

Term paper

Anchoring in exam score estimation: a classroom experiment

Matej Lorko

Abstract: Ability to correctly judge the demandingness of university exam can help students figure out how much time and effort they need to put into their preparation. Overly optimistic judgments can lead to underestimation of the exam demandingness and result in failure to pass the course. In this paper, I experimentally test the effect of low and high anchors on past average exam score estimates. I find that estimates under the low anchor are significantly lower than estimates without the anchor, but the estimates under the high anchor do not differ significantly with estimates produced without an anchor. Importantly, vast majority of the students in all treatments underestimate how demanding the exam is.

1. Introduction

The ability to make correct estimates is a vital skill for professionals in many different domains. However, Tversky & Kahneman (1974) propose that in various estimating situations, mistakes can be made as a result of using an initial value (an anchor) as a starting point and insufficiently adjusting the final figure away from it. The anchor can appear, for example, in form of a suggestion, or as a result of the partial computation. The estimates are usually systematically biased towards the anchor. Thus, the anchoring effect is observed, when considering the same problem, different starting points lead to different estimates or values.

Tversky & Kahneman (1974) demonstrate the anchoring effect on estimates of the percentage of African countries in the United Nations. The anchoring effect was later replicated in many other studies in the domain of general knowledge e.g., Jacowitz & Kahneman (1995), Epley & Gilovich (2001) Blankenship, Wegener, Petty, Detweiler-Bedell, & Macy (2008). However, the effect is not restricted to factual questions. The influence of numerical value presented to research subjects before they make a prediction, judgment or decision, was found in many other areas, such as negotiation (Galinsky & Mussweiler, 2001; Ritov, 1996), purchasing decisions and valuations (Ariely, Loewenstein, & Prelec, 2003), probability estimates (Plous, 1989), sentencing decisions (Englich, Mussweiler, & Strack, 2006), performance forecasts (Critcher & Gilovich, 2008), social judgments (Davis, Hoch, & Ragsdale, 1986), self-efficacy (Cervone & Peake, 1986), or project duration estimates (Lorko, Servátka, & Zhang, 2019). In the comprehensive review of anchoring effect literature, Furnham & Boo (2011) conclude, that “nearly 40 years worth of research on the anchoring has proved the effect to be extremely robust and it is actually unusual in experimental settings not to be able to demonstrate it”.

In this paper, I focus on the effect of anchors on estimates of course exam scores. Correct judgment of the exam difficulty can help students figure out how much time they would need to spend preparing for the exam and how much effort they will need to exert in order to successfully pass the course. Overly optimistic estimates of exam scores can lead to underestimation of the exam demandingness and result in failure to pass the course. Thus, it is in the best interest of students to assess how difficult the exam is as correctly as possible.

I ran a classroom experiment in which students individually and privately estimated the average course exam score in previous academic year. Students were randomly assigned into a Control treatment, a Low Anchor treatment in which they were requested to consider an extremely low average score before producing their estimate, and a High Anchor treatment in which they were requested to consider an extremely high average score before producing the estimate. I hypothesized that Low Anchor treatment would result in significantly lower exam score estimates than the Control treatment, while High Anchor treatment would result in significantly higher estimates.

- **Hypothesis 1a:** *Low anchor will cause significantly lower estimates of average exam score.*
- **Hypothesis 1b:** *High anchor will cause significantly higher estimates of average exam score.*

One might expect that the anchoring effect could disappear with the rising level of subjects' knowledge, expertise or experience in estimated task. In fact, Løhre & Jørgensen (2016) show that more experienced subjects are indeed less influenced by anchors thus provide more accurate estimates. To test whether the anchoring effect can be mitigated by having more prior knowledge about the course, for each subject I measured their number of older friends who had completed the course in the past.

- **Hypothesis 2:** *Having more friends who completed the course in the past mitigates the anchoring effect.*

2. Experimental design

I ran the experiment during the first lecture “Introduction to experimental methods in economics” course at the University of Economics in Bratislava. Students were requested to fill a short online form eliciting basic demographics, the number of their friends who already completed the course, and an incentivized estimate regarding the average exam score from the course in the previous academic year. Students were informed that if their estimate would be more accurate than the median estimate, they would be awarded with 2 points towards their final course grade.

The average exam score question read as follows: “What was the average number of points (out of 50) scored in this course final exam by students last year?”. Before placing their estimates, students randomly selected into the Low Anchor treatment were asked whether they thought the average was higher or lower than 5 points, while students randomly selected into the High Anchor treatment were asked whether they thought the average was higher or lower than 45 points.

3. Results

A total of 87 students, with a mean age of 22.4 years (standard deviation 2.4 years) and 48% female, participated in the experiment. The number of students in each treatment ranged from 24 to 33. The descriptive statistics are presented in Table 1.

Table 1: Descriptive statistics

Treatment	N	Estimate (SD)	Age (SD)	Female	No. of friends (SD)
Control	24	37.4 (7.3)	22.6 (2.5)	46%	0.3 (1.2)
Low Anchor	30	28.9 (12.7)	22.0 (1.4)	50%	0.9 (1.7)
High Anchor	33	38.1 (7.3)	22.6 (3.0)	48%	0.4 (0.9)

Note: SD refers to Standard Deviation.

I analyze the differences in estimates across treatments using ANOVA and Kruskal-Wallis tests. Both tests show that there is a significant difference between at least one of the three pairs of treatments (p -value $< .01$ in both tests). Table 2 shows p -values of pairwise comparisons (t-tests and Mann-Whitney tests) for the three treatments. Both tests show that estimates are significantly lower in Low Anchor treatment compared to Control treatment as well as compared to High Anchor treatment. I find no significant differences in estimates between the Control treatment and High Anchor treatment.

Table 2: Pairwise comparisons of estimates across treatments

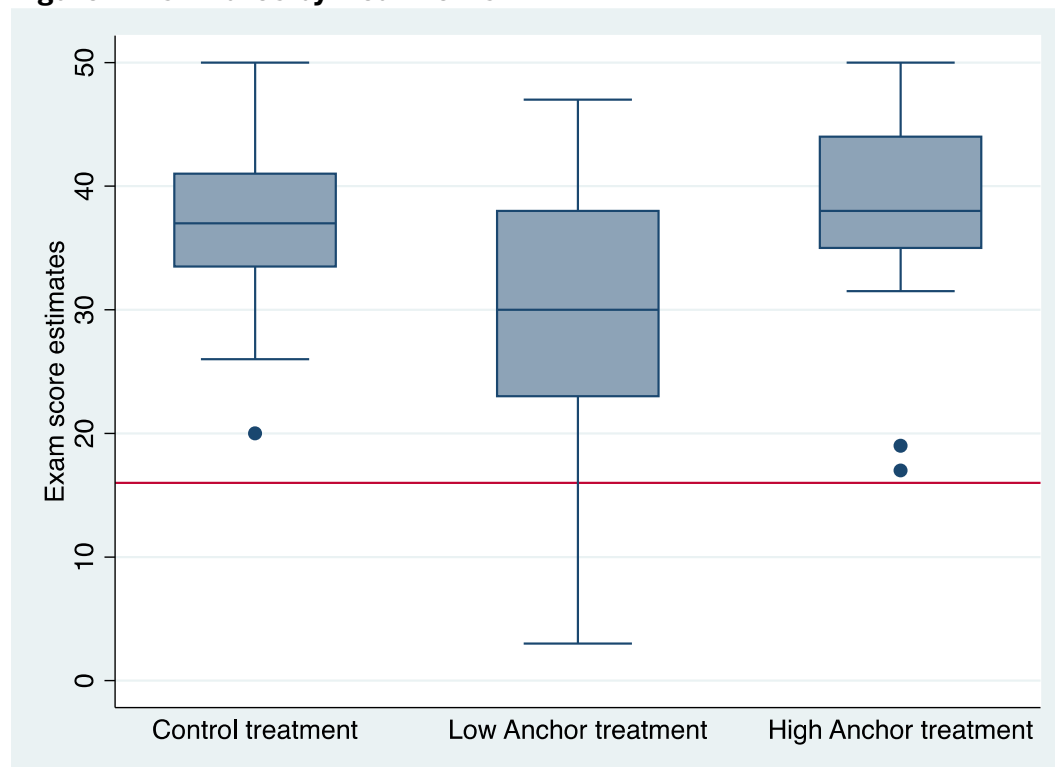
Comparison	t-test p-value	Mann-Whitney test p-value
Control vs. Low Anchor	0.01	0.02
Control vs. High Anchor	0.74	0.50
Low Anchor vs. High Anchor	$<.01$	$<.01$

Result 1a: Low anchor causes significantly lower estimates of average exam score.

Result 1b: High anchor does not cause significantly higher estimates of average exam score.

Figure 1 shows box plots of estimate by treatments. In addition, it shows the actual average exam score. As is apparent from the figure, students in all treatments were overly optimistic about the exam demandingness.

Figure 1: Estimates by treatments



Note: For reference, the red line indicates the actual average exam score.

To check the robustness of my finding and test my Hypothesis 2 stating that having more friends who completed the course in the past mitigates the anchoring effect, I ran OLS regressions presented in Table 3. Model 1 utilizes treatment dummies only and confirms that Low Anchor causes significantly lower estimates (by around 8.5 points, with the last model showing more precise difference of almost 12 points) while High Anchor does not have a significant effect on estimates. In Model 2 I add demographic controls, while in Model 3 I also test the interaction between the anchors and number of friends. Models 2 and 3 show that females estimated significantly higher exam scores than males. Interestingly, having more friends who completed the course in the past resulted in higher average score estimates (significantly higher in Low Anchor treatment), partially mitigating the anchoring effect.

Result 2: *Having more friends who completed the course in the past yields higher estimates in the Low Anchor treatment, effectively mitigating the anchoring effect.*

Table 3: OLS regressions

	(1)	(2)	(3)
	Estimate	Estimate	Estimate
Low Anchor treatment	-8.55*** (2.75)	-9.16*** (2.75)	-11.76*** (2.88)
High Anchor treatment	0.64 (1.94)	0.40 (2.14)	0.92 (2.12)
Age		0.26 (0.25)	0.27 (0.27)
Female		5.39*** (1.89)	5.93*** (1.88)
No. of friends		1.03 (0.92)	-1.63*** (0.36)
Low Anchor * No. of friends			4.61*** (0.81)
High Anchor * No. of friends			-0.69 (2.53)
Constant	37.42*** (1.47)	28.68*** (6.12)	29.12*** (6.61)
N	87	87	87
R²	0.17	0.26	0.36

Notes: Standard errors are reported in parentheses.

*, **, and *** indicate significance at the 10%, 5%, and 1%-level, respectively.

4. Discussion

The effect of low and high anchors on various estimates is one of the most robust findings in psychology and behavioral economics. In this paper, I tested the anchoring effect on estimates of average final exam scores. Interestingly, while I found that estimates produced under the influence of low anchor are significantly lower, I did not find the effect of high anchor. This result can, however, be explained by the fact that the estimates in the Control treatment were relatively high (over 37 points out of 50), which had left a little room for the estimates in High Anchor treatment to be significantly higher.

In fact, all treatments resulted in overly optimistic estimates of the exam scores, showing that students believed the exam is much easier than it really is. While having more friends who completed the exam in the past did slightly mitigate the effect of low anchor, it did not result in more realistic (accurate) estimates. Although the goal of this paper was to examine the anchoring effect, I believe that my results show an interesting avenue for the future research – to study student overconfidence regarding the demandingness of course exams.

References

- Ariely, D., Loewenstein, G., & Prelec, D. (2003). “Coherent Arbitrariness”: Stable Demand Curves Without Stable Preferences. *The Quarterly Journal of Economics*, 118(1), 73–106. <https://doi.org/10.1162/00335530360535153>
- Blankenship, K. L., Wegener, D. T., Petty, R. E., Detweiler-Bedell, B., & Macy, C. L. (2008). Elaboration and consequences of anchored estimates: An attitudinal perspective on numerical anchoring. *Journal of Experimental Social Psychology*, 44(6), 1465–1476. <https://doi.org/10.1016/j.jesp.2008.07.005>
- Cervone, D., & Peake, P. K. (1986). Anchoring, efficacy, and action: The influence of judgmental heuristics on self-efficacy judgments and behavior. *Journal of Personality and Social Psychology*, 50(3), 492–501. <https://doi.org/10.1037/0022-3514.50.3.492>
- Critcher, C. R., & Gilovich, T. (2008). Incidental environmental anchors. *Journal of Behavioral Decision Making*, 21(3), 241–251. <https://doi.org/10.1002/bdm.586>
- Davis, H. L., Hoch, S. J., & Ragsdale, E. K. E. (1986). An Anchoring and Adjustment Model of Spousal Predictions. *Journal of Consumer Research*, 13(1), 25–37. <https://doi.org/10.1086/209045>
- Englich, B., Mussweiler, T., & Strack, F. (2006). Playing Dice With Criminal Sentences: The Influence of Irrelevant Anchors on Experts’ Judicial Decision Making. *Personality and Social Psychology Bulletin*, 32(2), 188–200. <https://doi.org/10.1177/0146167205282152>
- Epley, N., & Gilovich, T. (2001). Putting adjustment back in the anchoring and adjustment heuristic: differential processing of self-generated and experimenter-provided anchors. *Psychological Science*, 12(5), 391–396. <https://doi.org/10.1111/1467-9280.00372>
- Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *Journal of Socio-Economics*, 40(1), 35–42. <https://doi.org/10.1016/j.socec.2010.10.008>
- Galinsky, A. D., & Mussweiler, T. (2001). First offers as anchors: The role of perspective-taking and negotiator focus. *Journal of Personality and Social Psychology*, 81(4), 657–669. <https://doi.org/10.1037/0022-3514.81.4.657>

- Jacowitz, K. E. , & Kahneman, D. (1995). Measures of anchoring in estimation tasks. *Personality and Social Psychology Bulletin*, 21(11), 1161–1166.
<https://doi.org/10.1177/01461672952111004>
- Løhre, E., & Jørgensen, M. (2016). Numerical anchors and their strong effects on software development effort estimates. *Journal of Systems and Software*, 116, 49–56.
<https://doi.org/10.1016/j.jss.2015.03.015>
- Lorko, M., Servátka, M., & Zhang, L. (2019). Anchoring in project duration estimation. *Journal of Economic Behavior & Organization*, 162, 49–65.
- Plous, S. (1989). Thinking the Unthinkable: The Effects of Anchoring on Likelihood Estimates of Nuclear War. *Journal of Applied Social Psychology*, 19(1), 67–91.
<https://doi.org/10.1111/j.1559-1816.1989.tb01221.x>
- Ritov, I. (1996). Anchoring in Simulated Competitive Market Negotiation. *Organizational Behavior and Human Decision Processes*, 67(1), 16–25.
<https://doi.org/10.1006/obhd.1996.0062>
- Tversky, A., & Kahneman, D. (1974). Judgments under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131. <https://doi.org/10.1126/science.185.4157.1124>