and an adjustment process in which iterative changes are made to an estimate (Kahneman, 2011), it also seems plausible that it would be reduced in people with better cognitive abilities – such as memory, numerical or reasoning skills.

Links with intelligence, however, have been weak. Oechssler, Roider and Schmitz (2009), for example, found no evidence of higher cognitive ability predicting lessened anchoring whereas Bergman et al (2010) found evidence for it reducing but not eliminating the effect. Welsh et al (2014) found suggestive evidence for it reducing susceptibility to anchoring and somewhat stronger evidence for a rational decision style and cognitive reflection (Frederick, 2005) doing the same - but only as a result of learning across the duration of an extended task in both cases.

Part of the problem here, though, as Welsh, Burns and Delfabbro (2013) point out, is that most tests of cognitive ability used in bias research are simple proxies for the far more complex hierarchical intelligence suggested by modern Catell-Horn-Carroll theory (CHC; McGrew, 2009) and selection of personality traits and decision style measures is often based on convenience and simplicity due to the time and cost of conducting complete measures.

Given this, a gap exists for a more complete look at how anchoring relates to the wide range of human abilities falling under the individual differences umbrella. This will enable us to state with greater certainty whether individual traits can predict susceptibility to this bias and, thus, the extent to which less biased personnel can be selected or identified for roles in which anchoring is likely to have significant economic or human consequences.

A wide-reaching exploration of individual differences also has the potential to shed light on the processes that give rise to the anchoring effect. For example, the CHC model of intelligence includes ten broad cognitive abilities that correspond to different aspects of memory, types of reasoning and cognitive speed. If anchoring is linked to these, this could indicate fruitful paths for further research into the cognitive processes underlying the bias.

Method

Participants

Participants were 301 university students, graduates and members of the general public recruited as part of a larger study; 172 identifying as female, 120 male and 9 non-binary. Participants listed 34 countries-of-origin but the majority (207) indicated English was their first language. Most were university educated, including 26 with post-graduate degrees, 101 with bachelor degrees (including 38 current post-grads) and 107 current university students. Ages ranged from 18 to 79 (M = 28.8, SD = 12.8).

Materials

Demographics

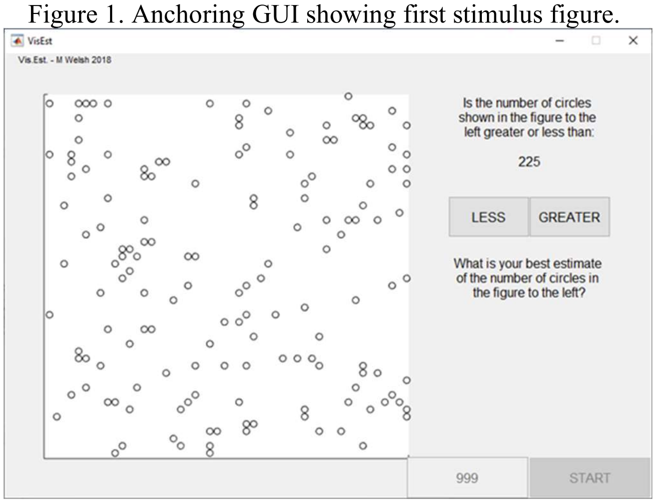
Demographics were gathered via online survey prior to online and in-person individual differences and experimental measures. These included: age; gender; native vs non-native English speaker; and educational attainment - recorded via options ranging from ‘Did not complete High School’ to ‘Doctorate’.

Anchoring

The anchoring task was written as a graphical user interface (GUI) in Matlab. The program showed participants stimuli consisting of a number of small circles. Pre-task instructions were as follows:

When the program begins, you will be shown a series of patterns of circles drawn here. For each, your task is to decide whether the number of circles is less or greater than a number you will be shown and then to give your best estimate of the number of circles. There is no time limit but the task is to estimate – not count – the circles. After a participant ID has been entered, press START when you are ready to see the first figure.

Figure 1. Anchoring GUI showing first stimulus figure.



On starting the task, participants were shown a stimulus figure as shown in Figure 1 and asked to indicate whether the number of circles was greater or less than the anchoring number shown before entering their best estimate.

Participants completed 10 trials, consisting of 5 pairs of figures with 150, 200, 250, 300 and 350 circles. One member of each pair was shown with a low anchor (50% of the actual value) and the other a high anchor (150%). All participants completed the tasks in the order shown in Table 1 but circle locations were randomly generated in each new stimulus figure.

Table 1. Order of stimuli and anchors (H-high or L-low)

1: 150H	2: 250L	3: 200H	4: 250H	5: 350L
6: 150L	7: 200L	8: 300L	9: 300H	10: 350H

Anchoring scores were calculated from the difference between estimates on the high and low anchor pairs, adjusted for the actual number of circles in the stimulus figure. That is:

$$\text{Anch} = (\text{Est}_{\text{HighAnchor}} - \text{Est}_{\text{LowAnchor}}) / \text{True \# of circles}$$

Higher numbers thus reflect greater impact of the anchors. The average value across the 5 pairs was taken as a participant’s final anchoring susceptibility score.

Cognitive Abilities

Tasks reflecting 8 broad abilities from the Cattell-Horn-Carroll model of intelligence (CHC; McGrew, 2009) were included. Six of these abilities were represented by 3 tasks each, allowing a factor score to be extracted using PAF with direct oblimin rotation in SPSS (NB – in each case a single factor was returned using eigenvalues>1). The remaining two were represented by single tasks, as described below.

Gf. Fluid ability extracted (KMO = .666, Bartlett’s Test of Sphericity $\chi^2(3) = 218$, p<.001) from 12-item Ravens APM (Arthur Jr & Day, 1994), CAB-I (Hakstian & Bennet, 1977) and WJ-IV Number Series (this and all subsequent measures labelled WJ-IV are from the Woodcock-Johnson IV Tests of Cognitive Abilities, Schrank & Wendling, 2018).

Gc. Crystallized ability extracted (KMO = .668, Bartlett’s Test of Sphericity $\chi^2(3) = 243$, p<.001) from the Mill-Hill Vocabulary Scale (Raven & Court, 1998), Spot-the-Word (Baddeley, Emslie, & Nimmo-Smith, 1993) and WJ-IV Oral Vocabulary.

Gsm. Short term memory extracted (KMO = .669, Bartlett’s Test of Sphericity $\chi^2(3) = 173$, p<.001) from WJ-IV Numbers Reversed, Memory Span Forward and Memory Span Backwards. The memory span tasks were written for this project in Matlab. Each displayed numbers of increasing length – from 1 to 10 – presented one digit at a time at 1 second intervals. After presentation, participants were asked to enter the digits in either the order presented or reversed order. Scores were the number of digits correctly recalled out of the total of 55 (10+9+8+… +1) from each task.

Glr. Long term retrieval ability extracted (KMO = .612, Bartlett’s Test of Sphericity $\chi^2(3) = 239$, p<.001) from WJ-IV Rapid Picture Naming, WJ-IV Retrieval Fluency and a Comprehension task. The comprehension tasks was written in Matlab for this project and required participants to read two ~500 word passages about historical events and answer four multiple (4) choice questions after each. Scores were the number of questions correctly answered out of eight.

Gq. Quantitative ability extracted (KMO = .642, Bartlett’s

Test of Sphericity $\chi^2(3) = 113$, p<.001) from the 12-Item Numerical Aptitude Test (Welsh et al., 2013), Berlin Numeracy Test (Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero, 2012) and Subjective Numeracy Test (Fagerlin et al., 2007).

Gt. Decision Speed extracted (KMO = .613, Bartlett's Test of Sphericity $\chi^2(3) = 136$, p<.001) from Inspection Time (Preiss & Burns, 2012), Simple Reaction Time and Go-No Go Reaction Time tasks written for this project in Matlab. The SRT asked participants to press a key as soon as a red 'R' appeared onscreen for 10 trials. The Go-No Go required responses only to the letter 'E', which occurred on 10 out of 100 trials (ten each of each letter from A to J, presented to participants in the same, randomized order.

Gs. WJ-IV Letter-Pattern.

Gv. WJ-IV Visualization.

All measures were scored such that higher numbers indicated better performance.

Personality Traits

The Big 5 personality traits were measured using the full version of the NEO-PI3 personality test (McCrae, Costa, & Martin, 2005). This yields 30 facets scales in addition to the five main traits.

Decision Styles

Six decision style measures were included. These were: a cognitive reflection test (CRT) incorporating both Frederick's (2005) and Thomson and Oppenheimer's (2016) questions (NB - whether CRT is a decision style, a cogntiive measure or a combination of the two is still debated in the literature, see, e.g., Welsh et al., 2013); the Rationality and Intuition scales from the Decision Styles Scale (Hamilton, Shih, & Mohammed, 2016); Need for Cognitive Closure (Webster & Kruglanski, 1994); Actively Open-Minded Thinking (Haran, Ritov, & Mellers, 2013); and the Brief Maximization Scale (Nenkov, Morrin, Schwartz, Ward, & Hulland, 2008), which distinguishes between preferences for satisficing versus optimization.

Attention Measures

Six tasks were included in the in-person tasks measuring the following aspects of attention: focused attention; sustained attention; selective attention; alternating attention; divided attention; and inhibition (defined as per Welsh, 2020a). These overlap the decision speed measures to some extent, being derived from a set of reaction time tasks, but additionally incorporate errors of omission and commission and changes in RT within and across conditions and tasks.

Procedure

The anchoring task and other measures described above were part of a larger study on biases. Participants were recruited online and on campus but needed to be available to come to campus for testing sessions. The testing was conducted in four parts and participants were paid $100 on completion. The first was an online survey that combined participant information, consent and demographic data collection. Four-hundred and four participants complete this initial survey.

Following this, participants invited to a second survey including: the personality and decision style questions; a subjective attention scale; the subjective numeracy scale; and the spot-the-word test. On average, participants spent 55 minutes on these tasks.

On completion, participants were invited to a third survey, which included: the cognitive reflection test; the cognitive abilities measures not conducted in person (see below); and a wide variety of bias tasks not discussed herein (e.g., framing, outcome bias, gamblers fallacy, sample size invariance, etc). Participants took around 2 hours to complete these tasks.

Finally, participants were invited to a 2 hour, in-person session, during which they completed the remaining cognitive ability measures including: the WJ-IV tasks; the numerical abilities test; the comprehension task; the memory span tasks; and a series of computerised tasks yielding the inspection time, reaction time and attention measures. The anchoring task was completed as part of the computerised tests noted above. Additional bias tasks were also included at the end of this session, examining tendencies such as hindsight bias and susceptibility to the availability effect. Overall, 301 of the 404 participants who signed up online completed the testing.

Results

Anchoring Bias

The mean anchoring score was 0.18 – indicating that estimates made after seeing the high anchor were 18% (of the true value) higher than those made after seeing the low anchor on equivalent stimuli. The overall effect of Anchoring bias was tested for using a single sample t-test, comparing participants' anchoring scores with the value of zero expected if anchors had no effect on estimates. This confirms a significant impact of anchors on participants' estimates, $t(300) = 18.46$, p<.001, $CI_{95\%}$ [0.161, 0.200].

Demographics

Participant demographics were examined to test whether any of these mediated anchoring susceptibility. Age did not correlate significantly with anchoring, r(299) =.059, p = .309. Likewise, an independent samples t-test indicated native English speakers did not differ from non-native in anchoring susceptibility, t(299) = .89, p = .369.

Female participants had a slightly higher mean anchoring scores (M = .191) than males (0.165) or non-binary participants (.164) but a one-way ANOVA indicated these differences were not significant F(2, 298) = 0.83, p = .437.

The effect of educational level was examined using a one-way ANOVA after dividing the participants into 5 groups – high school educated, post-secondary diplomas and certificates, university students, university graduates and post-doctoral graduates. The mean anchoring scores for these groups ranges from 0.167 (diplomas, etc) to 0.187 (postgraduate degrees) but the test indicated no significant differences between the means, $F(4,296) = .142$, $p = .966$.

Cognitive Ability

Pearson correlations calculated between anchoring scores and the eight cognitive ability scores are shown in Table 2.

Table 2. Correlations between anchoring and cognitive abilities.

	Gf	Gc	Gsm	Glr
Anch	-.119*	-0.098	-0.087	-0.001
	Gq	Gt	Gs	Gv
Anch	-.162**	-0.070	-0.063	-0.064

*- sig at .05 (2-tailed). **- sig at 0.01 (2-tailed). N=297-301.

Looking at Table 2, all eight cognitive abilities are negatively correlated with anchoring scores – i.e., suggesting higher cognitive abilities lessen anchoring susceptibility. Only the correlations with Gf and Gq, however, are significant and these are very weak relationships. (NB - a Bonferroni correction for family-wise alpha accounting for 8 comparison would result in only the Anchoring-Gq correlation remaining significant, p=.04, two-tailed, but given the overall pattern of results, known relationships between the cognitive measures and two significant results out of eight, such a correction seems unnecessary).

Personality Traits

Pearson correlations calculated between anchoring scores and the Big 5 personality traits are shown in Table 3.

Table 3. Correlations between anchoring and Big 5 personality traits (N=301).

	N	E	O	A	C
Anch	-0.074	-0.060	-0.064	0.056	0.076

Looking at Table 3, the personality traits show no clear relationship to anchoring, with all correlations less than ±0.1 and none significant. Further examination using the 30 facets showed four weak, significant relationships with anchoring: Neuroticism-Depression (r = -.113, p=.049); Openness-Fantasy (r= -.139, p=.016); Openness-Ideas (r= -.113, p=.05); and Conscientiousness-Deliberation (r= .155, p=.007). Correcting the family-wise alpha values for 30 comparisons, however, left none of these as significant.

Decision Styles

Pearson correlations were calculated between anchoring scores and the decision style measures. Given the relationships between the different decision style measures are less well-known than those between the cognitive and personality traits, the full correlation matrix is shown in Table 4 - rather than just the relationships between the decision styles measures and anchoring.

Looking at Table 4, one sees that all correlations with anchoring are less than ±0.1 and none are significant. By contrast, there is evidence that the different decision styles measures are related, with several moderate relationships between them. Interestingly, in light of debate about what the CRT measures, it shows the weakest relationships to the other decision style measures. (Additional analyses showed it correlated significantly with all of the intelligence measures - including r = 0.59 with Gf and r = 0.45 with Gq, p<.001 in both cases.)

Attention Measures

Correlations coefficients were calculated for twenty-one measures extracted from the six attention tasks with anchoring, as shown in Table 5.

Looking at Table 5, ones sees all correlations are less than ±0.1 and none are significant, showing no evidence of relationships between attention and anchoring susceptibility.

Changes in Anchoring

In light of previous work finding changes in anchoring across an experiment (Welsh et al., 2014), an examination of anchoring across the 10 trials was also conducted. For this, an anchoring reliance score was calculated for each trial as the absolute difference between the anchor and the estimate, adjusted for the magnitude of the estimated value (i.e., |A-E|/E), with lower scores thus indicating greater reliance.

To see if people's reliance on the anchor changed across the task, average anchor reliance scores were calculated for each participant on the first 5 and last 5 trials on the overall task. These were compared using a Wilcoxon sign-rank test, revealing reliance on the anchor values increased as the task went on, $Median_{1-5} = 0.662$, $Median_{6-10} = 0.445$, $z(300) = -11.38$, $p < .001$.

Changes in reliance were largely unrelated to the demographic, cognitive ability, decision style and attention measures. Of the personality traits, only Neuroticism correlated significantly with the change in reliance, r = -.123, p = .033 (two-tailed). That is, more neurotic people showed greater increases in reliance across the task.

Table 4. Correlations between anchoring and decision style measures.

	1	2	3	4	5	6	7
1.Anch		-.09	.07	-.01	-.03	-.02	-.04
2.CRT	.103		***.12***	***-.11***	-.03	***-.17***	-.10
3.DSR	.235	**.039**		***-.24***	***.16***	.03	***.20***
4.DSI	.808	**.049**	**<.001**		***.13***	***.31***	.06
5.NFCC	.659	.596	**.005**	**.021**		***.13***	***.33***
6.AOT	.752	**.004**	.649	**<.001**	**.020**		***.29***
7.BMax	.469	.084	**.001**	.287	**<.001**	**<.001**	

Note: top triangle shows correlation coefficients; bottom triangle shows p-values (2-tailed). Sig. results are bolded. N=301 except for CRT where N=300.

Table 5. Correlations between anchoring score and attention measures.

	Foc. RT	Foc. Err.	Sus. RT	Sus. Err.	Sus. Corr.	Sel. dRT	Sel. dErr
r	.091	.054	.030	.014	-.007	-.028	-.012
	Alt. dRT	Alt. dErr	Div. RT Corr.	Div. Corr.	Div. RT Err.	Div. Err.	Inh. RT Corr.
r	.067	-.083	-.015	-.097	-.057	.036	-.026
	Inh. Corr.	Inh. RT Err.	Inh. Err.	Sus. dRT	Inh dRT	Div. dRT	All Err.
r	-.035	.013	.035	-.059	-.063	-.054	-.022

Note: N=259-301. No correlations are significant at the .05 level. Foc. = focused, Sus. = sustained, Sel. = selective, Alt. = alternating, Div = divided, Inh. = inhibition. RT = reaction time, Err. = number of errors, Corr. = number correct, dRT = change in RT, dErr. = change in # errors.

Discussion

The above results paint a clear, if bleak, picture: even with a wide range of personal traits and a sizeable sample, our ability to predict susceptibility to anchoring is poor. None of the personality, decision style or attention measure considered herein showed any significant relationship to people's susceptibility to anchoring.

The only two traits that correlated significantly with anchoring susceptibility in our experiment were two of the eight intelligence measures – quantitative intelligence (Gq) at ~0.16 and fluid intelligence (Gf) at ~0.12. That these would predict performance does seem reasonable – as they reflect a person's ability with numbers and new problems respectively – and the anchoring task was a novel, numerical estimation task. The relationships are weak, however, offering little predictive power for those interested in personnel selection or identification and, given they correlate with one another at 0.6 in our sample, there is no additional predictive power from considering both.

This is, at once, disappointing and not entirely surprising. Anchoring has long stood somewhat apart from other biases – an observation recently confirmed in Ceschi et al's (2019) taxonomy wherein anchoring emerged as distinct from all 16 other biases they examined.

This seems to reflect a fundamental difference in the underlying processes involved. While a case can be made that a number of other biases are related to memory processes and thus to mnemonic aspects of intelligence (e.g., hindsight bias, overconfidence, availability. See, e.g., Welsh, 2018; Welsh, 2020b) or emerge from simple logical/numerical problems, anchoring is more complex. The discussion on the root cause of anchoring has focused on whether it is a conscious adjustment process (see, e.g., Tversky & Kahneman, 1974) or a priming/activation effect linked to confirmatory search (Chapman & Johnson, 1999).

In either case, the underlying processing involves consideration of values for their adequacy – which sounds like it should involve Gq and Gf (and memory in the form of domain specific knowledge) but would allow a wide variety of potential strategies to be used. This lack of constraints on how to solve the problem could enable a much greater diversity of outcomes from people with the same level of ability as differences in performance could be caused by differences in the adopted strategy rather than ability per se.

Caveats and Future Research

A key result, however, needs to be discussed in greater detail as it may impact the interpretation of the experiment. This is the observation that reliance on anchoring values increased during the task. The simplest explanation of this is that participants were getting increasingly tired or unhappy, making them more susceptible to bias as the trials continued (see, e.g., Danziger, Levav, & Avnaim-Pesso, 2011; Englich & Soder, 2009).

If this is correct, the fact that higher neuroticism may exacerbate this – correlating with increased reliance on the anchor as the task proceeded ($r = -0.123$) - is unsurprising as this trait measures emotional stability.

It does, however, indicate a potential confound in this experiment's design – where the anchoring task – in addition to being included near the end of a two-hour session incorporating multiple tasks, required participants to complete a series of trials and may, thus, have been doubly taxing. (NB – this may also hold true for the attention tasks, which occurred around the same time and also required multiple trials.)

The relationship seen here with anchoring reliance and neuroticism indicates the potential for interaction effects between the durations of the task and the experiment and specific traits that could obscure relationships of greater interest. For example, the role of intelligence, openness or decision style traits could be larger in less fatigued participants. (Of course, the relevance of this should be viewed through the lens of personnel selection and the specific tasks being selected for – i.e., whether the role requires estimates to be made continuously over an extended period or more occasionally.)

This could be tested in future research, with the results herein shedding light on the most promising predictors of anchoring to test in a shorter format – Gq and Gf. Another open question, given the discussion above, is whether anchoring susceptibility will remain stable across domains.

Conclusions

Despite the use of best-practice measures for individual differences including the CHC model of intelligence, the Big 5 personality measures and a range of attention and decision style measures, we did not find any traits that predict anchoring strongly. In fact, only two measures – quantitative (Gq) and fluid (Gf) intelligence from the CHC model - showed significant relationships with anchoring and both of these were weak (~0.16 and ~0.12).

These relationships may, however, have been deflated by fatigue, given the length of the experiment and the observation that participants were increasingly relying on the anchors as the task continued (and high neuroticism may have increased this somewhat). Future research should focus on these traits in smaller experiments where fatigue can be eliminated as a covariate.

References

Arthur Jr, W., & Day, D. V. (1994). Development of a short form for the Raven Advanced Progressive Matrices Test. *Educational and Psychological measurement, 54*(2), 394-403.

Baddeley, A., Emslie, H., & Nimmo-Smith, I. (1993). The Spot-the-Word test: A robust estimate of verbal intelligence based on lexical decision. *British Journal of Clinical Psychology, 32*(1), 55-65.

Bergman, O., Ellingsen, T., Johannesson, M., & Svensson, C. (2010). Anchoring and cognitive ability. *Economics Letters, 107*(1), 66-68.

Ceschi, A., Costantini, A., Sartori, R., Weller, J., & Di Fabio, A. (2019). Dimensions of decision-making: An evidence-based classification of heuristics and biases. *Personality and Individual Differences, 146*, 188-200.

Chapman, G. B., & Johnson, E. J. (1994). The limits of anchoring. *Journal of Behavioral Decision Making, 7*(4), 223-242.

Chapman, G. B., & Johnson, E. J. (1999). Anchoring, activation, and the construction of values. *Organizational Behavior and Human Decision Processes, 79*(2), 115-153.

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin numeracy test. *Judgment and Decision Making, 7*(1), 25.

Costa, P. T. J., & McCrae, R. R. (1992). *NEO PI-R Professional Manual*. Odessa, FL: Psychological Assessment Resources.

Danziger, S., Levav, J., & Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences, 108*(17), 6889-6892.

Englich, B., & Soder, K. (2009). Moody experts---How mood and expertise influence judgmental anchoring. *Judgment and Decision Making, 4*(1), 41.

Eroglu, C., & Croxton, K. L. (2010). Biases in judgmental adjustments of statistical forecasts: The role of individual differences. *International Journal of Forecasting, 26*(1), 116-133.

Fagerlin, A., Zikmund-Fisher, B. J., Ubel, P. A., Jankovic, A., Derry, H. A., & Smith, D. M. (2007). Measuring numeracy without a math test: Development of the subjective numeracy scale. *Medical Decision Making, 27*(5), 672-680.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*(4), 25-42.

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

Hakstian, A. R., & Bennet, R. W. (1977). Validity Studies Using the Comprehensive Ability Battery (CAB): 1. Academic Achievement Criteria. *Educational and Psychological measurement, 37*(2), 425-437.

Hamilton, K., Shih, S.-I., & Mohammed, S. (2016). The development and validation of the rational and intuitive decision styles scale. *Journal of Personality Assessment, 98*(5), 523-535.

Haran, U., Ritov, I., & Mellers, B. A. (2013). *The role of actively open-minded thinking in information acquisition, accuracy, and calibration*. ScholarlyCommons, (5-2013).

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

Lawrence, M., & O'Connor, M. (1992). Exploring judgemental forecasting. *International Journal of Forecasting, 8*(1), 15-26.

McCrae, R. R., Costa, J., Paul T, & Martin, T. A. (2005). The NEO–PI–3: A more readable revised NEO personality inventory. *Journal of Personality Assessment, 84*(3), 261-270.

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.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 37*(1), 1-10. doi:https://doi.org/10.1016/j.intell.2008.08.004

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.

Nenkov, G. Y., Morrin, M., Schwartz, B., Ward, A., & Hulland, J. (2008). A short form of the Maximization Scale: Factor structure, reliability and validity studies. *Judgment and Decision Making, 3*(5), 371-388.

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.

Oechssler, J., Roider, A., & Schmitz, P. W. (2009). Cognitive abilities and behavioral biases. *Journal of Economic Behavior & Organization, 72*(1), 147-152.

Preiss, A. K., & Burns, N. R. (2012). Accurately measuring inspection time with computers. *International Journal of Intelligence Science,* 2(4). doi:DOI:10.4236/ijis.2012.24013

Raven, J. C., & Court, J. H. (1998). *Raven's progressive matrices and vocabulary scales* (Vol. 759): Oxford pyschologists Press Oxford, England.

Schrank, F. A., & Wendling, B. J. (2018). *Woodcock–Johnson IV Tests of Cognitive Abilities Examiner's Manual, Standard and Extended Batteries.* Itasca, IL.: Riverside Publishing.

Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision Making, 11*(1), 99.

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

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. (2018). *Bias in Science and Communication*. Bristol, UK: IOP Publishing.

Welsh, M. B. (2020a). May I Have Your Attention? Testing a Subjective Attention Scale. In S. Denison, M. Mack, Y. Xu, & B. C. Armstrong (Eds.), *Proceedings of the 42nd Annual Conference of the Cognitive Science Society* (pp. 2535-2541): Cognitive Science Society.

Welsh, M. B. (2020b). Overconfident in Hindsight: Memory, Hindsight Bias and Overconfidence. In S. Denison, M. Mack, Y. Xu, & B. C. Armstrong (Eds.), *Proceedings of the 42nd Annual Meeting of the Cognitive Science Society* (pp. 822-828): Cognitive Science Society.

Welsh, M. B., & Begg, S. H. (2016). What have we learnt? Insights from a decade of bias research. *APPEA Journal, 56*, 435-450.

Welsh, M. B., Burns, N., & Delfabbro, P. (2013). *The Cognitive Reflection Test: how much more than numerical ability.* Paper presented at the Proceedings of the 35th Annual Conference of the Cognitive Science Society.

Welsh, M. B., Delfabbro, P. H., Burns, N. R., & Begg, S. H. (2014). Individual differences in anchoring: traits and experience. *Learning and Individual Differences, 29*, 131-140. doi:10.1016/j.lindif.2013.01.002

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.