Anchoring does not activate examples associated with the anchor value Štěpán Bahník¹

Faculty of Business Administration, Prague University of Economics and Business

¹ Faculty of Business Administration, Prague University of Economics and Business, náměstí Winstona Churchilla 4, Prague, 130 67, Czech Republic, bahniks@seznam.cz, ORCID: 0000-0002-0579-6808

Abstract

In the standard anchoring paradigm, people first compare a selected attribute of a target to a numeric value—an anchor. A subsequent absolute judgment of the target's attribute value is biased in the direction of the anchor. A prominent theory of the anchoring effect, the selective accessibility model, argues that people make the initial comparison by focusing on similarities between the target and the anchor, which activates information compatible with the anchor value being the target value. This activated information biases the subsequent estimate of the target value. To test the selective activation of information, the present study asked people to provide an example of the target's category following its comparison with an anchor. The attribute values of the provided examples were not biased in the direction of the anchor. However, they were positively associated with estimates of the target value. The study thus provides evidence for the use of activated information in the absolute judgment in the standard anchoring paradigm, but it does not show the selective activation of information compatible with the anchor value predicted by the selective accessibility model.

Keywords

anchoring, selective accessibility, judgment

Acknowledgments

I would like to thank Nikola Frollová, Markéta Sýkorová, Dominik Stříbrný, Marek Vranka, and Vojtěch Zíka for their help with data collection. The work was supported by the Internal Grant Agency of the Faculty of Business Administration, Prague University of Economics and Business (grant no. IP300040).

Human judgment is influenced by a number of biases (Pohl, 2017). One of the notable biases is the anchoring effect—assimilation of judgment toward a previously considered standard (Bahník, Englich, & Strack, 2017; Tversky & Kahneman, 1974). Anchoring can be observed in a variety of domains: First offers in negotiations serve as anchors (Galinsky & Mussweiler, 2001). Sentencing demands play a role of anchors in legal judgments (Englich,

Mussweiler, & Strack, 2006). Anchoring also influences valuation of products (Yoon, Fong, & Dimoka, 2019), probability assessments (Plous, 1989), and answers to general knowledge questions (Tversky & Kahneman, 1974).

In an early study demonstrating the anchoring effect (Tversky & Kahneman, 1974), participants were asked to compare the proportion of African nations in the United Nations to an anchor—a number determined by a wheel of fortune. Afterward, when asked to estimate the proportion of African nations in the United Nations, participants who had compared the proportion with a high number gave higher estimates than participants who had compared the proportion with a low number. The procedure used in the study constitutes the so-called standard anchoring paradigm, which consists of comparison and absolute judgment questions. First, a participant is asked to compare a target value to an anchor value. Second, a participant is asked to estimate the target value. The anchoring effect means that this estimate of the target value is biased in the direction of the anchor value.

Multiple explanations have been proposed for the anchoring effect. The selective accessibility model argues that anchoring is a result of positive hypothesis testing, which leads to activation of information compatible with the anchor value. This activated information is subsequently used in the absolute judgment (Mussweiler & Strack, 1999a, 1999b; Strack, Bahník, & Mussweiler, 2016; Strack & Mussweiler, 1997). Anchoring can also be a result of insufficient adjustment from the anchor (Epley & Gilovich, 2001; Tversky & Kahneman, 1974). Another explanation argues that anchors influence judgment through numeric priming (Wilson, Houston, Etling, & Brekke, 1996; Wong & Kwong, 2000), Anchors may be used in judgment because they are perceived to be informative (Schwarz, 1994). Anchors can also distort the perception of the scale on which the subsequent absolute judgment is made (Frederick & Mochon, 2012). Understanding the processes behind anchoring is important for debiasing the anchoring effect. Different processes could require different tools for preventing or reducing their impact on judgment. For example, the biasing effect of positive hypothesis testing may be reduced by a prompt to consider reasons why the anchor value may not be the true value (Mussweiler, Strack, & Pfeiffer, 2000). Scale distortion may be reduced by asking people to imagine objects with different values on the judgment scale before considering an anchor (Bahník, Houdek, Vrbová, & Hájek, 2019). The evidence for and against the various theories of anchoring is summarized in Bahník et al. (2017). The present study examines the activation of information compatible with the anchor value and its subsequent use in judgment, which lie behind the anchoring effect according to the selective accessibility model.

According to the selective accessibility model, people answer the comparison question using positive hypothesis testing (Klayman & Ha, 1987). When asked to compare the target value to an anchor, people consider reasons why the two values may be equal (Mussweiler, 2003). This process activates information compatible with the anchor value. For example, according to the selective accessibility model, participants in the aforementioned study by Tversky and Kahneman (1974) answer the question asking them to compare the proportion of African nations in the United Nations with 10% by considering the reasons why the proportion may be equal to 10%. They therefore think of information that is compatible with the possibility that only a small proportion of African nations is in the United Nations. They may also come up with examples of African nations that are not members of the United Nations. When subsequently asked for the estimate of the target value, participants are more likely to use the activated information than if they had not considered the anchor. The participants who

considered the anchor 10% therefore estimate that the proportion of African nations in the United Nations is lower than if they had not been presented with the anchor.

Various evidence for the selective accessibility model has been provided. The precise wording of the comparison question influences subsequent absolute judgments. When the comparison question asks whether the target value is higher than the anchor, participants give higher estimates of the target value than when the question asks whether the target value is lower than the anchor (Mussweiler & Strack, 1999b). According to the selective accessibility model, the former wording should lead to testing the corresponding hypothesis that the target value is higher than the anchor and it should thus lead to activation of information that is compatible with higher target values than when the comparison question asks whether the target value is lower than the anchor. Further evidence for the role of positive hypothesis testing comes from a study by Chapman and Johnson (1999), which showed that the anchoring effect is reduced by the instruction to consider differences of the target and the anchor, but it is not affected by the instruction to consider similarities; likely, because people consider similarities between the target and the anchor even without the instruction. Given that, according to the selective accessibility model, an anchor influences subsequent judgment through the influence of information activated by the comparison question, the anchoring effect should disappear if this information would have been activated even without consideration of the anchor. Bahník and Strack (2016) provided support for this prediction. For example, in one of their experiments, participants judged the average summer temperature in New York City. Participants' absolute judgments were affected by a preceding question asking for the comparison of the average annual temperature in New York City with a low anchor, but not with a high anchor. The latter anchor is likely to activate information related to high temperatures within the context of the whole year, which presumably largely overlaps with information that is associated with summer and would have been therefore used for the absolute judgment anyway. While the described studies provide indirect evidence for selective accessibility of information compatible with the anchor value following the comparison question, most direct evidence comes from studies trying to directly evaluate the accessibility of information.

A key study examining the information activated by an anchor showed that people are faster to categorize words that are associated with the anchor value (e.g., words associated with winter following a low temperature anchor; Mussweiler & Strack, 2000; see also Englich et al., 2006; Mussweiler & Englich, 2005). However, a recent study by Harris et al. (2019) failed to replicate this result. Furthermore, Harris et al. argue that the lexical decision task used in the study is not sensitive enough to reasonably measure any effect on categorization time that could follow from the consideration of an anchor. The key study demonstrating the activation of information compatible with the anchor value following the comparison question therefore does not seem to provide reliable evidence for the process posited by the selective accessibility model. In the present study, I use a simpler method to test whether information consistent with an anchor value is more accessible after presentation of the anchor. In particular, I directly ask participants to provide an example of a category to which the anchor is compared. For example, after comparing an average annual temperature with an anchor, I ask people to give an example of a month, hypothesizing that participants will be more likely to provide examples of cold months following a low anchor and warm months following a high anchor. This test assumes that positive hypothesis testing would directly lead people to come up with examples of the category whose values of the target attribute are in line with

the anchor value or that such examples are more likely to be provided due to their semantic association with other activated information. Such an assumption is in accord with studies showing that people are more likely to provide examples of a category to which they had been previously exposed and which are therefore more accessible (e.g., Hamann, 1990).

A somewhat similar design was used by Mussweiler and Strack (1999b, Study 4), who asked participants to list features of the target that came to their mind while answering two comparison questions. Two judges subsequently evaluated whether the listed thoughts indicated that the target value is high or low. The listed thoughts were more likely to indicate that the target value is high following high anchors in comparison to low anchors. While the study provides some evidence for activation of information compatible with an anchor, the study had a relatively low sample size (90 participants, out of whom only around one third listed the features), it used only two items, the compatibility of the listed thoughts with the anchor value was not assessed directly by the participants or by some objective standard, and even though the effect of anchors on the listed thoughts was significant, the statistical test yielded a relatively high p-value of about .03. The present study uses a much larger sample of participants (501) and a larger sample of items (10), which are also properly treated as a sample from a larger possible population of items by using random effects in mixed-effect models, allowing broader generalizability of the results (Bahník & Vranka, 2017; Yarkoni, 2020). Furthermore, the employed method allowed participants themselves to indicate whether the provided example is higher or lower in the evaluated attribute than the target value.

In sum, the present study uses a novel design to examine whether people activate information compatible with the anchor value. As a secondary goal, it indirectly assesses whether the activated information is used in subsequent judgment. In particular the study examines the following pre-registered hypotheses:¹

- H1) In comparison to a low anchor, when given a high anchor, participants are more likely to provide an example that they subsequently classify as higher than the average value for the category than lower than the average value. That is, the activated information is compatible with the anchor value.
- H2) An estimate of the target value is higher for high anchors than for low anchors. That is, the anchoring effect occurs even after a delay introduced between comparison and absolute judgments by the employed procedure.
- H3) Providing an example classified as higher than the average value for the category is associated with higher estimates of the target value in comparison to providing an example classified as lower than the average value for the category when the effect of the anchor is controlled for. That is, the activated information is associated with the answer to the absolute judgment question.

.

¹ The remaining tested hypotheses were not pre-registered and should be treated as exploratory.

Methods²

Participants

Five hundred and one participants ($Mdn_{age} = 22$, $IQR_{age} = 6$, 70% students, 61% women) were recruited from a laboratory subject pool to participate in a batch of unrelated studies, one of which was the present study. The experiment was administered in groups of up to 17 and it was conducted in Prague, Czech Republic in November and December 2019. The study was administered in Czech.

Materials

The study consisted of two parts. Each part comprised 10 trials, which related to 10 categories presented in the same order in both parts. In the first part, each trial consisted of a comparison question and an instruction asking for an example of the associated category. The comparison question asked participants to compare a certain attribute with a high or low anchor. For example, the question for the attribute "average temperature in Prague" asked "Is the average annual temperature in Prague higher or lower than 1°C?" for a low anchor, and the comparison was with 25°C when the anchor was high. The instruction for an example of the associated category followed the comparison question in each trial. For the above example, the category was "a month" and the question thus instructed: "Provide an example of a name of a month." The first part of the study thus presented the comparison question from the standard anchoring paradigm and assessed the information activated by the comparison question by directly asking participants for an example of a related category.

In the second part of the study, participants were asked absolute judgment questions for the 10 attributes and then compared their examples from the first part with the attributes. The absolute judgment question, for example, asked: "What is the average annual temperature in Prague (in °C)?" The participants were then asked: "Is the average temperature in [month] higher or lower than the average annual temperature in Prague?", where [month] was the example provided by the participant in the first part of the study. The second part of the study thus presented the absolute judgment question from the standard anchoring paradigm, allowing to test the anchoring effect. Moreover, it assessed whether the provided example was compatible with the anchor, or not. In particular, if the example was compatible with a high anchor, it should be judged to be higher in the estimated attribute than the target of the absolute judgment (and vice versa for a low anchor).

The 10 categories were selected so that: (i) they were likely to be known to participants; (ii) they were clearly defined; (iii) at least one example was likely to be known to participants; (iv) familiarity with examples of the category was likely not to be strongly associated with the attribute value of the example. These requirements were important to ensure that participants were always able to provide a suitable example which could then be compared with the target of the absolute judgment in the second part of the study. A high and a low anchor were chosen for each comparison question. The anchors were chosen so that they

² Pre-registration, software used for data collection, data, and analysis scripts can be found on: https://osf.io/9qbwv/

would be clearly lower or higher than the target value, and so that they still fell within a reasonable distance from values of the target attributes of possible examples. An anchor was randomly selected from the preselected low and high anchor on each trial. A list of the categories used in the study and anchor values for the associated items is shown in Table 1.

Category	Attribute	Target of the absolute judgment question	Low anchor	High anchor
Car brands	Price	Average price of a new car bought in the Czech Republic	200 thousand CZK	900 thousand CZK
Months	Temperature	Average annual temperature in Prague	1°C	25°C
Members of the Czech parliament	Age	Average age of the current members of the Czech parliament	25 years	80 years
EU countries	Population	Average population of EU countries	1 million	80 million
Occupations	Salary	Average salary in the Czech Republic	15,000 CZK	90,000 CZK
Foreign European capitals	Distance from Prague	Average distance of European capitals from Prague	300 km	2500 km
Czech regions	Unemployment rate	Unemployment rate in the Czech Republic	1%	10%
Mammal species	Weight	Average weight of mammals in Prague Zoo	1 kg	3,000 kg
First names	Length	Average length of a Czech first name	4 letters	9 letters
Planets	Orbital period	Average orbital period of Solar System planets around the Sun	100 days	25,000 days

Table 1. A list of categories used in the study. Note that some of the attribute names are shortened for simplicity.

Results

Examples

The effect of anchors on provided examples was tested using a mixed-effect logistic regression where valence of the anchor served as a predictor and classification of the example as higher than the average value for the category served as the dependent variable. Random intercepts for participants and items and random slopes for the anchor valence for items were included in the model. Contrary to H1, participants were *less* likely to classify their provided example as higher than the average value for the category when they were given a high anchor than when they were given a low anchor, z = -6.89, p < .001, OR = 0.48, 95% CI [0.38, 0.59]. Participants classified 38.5% of the provided examples as higher than the average value for the category after a high anchor and 56.4% after a low anchor. The effect was in the same direction for all categories and was significant for all categories apart from one (see Figure 1).

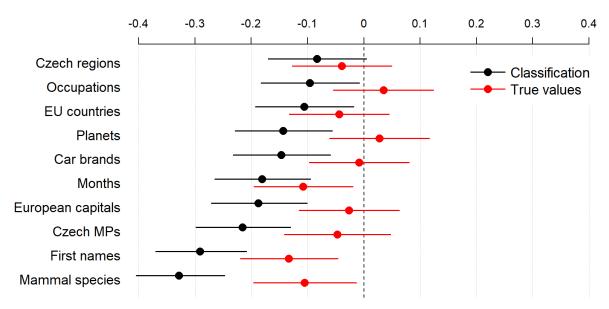


Figure 1. Association between provided examples and anchor valence by category. The figure shows Pearson correlations coefficients between an anchor condition and participants' classification of provided examples as well their true values. The error bars show 95% confidence intervals.

Given the unexpected result, I next searched for a true value of the attribute for each of the provided examples. For some of the categories, the values were easy to obtain (planets, names, months, cities, countries, municipalities, members of the parliament) while the true values for examples of other categories were more uncertain (mammals, car brands, and especially occupations). The results for the latter three categories should be therefore interpreted with caution. I computed a z-score for true values of examples of each category and conducted a mixed-effect linear regression with the z-scores as the dependent variable and valence of the anchor as a predictor. Random intercepts for participants and random slopes for the anchor valence for items were included in the model. The provided examples were somewhat lower in the target attribute for high anchors, t(9.0) = -2.47, p = .036, b = -0.089, 95% CI [-0.160, -0.018].³ Figure 1 shows that the effect was significant only for three of the categories, and the estimated effect was positive only for two categories.⁴ Even though the effect of anchors on provided examples was smaller when true attribute values were evaluated, the data still showed that people do not generally provide examples compatible with the anchor value.

While high anchors led people to provide examples that were lower on the target attribute, it is also possible that the anchor could have also directly influenced the classification of the examples. Given that high anchors lead people to estimate the target value to be higher, the provided examples might have been more likely to be classified as lower than this higher target value even if the examples actually had not differed in the target attribute. I therefore conducted a mixed-effect logistic regression with the classification of the example as the dependent variable and an anchor (using effect coding), the standardized true value of the

7

³ Using McCall's transformation (see footnote 5) instead of standardization with z-scores yields the same results.

⁴ The results were similar when the three categories with more uncertain true values were not included in the analysis, t(6.2) = -2.59, p = .040, b = -0.106, 95% CI [-0.186, -0.026].

provided example, and their interaction as predictors. Examples with higher true values were more likely to be classified as higher in the target attribute than the average value for the category, z = 31.44, p < .001, OR = 8.59, 95% CI [7.51, 9.82], showing that participants' classification was influenced by the true value of the provided examples. Even when taking the true values into account, higher anchors still led examples to be less likely to be classified as higher in the target attribute than the average value for the category, z = -6.51, p < .001, OR = 0.35, 95% CI [0.26, 0.48], showing a direct effect of anchors on classification. The interaction of the two predictors was not significant, z = 0.26, p = .796, ratio of OR = 1.03, 95% CI [0.80, 1.34].

Estimates

Two of the items asked for estimates in higher orders of numbers, i.e., thousands of CZK and millions of people. Following pre-registration, answers higher than 10,000 were divided before analysis by 10,000 and 1,000,000, respectively, assuming that no one would estimate the average price of a new car to be higher than 10 million CZK (~430,000 USD) and the average population of an EU state to be higher than 10 billion. The absolute judgments were transformed using McCall's transformation⁵ for each item before analysis to reduce the impact of outliers and transform all the items to the same scale.

The anchoring effect was tested using a mixed-effect linear regression where valence of the anchor served as a predictor and the absolute judgment of the target value served as the dependent variable. Random intercepts for participants and items and random slopes for the anchor valence for items were included in the model. Confirming H2, participants gave higher estimates when they were presented with a high anchor, t(9.0) = 10.61, p < .001, b = 0.839, 95% CI [0.684, 0.994].

The association between provided examples and absolute judgments was tested using a mixed-effect linear regression where valence of the anchor and classification of the example served as predictors and the absolute judgment of the target value served as the dependent variable. Random intercepts for participants and items and random slopes for the anchor valence and classification of the example for items were included in the model. Contrary to H3, when participants classified a provided example as higher in the target attribute than the average value for the category, they gave lower estimates than when they classified the provided example as lower in the target attribute than the average value for the category, t(9.0) = -2.79, p = .021, b = -0.141, 95% CI [-0.241, -0.042]. However, given that anchors directly influenced classification of examples, this analysis may confound the effect of anchors on classification with the effect of provided examples on estimates. I therefore conducted a mixed-effect linear regression with an anchor and standardized true values of the provided examples as predictors and absolute judgments as the dependent variable. Participants gave higher estimates for high anchors, t(9.0) = 10.92, p < .001, b = 0.833, 95% CI [0.684, 0.982], as well as when the provided examples had higher values, t(9.0) = 3.41, p= .008, b = 0.078, 95% CI [0.033, 0.123] (see Figure 2).

_

⁵ McCall's transformation first transforms the variable to percentiles and then transforms the percentiles to z-values corresponding to the percentiles for the standardized normal distribution.

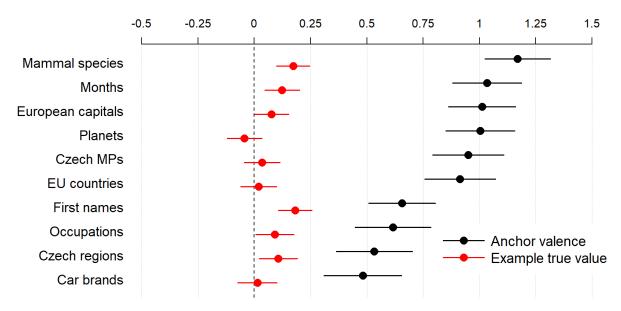


Figure 2. The effect of anchor valence and true values of provided examples on absolute judgments by category. The figure shows regression coefficients and their 95% confidence intervals for regressions with anchor valence and standardized true values of provided examples as predictors and absolute judgments as dependent variables.

Discussion

Contrary to a prediction derived from the selective accessibility model, examples provided after consideration of an anchor were not closer to the anchor in the target attribute. That is, anchors did not seem to make information compatible with the anchor value more accessible. This effect was observed using both participants' own classification of the target attribute of the example as well as using true values of the examples' target attribute. The selective accessibility model also predicts that the information activated after consideration of the anchor is used in subsequent judgment. In accord with this prediction, absolute judgments were influenced by anchors and positively associated with true values of provided examples.

Even though it is possible that consideration of an anchor does not make information compatible with the anchor value more accessible, there are also other possible explanations for the lack of the expected effect. It is possible that the information activated while answering the comparison question is propositional in nature and does not involve generating examples. Any effect on higher accessibility of examples, such as has been studied by Mussweiler and Strack (2000) and by the present study, would then be caused only by semantic priming. This semantic priming could be too weak to be reliably measured using the methods of the present study or by the lexical decision task used by Mussweiler and Strack (2000), as suggested by Harris et al. (2019).

It is also possible that the lack of the expected effect of anchors on generation of examples biased in the direction of the anchor value might have been due to the use of relatively extreme anchors. For both low and high anchors, participants answered the comparison question incorrectly only in 9% of trials. Mussweiler and Strack (2000) argue that the comparison question is answered using only knowledge about a general category when an

anchor is implausible. However, in the present study, comparison questions actually mostly asked about general categories, so it is not clear what kind of a more general category could have been considered while answering the comparison question. For example, the study did not ask about the price of a particular car model, but about prices of cars in general. It is doubtful that participants would answer the comparison question using prices of all vehicles or another more general category.

Unlike in other studies using non-random anchors, participants were not misled to believe that the anchors were randomly determined. Participants could have therefore believed that the anchors are informative (Schwarz, 1994). Yet, even if the anchors were considered to be informative, it is not clear why the process of comparison should have been different as a result. Perceived informativeness of the anchors could have contributed to their effect on absolute judgments, but their comparison to the targets of comparison should have been normally based on focusing at their similarities rather than differences as posited by the selective accessibility model (Mussweiler, 2003).

The examples of objects were classified by participants as lower than the average of the category in the target attribute following a high anchor and as higher than the average following a low anchor. However, the classification seemed to be directly influenced by anchors. For example, because high anchors increased participants' estimates of the target values, examples were more likely to be judged to be lower in the target attribute than this higher estimated value. Participants' classification of the examples cannot therefore speak to the activation of information compatible with the anchor value. However, the effect of anchors on provided examples was still in the same direction when true values for the provided examples were extracted, even though the effect was smaller and not consistently observed for all categories.

While true values of target attributes of provided examples were in the opposite direction from the anchor value, the p-value of the associated statistical test was only slightly below the significance threshold. The reliability of the effect should thus be confirmed with a replication study. Nevertheless, if the result is taken seriously, it is possible to speculate that participants might have searched for examples disconfirming the anchor value as a possible target value. While Mussweiler (2003) argues that focus on similarities is the default in comparison judgments, he acknowledged that people may focus on dissimilarities in certain situations. However, it is not clear why participants would focus on dissimilarities between the target and the anchor in this particular study.

The anchoring effect is a robust phenomenon (Bahník et al., 2017; Klein et al., 2014), so its successful replication is not surprising. Yet, the absolute question was asked in the present study only after other questions related to other targets. The present study thus provides further evidence that anchoring can be reliably shown even after a delay and possible interference by other numerical judgments. A similar finding has been shown in multiple studies (Bahník, 2021; Mussweiler, 2001; Yoon & Fong, 2019). The presence of anchoring after a delay and possible interference by other numeric judgments speaks against a simple numeric priming account of anchoring.

While the use of accessible information in judgment has been studied extensively in social psychology (Higgins, 1996), the evidence for the role of accessible information in the anchoring effect is mostly indirect. For example, Mussweiler and Strack (1999b) showed that

participants made the absolute judgment slower when they answered the comparison question under time pressure, presumably because when making the absolute judgment they had to activate information which was not activated to make the comparison judgment. Similarly, absolute judgments are slower when the anchor is implausible and answering the comparison judgment thus does not require activation of information relevant to the target (Mussweiler & Strack, 1999b; Strack & Mussweiler, 1997). The present study adds further evidence of the use of accessible information in anchoring, directly relating activated information to absolute judgments.

The results showed that true values of the considered attribute of the provided examples were positively associated with absolute judgments. It is possible that participants indeed used activated information in their absolute judgments as posited by the selective accessibility model, but that the assimilation of the absolute judgment toward an anchor values was not driven by the biased activation of information. That is, people may retrieve relevant information when making the absolute judgment, but the effect of an anchor may be driven by other means, such as scale distortion (Frederick & Mochon, 2012). However, it is also possible that the information activated by consideration of an anchor actually does not play an important role in the anchoring effect, and that the observed association was merely an artifact of the used method, which forced people to come up with an example, and thus activate relevant information, even if they would not have done so normally.

In sum, the study did not support the prediction of the selective accessibility model that anchors activate information compatible with the anchor value, but it is consistent with the prediction that participants use activated information in their absolute judgment. Given the results of the present study and the recent failure to replicate the study by Mussweiler and Strack (2000) by Harris et al. (2019), the evidence for the selective accessibility model deserves to be revisited. Future replications of existing findings as well as novel studies testing predictions of the model may help to further clarify the role of selective accessibility in the anchoring effect.

References

Bahník, Š. (2021). Anchoring without scale distortion. *Judgment and Decision Making, 16*(1), 131-141.

Bahník, Š., Englich, B., & Strack, F. (2017). Anchoring effect. In R. F. Pohl (Ed.). *Cognitive Illusions: Intriguing Phenomena in Thinking, Judgment, and Memory (2nd ed.)* (pp. 223-241). Hove, UK: Psychology Press.

Bahník, Š., Houdek, P., Vrbová, L., & Hájek, J. (2019). Variations on anchoring: Sequential anchoring revisited. *Judgment and Decision Making*, *14*(6), 711-720.

Bahník, Š., & Strack, F. (2016). Overlap of accessible information undermines the anchoring effect. *Judgment and Decision Making*, *11*(1), 92-98.

Bahník, Š., & Vranka, M. A. (2017). If it's difficult to pronounce, it might not be risky: The effect of fluency on judgment of risk does not generalize to new stimuli. *Psychological Science*, *28*(4), 427-436.

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

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.

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.

Frederick, S. W., & Mochon, D. (2012). A scale distortion theory of anchoring. *Journal of Experimental Psychology: General*, 141(1), 124-133.

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.

Hamann, S. B. (1990). Level-of-processing effects in conceptually driven implicit tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(6), 970-977.

Harris, A., Blower, F., Rodgers, S., Lagator, S., Page, E., Burton, A., ... Speekenbrink, M. (2019). Failures to replicate a key result of the selective accessibility theory of anchoring. *Journal of Experimental Psychology: General, 148*(9), e30-e50.

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social Psychology: Handbook of Basic Principles* (pp. 133–168). New York: Guilford Press.

Klayman, J., & Ha, Y. W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, *94*(2), 211-228.

Klein, R. A., Ratliff, K. A., Vianello, M., Adams, R. B., Jr., Bahník, Š., Bernstein, M. J., ... Nosek, B. A. (2014). Investigating Variation in Replicability: A "Many Labs" Replication Project. *Social Psychology*, *45*(3), 142-152.

Mussweiler, T. (2001). The durability of anchoring effects. *European Journal of Social Psychology*, 31(4), 431–442.

Mussweiler, T. (2003). Comparison processes in social judgment: mechanisms and consequences. *Psychological Review, 110*(3), 472-489.

Mussweiler, T., & Englich, B. (2005). Subliminal anchoring: Judgmental consequences and underlying mechanisms. *Organizational Behavior and Human Decision Processes*, *98*(2), 133-143.

Mussweiler, T., & Strack, F. (1999a). Comparing is believing: A selective accessibility model of judgmental anchoring. *European Review of Social Psychology*, *10*(1), 135-167.

Mussweiler, T., & Strack, F. (1999b). 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. (2000). The use of category and exemplar knowledge in the solution of anchoring tasks. *Journal of Personality and Social Psychology, 78*(6), 1038-1052.

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.

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.

Pohl, R. F. (Ed.). (2017). Cognitive Illusions: Intriguing Phenomena in Judgement, Thinking and Memory. Psychology Press.

Schwarz, N. (1994). Judgment in a social context: Biases, shortcomings, and the logic of conversation. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (pp. 123-162). San Diego, CA: Academic Press.

Strack, F., Bahník, Š., & Mussweiler, T. (2016). Anchoring: Accessibility as a cause of judgmental assimilation. *Current Opinion in Psychology*, *12*, 67–70.

Strack, F., & Mussweiler, T. (1997). Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility. *Journal of Personality and Social Psychology*, 73(3), 437-446.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124-1131.

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

Wong, K. F. E., & Kwong, J. Y. Y. (2000). Is 7300 m equal to 7.3 km? Same semantics but different anchoring effects. *Organizational Behavior and Human Decision Processes, 82*(2), 314-333.

Yarkoni, T. (2020). The generalizability crisis. *Behavioral and Brain Sciences*, 1-37. doi:10.1017/S0140525X20001685

Yoon, S., & Fong, N. (2019). Uninformative anchors have persistent effects on valuation judgments. *Journal of Consumer Psychology*, *29*(3), 391–410.

Yoon, S., Fong, N. M., & Dimoka, A. (2019). The robustness of anchoring effects on preferential judgments. *Judgment and Decision Making*, *14*(4), 470-487.