Debiasing through raising awareness reduces the anchoring bias

Dr Carolyn Mair[1], Prof Martin Shepperd[2] & Prof Magne Jorgensen[3]

[1]London College of Fashion, University of the Arts London, UK
[2]Department of Computer Science & Maths, Brunel University, London, UK
[3]Simula Labs, Oslo, Norway

ABSTRACT

Objective: To investigate the impact of the anchoring bias and debiasing interventions with professional software engineers. We hypothesised that debiasing strategies would reduce bias

Method: Ethical approval was given prior to conducting the study. Five experiments were conducted with software engineers (*n*=126) in three locations. At each location, participants attended a workshop raising awareness of the impact on estimation of a range of cognitive biases, including anchoring and offering debiasing strategies. Immediately following the workshop, the participants at each location were divided into two groups (high versus low anchor) and asked to provide an estimate of effort for a familiar task. Data were analysed using Mann-Whitney tests. ANOVAs were used to test the impact of the workshop on the estimation task. Effort estimation data collected previously with different samples of software engineers at four other locations were compared with the effort estimation data collected following the workshop.

Results: The anchor had a large and significant impact (effect size, large, 0.3; Mann-Whitney, p<0.0001) on the effort estimates. The impact of the anchor was significantly reduced following the workshop (effect size, large 0.14; 2-way ANOVA, p<0.0001).

Conclusions: The impact of the anchoring bias on expert judgement is substantial and difficult to eradicate. Nevertheless, a debiasing activity such as attending a workshop to raise awareness can significantly reduce the effect. The study is limited in that the samples involved in the analysis of the impact of the workshop on effort estimation were independent. This will be addressed in further work.

BACKGROUND

The influence of information on decision making has been studied for decades by cognitive psychologists (e.g., Tversky & Kahneman, 1974; Kahneman, Slovic, & Tversky, 1982). An important finding is that humans typically use 'rules of thumb' heuristics when making decisions. Heuristic-thinking tends to be automatic, therefore less cognitively demanding allowing us to apply cognitive effort to more seemingly demanding tasks. West and Stanovich (1997) described a 2-system cognitive processing model. The fast, automatic system better suited to handling basic and simple processing; and the slow conscious system better suited to completing cognitively demanding tasks. The fast system uses heuristics which generally suffice for quick decision making, but when accuracy is more important than speed, can provide suboptimal solutions. The influence of heuristic thinking on decision making is known as cognitive bias. Empirical studies in cognitive science and social psychology have identified a wide variety of cognitive biases, common in many domains.

Professionals are frequently required to make decisions. Typically, solutions are based on a combination of past experience, decision support tools and information from multiple sources. However, recall is prone to error and this can be significantly influenced by misleading information (e.g., Loftus, 1975). Although this is well understood in eye witness testimony research, it is an under-unexplored area of research in the context of software engineering: the context of interest in this paper.

Aranda and Easterbrook (2005) explored the effects of anchoring on software estimation. They found that anchoring occurs in software estimation, and this can significantly change the resulting estimates, no matter what estimation technique is used. They also found that software estimators tend to be too confident of their own estimations. More recently Magazinius, Börjesson and Feldt (2012) investigated the impact of bias on prediction in software engineering and found human and organizational factors should be considered when addressing estimation problems. This paper focuses specifically on the anchoring bias, because it has been found to lead to significant distortions (Klayman & Brown, 1993; Buehler, Peetz & Griffin, 2010). The anchoring bias results from over-reliance on specific information during decision making such that 'solutions' are adjusted to that information to account for other elements of the circumstance (Mussweiler & Strack, 2001). Typically, once set, there is a bias toward the anchoring information. Jorgensen and Grimstad (2012) investigated the anchor bias within the software engineering domain. They explicitly requested software engineers to ignore misleading information, but still found significant differences in productivity estimations. between randomly assigned groups of software engineers who had been given either a high or a low anchor prior to being asked to reflect on previous performance. Participants were asked to estimate the number of lines of code (LOC) they had written on average in their most recently completed project. Participants in the low anchor group were asked "Did you write more than 1 LOC/hour?"; participants in the high anchor group were asked "Did you write less than 200 LOC/hour?". Jorgensen found that participants consistently anchored their estimates to the low or high anchor they had been given.

Most of the literature on cognitive bias reports its impact rather than strategies to reduce it. Hence the purpose of our experiment is to investigate whether it is possible to reduce or even eliminate the anchor effect. The context for this investigation was software professionals.

The contributions of this paper are twofold. First, past work has predominantly focused upon understanding factors that contribute to bias in decision-making whilst we examine interventions that potentially can reduce bias. Second, the experiment with 118 professional participants performed in four different countries is then pooled with previous, similarly designed experiments enabling analysis of results derived from in excess of 400 software engineers completing a highly relevant estimation task. We believe this to be one of the largest experimental studies of software developer decision making.

METHOD

Ethical approval was given prior to conducting the study which involved participants attending a workshop and completing an estimation task. The task was based on a series of experiments using the estimation task with 295 participants from industry (Jorgensen & Grimstad, 2012, Estimation Task 1). Five experiments were conducted with software engineers (*n*=118) in two locations. At each location, participants attended the workshop aimed at raising awareness of the impact of a range of cognitive biases, including anchoring, on estimation. Immediately following the workshop, participants at each location were randomly divided into two groups (high versus low anchor) and asked to estimate their own productivity for a familiar task. The low anchor was "Do you believe your coding productivity was more than 1 LOC per hour on your last project?" The high anchor was "Do you believe your coding productivity was less than 200 LOC per hour on your last project?". Participants recorded 'Yes' or 'No' and their actual estimate of programming productivity in LOC per hour. The actual estimates are used for this analysis in which we compare the task results with those of participants from Jorgensen and Grimstad (2012) in which participants completed the same task without attending a workshop. Data were analysed using Mann-Whitney tests. To test the impact of the workshop on the anchoring bias, effort estimation data collected previously with different samples of software engineers at each location, were compared with effort estimation data following the workshop (ANOVA)s.

RESULTS

Seven experiments were conducted with a total of 410 participants. Of these, 202 were in the high anchor group; 208 were in the low anchor group (see table 1). All participants completed an estimation task (see Table 2). The impact of the anchor is statistically significant (p≤ 0.0001) and the effect size is large (eta-squared=0.247). The intervention, a workshop, was given to 126 of the participants. The impact of the intervention on estimation is significant with a large effect (Tables 3 and 4) and Figure 1 estimated productivity by anchor value, and Figure 2 estimated productivity by workshop.

Country	Count	Attended Workshop	High Anchor	Low Anchor	Total
Nepal	59	0	29	30	59
NZ	18	18	9	9	18
Poland	92	92	48	44	92
Romania	48	0	26	22	48
UK	16	16	9	7	16
Ukraine	114	0	54	60	114
Vietnam	63	0	27	36	63
Total	410	126	202	208	410

Table 1 Summary statistics for estimated productivity

Estimated productivity	Count	Mean	Median	SD	Min	Max
All	410	52.7	30	58.7	0.5	300
High	202	82.1	80	59.7	0.5	300
Low	208	24.2	10	39.4	0.5	250

Table 2 Summary statistics for estimated productivity by anchor value

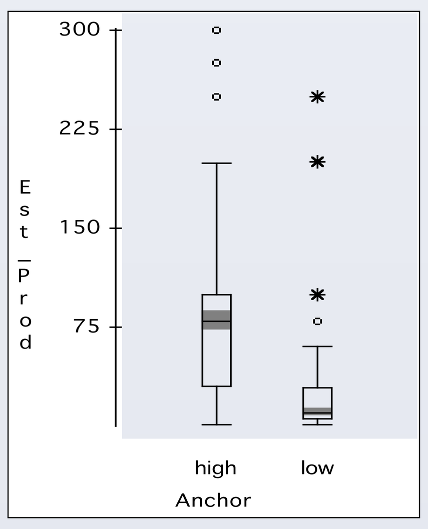

Figure 1: Boxplots of estimated productivity by anchor value

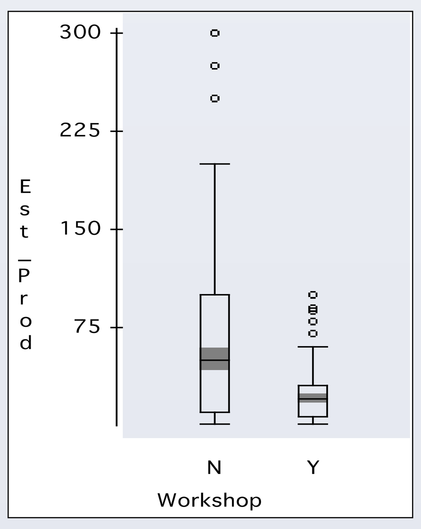

Figure 2: Boxplots of estimated productivity by workshop

Attended Workshop	High Anchor	Low Anchor	Difference
N	28.8 (sd=44.9)	101.9 (sd=58.5)	73.1
Y	12.4 (sd=12.3)	35.2 (sd=27.8)	22.8

Table 3 Means and standard deviations of productivity estimates by intervention

Source	df	Sums of Squares	Effect size	F-ratio	Prob
Anchor	1	193112	24.9%	93.6	≤ 0.0001
Workshop	1	144824	10.5%	70.2	≤ 0.0001
Anchor* Workshop	1	53032		25.7	≤ 0.0001
Error	406	837686			
Total	409	1379207			

Table 4 2-way Analysis of Variance for estimated productivity

When software professionals are asked to estimate with differing anchor values, the means differ by approximately 350%. Whilst the anchor is significant and 'explains' about a quarter of the variation in the estimates it does not account for most of the variance hence the high error term. However, analysing the data using a 2-way ANOVA Overall model is able to account for approximately 40% of the response variable variance, however the error term still represents 60% of the variability in the estimates. This suggests that there are many other factors potentially including individual differences, variation in use of different development tools and so on that impact productivity and tasks differences.

CONCLUSIONS

Seven experiments were conducted with a total of 410 participants. Of these, 208 were in the high anchor group; 202 were in the low anchor group. 126 attended a workshop aimed at debiasing cognitive biases through raising awareness and other strategies. All participants completed an estimation task (see Table 2). The impact of the anchor was statistically significant with large effect size. This supports previous work on cognitive bias (e.g., Tversky & Kahneman, 1974); West & Stanovich, 1997; Aranda & Easterbrook, 2005). 126 of the participants at 3 separate sites attended a workshop immediately prior to completing the task. The impact of the workshop on estimation was significant with a large effect. Therefore, we conclude that debiasing workshop significantly reduced, but did not eliminate this bias. It also reduced the variability in the estimates of professionals leading to more realistic (i.e., lower) estimates. Interestingly the debiasing workshop had a greater impact for the high than the low anchor.

The study is limited in that the samples involved in the analysis of the impact of the workshop on effort estimation were independent. This will be addressed in further work.

ACKNOWLEDGEMENTS

This work was funded by EPSRC Grants EP/I038225/1 and EP/1037881/1.
We are also grateful to the participants of the experiments.

REFERENCES

Aranda, J., & Easterbrook, S. (2005). Anchoring and adjustment in software estimation. In *ESEC-FSE'05* (pp. 346–355). Lisbon, Portugal: ACM Press.
Buehler, R., Peetz, J., & Griffin, D. (2010). Finishing on time: When do predictions influence completion times? *Organizational Behavior and Human Decision Processes*, *111*(1), 23–32.
Jorgensen, M., & Grimstad, S. (2012). Software development estimation biases: the role of interdependence. *IEEE Transactions on Software Engineering*, *38*(3), 677–693.
D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases. Cambridge University Press, Cambridge, UK, 1982.* Cambridge, England: Cambridge University Press.
Klayman, J., & Brown, K. (1993). Debias the environment instead of the judge: an alternative approach to reducing error in diagnostic (and other) judgment. *Cognition*, *49*(1–2), 97–122. doi:10.1016/0010-0277(93)90037-V
Loftus, E.F. (1975) Leading questions and the eye witness report, Cognitive Psychology, 7, 560-572.
Magazinius, A., Börjesson, S., & Feldt, R. (2012). Investigating intentional distortions in software cost estimation – An exploratory study. *Journal of Systems and Software*, *85*(8), 1770–1781.
Mussweiler, T., & Strack, F. (2001). The semantics of anchoring. *Organizational Behavior and Human Decision Processes*, *86*(2), 234–255.
Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124–1131.
West, R. F., & Stanovich, K. E. (1997). The domain specificity and generality of overconfidence: Individual differences in performance estimation bias. *Psychonomic Bulletin & Review*, *4*, 387-392.