Trust the Experienced? Investigating the Effect of Experience on Decision Making in the Crowd

Thomas Görzen

Paderborn University, Business Administration and Economics, Paderborn, Germany
{Thomas.Goerzen}@wiwi.uni-paderborn.de

Abstract. Companies increasingly use the crowd for collective decision making. Currently the most popular approach used to aggregate multiple decisions is majority voting. This approach however can lead to suboptimal outcomes by, on the one hand, neglecting different levels of individuals' experiences and, on the other, ignoring decision biases to which individuals are prone. As an alternative to simply averaging judgements, researchers have proposed weighted models that favor more experienced judges in the crowd. However, previous research is inconclusive as to whether more experienced judges are any better at avoiding decision biases. To shed light on this question, we conduct several experiments using the well-established bias known as the anchoring effect. Results indicate that a greater level of experience does not decrease reviewers' tendency to follow the displayed anchor, and consequently does not reduce their proneness to decision bias.

Keywords: Crowdsourcing, anchoring effect, biased decision, experience.

1 Introduction

With the advent of online ideation communities, the challenge of effectively screening a large number of ideas is more relevant today than ever [1]. The crowd is increasingly used not only to generate new ideas but also to evaluate these ideas, which can be numerous [2]. The method of using a large number of judgements - the so-called wisdom of the crowd - offers several advantages since it i) maximizes the amount of information that a decision draws upon, given the sheer diversity of information backgrounds that a large number of different people bring to a decision and, ii) it reduces the potential impact of 'outliers' - extreme decisions based on unreliable or simply inaccurate information sources. The currently favorite application for using the wisdom of the crowd approach is to simply average the judgements of each individual decision in the crowd [3]. Though, this can lead to suboptimal outcomes since it neglects external information such as experience [4]. As an alternative to simply averaging individual judgements, researchers have proposed weighted models that favor more experienced judges in the crowd (e.g. [5]). The assumption behind this approach is that more experienced judges are less likely to be biased in their decision making. However, previous studies on decision making in the offline context show

Multikonferenz Wirtschaftsinformatik 2018,
March 06-09, 2018, Lüneburg, Germany

contrasting results. While some studies indicate that knowledgeable people are less influenced by biases (e.g. [6]), others show contrasting results, suggesting that even knowledgeable people with experience in a given context are significantly biased (e.g. [7]). With respect to these conflicting results, we want to shed light upon the question of whether more experienced people in the crowd are less likely to make biased decisions. Hence, we aim to answer the following research question:

Are more experienced people in the crowd less inclined to make distorted decisions?

To answer this question, we use the established anchoring effect [8] in experiments on a commercial crowdworking platform, giving the crowd 80 different business model ideas to rate. We implement the anchor, i.e. information about previous rating results, using it as a treatment, and analyze the effect on individual raters' decisions. We then ask each individual rater about her background experience in several dimensions such as business models, the product category (perfume) and their experience in retail and analyze the effect of experience on the occurrence of the anchoring effect. Results indicate that experience in different dimensions does not decrease the probability to follow the anchor and therefore does not protect against distorted decisions.

2 Related Literature

2.1 Anchoring Effect in Decision Making

The anchoring effect describes the disproportionate influence of an initially presented value on decision makers [8]. The occurrence of this effect has been shown in a variety of different domains such as general knowledge [9] or probability estimation [10]. Previous studies show that both uninformative as well as anchors with informational relevance to the task itself can influence the decision. In a classic study, [8] randomly generated anchor values by spinning a wheel of fortune between 0 and 100 and afterwards asked the participants to estimate the percentage of African countries in the United Nations. The given anchor values had a strong influence on their following estimation, since high anchor values led to significant higher estimations than low anchor values [8]. Another study found that the estimation of an athlete's performance could be anchored by the number on his jersey [11]. Anchors with relevance to the task can also lead to the anchoring effect: in an example from the legal domain, higher damage awards were obtained when higher compensations were requested in court [12].

2.2 Influence of Experience on Anchoring Effect

Previous literature on the influence of experience on the anchoring effect show contradictory results. On the one hand one could argue that people with high expertise should have greater knowledge and experience, which would mean they would be less influenced by the anchors provided. This is supported by research which shows that both people with high certainty about an answer [10] and with greater relevant knowledge are less influenced by anchors [6]. On the other hand, it could be argued that the underlying mechanisms to the anchoring effect - notably the "conformity

search" [10] and "selective accessibility" [13] - are such fundamental cognitive processes that people regardless of their experience and knowledge could fall victim to this effect. Previous research also supports this position, indicating that people are influenced by anchors regardless of their experience ([14], [7]).
Thus, research in the field shows contradicting results, on the one hand indicating that experienced people are less influenced by the anchoring effect while others show the anchoring effect to be a robust effect in human decision making regardless of the experience of decision makers.

3 Methodology

To answer our research question we design an experiment which enables us (1) to analyze the occurrence of the anchoring effect in idea evaluation on a crowdworking platform and (2) to investigate whether experience might protect people from being influenced by the anchoring effect. The ideas in our experiment were taken from previous research [15], in which students generated business model ideas for perfume in a classroom experiment. All ideas are presented in the same way, consisting of the nine elements of the Business Model Canvas [16]. We designed evaluation tasks on *Crowdflower*, a commercial crowdsourcing platform. We divided the ideas into eight blocks of ten ideas each and randomly assigned each idea to one block. Each participant had to rate the displayed ideas on a seven-point scale in terms of creativity, novelty and usefulness [17].
To avoid the occurrence of a learning effect and potential rating bias through users who assign themselves to several tasks in succession, we allowed each participant only to assign themselves to the job once. Each block of ideas was evaluated by twenty different contributors, each earning 0.50€ for the evaluation of one block of ideas (= ten ideas). We calculated the average rating of each idea based on crowd evaluation. This part of our experiment represents the control condition in which each individual had to evaluate the ideas without the display of an anchor. To analyze the influence of the given anchor and of the experience on the anchoring effect, we designed two additional experiments with two different treatments, enabling us to investigate the anchoring effect. First, we designed an experiment *(Crowd_Anchor)* where the information about the previous evaluation resulting from the control condition of each idea was displayed above the rating scale. Second, we designed another additional experiment *(Random_Anchor)* where the only difference was that for the displayed rating, each idea was assigned a randomly generated rating between 1.0 and 7.0. The general task design of the control condition was retained for both additional experiments.
Further, all participants had to answer a short survey with respect to their experience in addition to the evaluation task. Experience can be defined as the acquisition and cumulative knowledge of reality, mechanisms, rules, and procedures related to a specific domain [18]. To address different dimensions of individuals' experience, we ask each participant about their experience regarding the following dimensions: Context/market mechanism, platform experience, product knowledge and business models. After the idea evaluation was completed, we further asked each subject about their confidence in evaluating the displayed business model ideas. All scales ranged

from 1 = “Not at all” to 7 = “Absolutely agree”. Further, since we need a benchmark for idea quality to know whether the randomly assigned rating of an idea strongly deviates from a more professional evaluation of idea quality, we recruited two experts to evaluate the ideas, one from our university’s entrepreneurship center and another who works as a senior in-house consultant in a large corporation. Experts commonly represent a benchmark since they are characterized by their experience and typically consider relevant characteristics such as feasibility and business potential when evaluating ideas [19].

4 Results

4.1 Analyzing the General Occurrence of an Anchoring Effect

To analyze the general occurrence of an anchoring effect, we first calculated the average rating of each idea in each experimental condition. Table 1 presents the descriptive statistics for all 80 ideas for the control condition *(No_Anchor)* as well as both experimental conditions.

Table 1. Descriptive Statistics

Condition	Obs.	Mean	Std. Dev.	Min.	Max.
No_Anchor	80	4.58	.412	2.60	5.87
Crowd_Anchor	80	4.59	.479	3.35	5.5
Random_Anchor	80	4.39	.745	2.60	5.87

To investigate the general occurrence of the anchoring effect, we first analyzed the distributions of individual ratings from each experiment (Figure 1). As can be seen, the given distributions differ depending on whether and which anchor was shown in the condition. For example, in experiment 2 we can see a “range reduction effect” with people tending to provide a rating close to the average rating of experiment 1. In contrast, if showing a randomly generated anchor in experiment 3, we see a much broader distribution.

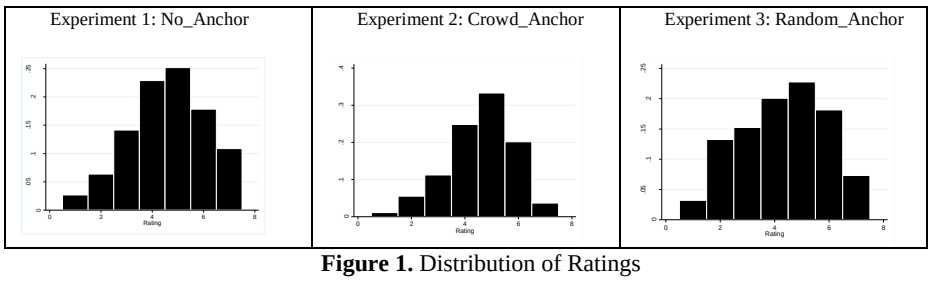

Figure 1. Distribution of Ratings

In addition to a visual inspection of the distributions, we use a common measure of the distance between two distributions, the Kolmogorov-Smirnov statistic. The Kolmogorov-Smirnov statistic between two cumulative distributions is defined as the maximum absolute difference between two distributions.

Table 2. K-S Statistics for Comparison of Distribution

Conditions	Difference (Combined K-S)
No_Anchor - Crowd_Anchor	.0716 (.001)**
No_Anchor - Random_Anchor	.0851 (< .001)***
Crowd_Anchor - Random_Anchor	.1378 (< .001)***

Note: *p <.10; **p < .05; ***p< .01

Results (Table 2) indicate statistically significant differences in the distribution of ratings for all experiments. We can assume, therefore, that a general anchoring effect occurred in experiment 2 and experiment 3, resulting in a significantly different rating behavior displayed by participants.

4.2 Analyzing the Anchoring Effect Depending on Experience

To analyze whether more experienced people are less inclined to be influenced by the displayed anchor, we further investigate the ratings of experiment 3, where we displayed a randomly generated rating between 1.0 and 7.0 for each idea. Although the probability is quite low, it still could be the case that even the random rating reflects the professionals' rating of the idea (= experts' rating) and therefore, following the anchor might only represent a rational decision. To address this issue, we only analyze ideas where the difference between the displayed anchor and the experts' rating was ≥ 2, which means that for example an anchor of two or lower was assigned to an idea rated four by the experts. In sum, we analyzed 740 observations. We used a logistic regression with dependent variable =1 if the person's rating was consistent with the displayed anchor. Due to the fact that people could only rate integers (for example: 4) while the anchor was provided as a float (for example: 4.2), the dummy variable is =1 if the person's rating was closer to the provided anchor. If the anchor was 4.4, for example, and the person rated the idea with 4, the dummy variable would be 1. In contrast, if the provided anchor was 4.6, the dummy variable would be 1 if the person rated the idea with 5 since this value is closer to the provided anchor than 4. To avoid multicollinearity in our model, since some variables consisted of several items, we conducted a principal component analysis [20] to summarize multi-item variables. Thus, we consider the following model [21]:

$$Y^* = \beta_0 + \beta_1 Mem_i + \beta_2 BMC\_Exp_j + \beta_3 Perf\_Exp_j + \beta_4 Retail\_Exp_i + \beta_5 Conf\_Rating_j + \varepsilon,\ Y = 1[Y^* > 0].$$

Mem_j represents the reported length of membership on the platform of a person. Further, BMC_Exp_j captures the multi-items for the reported experience in business models, while $Perf\_Exp_j$ captures the different items for the reported experience and knowledge of perfume. $Retail\_Exp_i$ represents the reported experience in months of a participant working in retail, while $Conf\_Rating_j$ covers the reported confidence of the participant when evaluating the business model ideas. Results (Table 3) indicate that experience in different dimensions do not significantly decrease the probability to follow the displayed anchor. In other words, higher levels of experience do not increase protection against making a distorted decision following a randomly generated anchor.

Table 3. Results of Logistic Regression

Variable	Coef.	Std. Dev.	z	P > |z|
Mem_i	.0096	.0066	1.45	.147
BMC_Exp_j	-.0423	.1278	-.33	.740
$Perf\_Exp_j$	-.0029	.0880	-.03	.974
$Retail\_Exp_i$	.0001	.0018	.08	.937
$Conf\_Rating_j$	.2292**	.1380	1.66	.097
Constant	-2.859***	.7665	-3.73	.000
Controls[1]	✓			

Notes: *p < .10; **p < .05; ***p < .01; [1]Gender and age

Surprisingly, results indicate that people who follow the displayed anchor feel more confident about their rating. One possible explanation for this result could be the so-called "conformity bias" [22], meaning that when people become uncertain about what to do, they are more likely to depend on others for the answer. Hence, when participants in our experiment rated the idea conform to the displayed anchor, they feel more confident about their decision since it seems to be in line with previous others.

5 Conclusion

In this study we investigate whether more experienced people in a crowd are less prone to distorted decision making. Our results indicate that experience in different dimensions does not decrease the probability to follow an anchor and therefore does not protect against distorted decision-making. We can conclude, therefore, that anchoring has a robust effect on human decision making regardless of the different experience levels of decision makers. Our study contributes to the body of literature on aggregating multiple judgements in crowdsourcing as well as to literature on the anchoring effect. In addition to contributing to research, our results also have managerial implications for task designers in crowdsourcing, since our results illustrate that weighted mechanisms to aggregate multiple judgements can also lead to biased results. However, we also acknowledge certain limitations, for example, the specific task assigned to the crowd (evaluation of business model ideas for perfume). Thus, we suggest that additional studies with other ideas or tasks should be conducted to investigate whether the effect is constant for different tasks or types of ideas. In the future, we seek to gain additional insights into the reasons for distorted decision making. Another focus could be to develop potential solutions in crowdsourcing task design aimed at minimizing distorted decisions and de-biasing decision making.

References

1. Simon, R.: One week, 3,000 product ideas. Wall Street Journal (July 3), http://www.wsj.com/articles/one-week-3-000-product-ideas-1404332942. (2014)

2. Soukhoroukova, A., Spann, M., Skiera, B.: Sourcing, Filtering, and Evaluating New Product Ideas: An Empirical Exploration of the Performance of Idea Markets. Journal of Product Innovation Management (29:1), 100-112 (2012)
3. Larrick, R.P., Mannes A.E. and Soll, J.B.: The Social Psychology of the Wisdom of Crowds. Krueger, J.I., ed. Frontiers of Social Psychology: Social Judgment and Decision Making, 227-242 (2011)
4. Budescu, D.V. and Chen, E.: Identifying Expertise to Extract the Wisdom of Crowds. Management Science (61:2), 267-280 (2015)
5. Aspinall, W.: A Route to more Tractable Expert Advice. Nature (463:7279), 294–295 (2010)
6. Wilson, T.D., Houston, C.E., Etling, K.M., and Brekke, N.: A New Look at Anchoring Effects: Basic Anchoring and its Antecedents. Journal of Experimental Psychology: General 125, 387–402 (1996)
7. Englich, B., Mussweiler, T.: Sentencing under Uncertainty: Anchoring Effects in the Courtroom. Journal of Applied Social Psychology 31, 1535–1551 (2001)
8. Tversky, A. and Kahneman, D.: Judgment Under Uncertainty: Heuristics and Biases. Science 185, 1124-1131 (1974)
9. McElroy, T. and Dowd, K.: Susceptibility to Anchoring Effects: how Openness-to-Experience Influences Responses to Anchoring Cues. Judgment and Decision Making 2, 48-53 (2007)
10. Chapman, G.B., Johnson, E.J.: The limits of anchoring. Journal of Behavioral Decision Making 7, 223–242 (1994)
11. Critcher, C.R. and Gilovich, T.: Incidental Environmental Anchors. Journal of Behavioral Decision Making 21, 241-251 (2008)
12. Marti, M.W., Wissler, R.L.: Be Careful What You Ask For: the Effects of Anchors on Personal Injury Damages Awards. Journal of Experimental Psychology: Applied 6, 91-103 (2000)
13. Strack, F., Mussweiler, T.: Explaining the Enigmatic Anchoring Effect: Mechanisms of Selective Accessibility. Journal of Personality and Social Psychology 73, 437–446 (1997)
14. Mussweiler, T., Strack, F., Pfeiffer, T.: Overcoming the Inevitable Anchoring Effect: considering the opposite compensates for selective accessibility. Personality and Social Psychology Bulletin 26, 1142–1150 (2000)
15. John, T. and Kundisch, D. 2015. "Why Use the Canvas for Idea Generation? A Design Theory and First Evidence." Working Paper, University of Paderborn.
16. Osterwalder, A. and Pigneur, Y.: Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers. Hoboken, NJ: Wiley (2010)
17. Amabile, T.M.: Creativity in Context: Update to the Social Psychology of Creativity. Boulder, CO: Westview Press (1996)
18. Galati, F. Complexity of Judgement: What Makes Possible the Convergence of Expert and Nonexpert Ratings in Assessing Creativity. Creativity Research Journal (27:1), 24-30 (2015)
19. Bayus, B.L. Crowdsourcing New Product Ideas Over Time: An Analysis of the Dell IdeaStorm Community. Management Science (59:1), 226-244 (2013)
20. Dunteman, G.H.: Principal Component Analysis. Sage Publications (1989)
21. Wooldridge J.M.: Econometric Analysis of Cross Section and Panel Data. MIT Press, Cambridge, MA. (2010)
22. Kameda, T., Nakanishi, D.: Cost–Benefit Analysis of Social/Cultural Learning in a Nonstationary Uncertain Environment: An Evolutionary Simulation and an Experiment with Human Subjects. Evolution and Human Behavior 23, 373–393 (2002)