Judgment and Decision Making, Vol. 14, No. 6, November 2019, pp. 711–720

Variations on anchoring: Sequential anchoring revisited

Štěpán Bahník* Petr Houdek† Lucie Vrbová† Jiří Hájek†

Abstract

The anchoring effect, the assimilation of judgment toward a previously considered value, has been shown using various experimental paradigms. We used several variations of the sequential anchoring paradigm, in which a numeric estimate influences a subsequent numeric estimate on the same scale, to investigate how anchoring is influenced by multiple anchors, a comparison question, and by a newly introduced debiasing procedure. We replicated the anchoring effect using the sequential anchoring paradigm and showed that, when two anchors of opposite directions are presented, the second seems to influence a subsequent judgment somewhat more. A comparison of a target with another object before the numerical estimate was not sufficient to elicit anchoring, but it might have increased the sequential anchoring effect. The debiasing procedure, based on providing reference points on the numerical scale, prevented the sequential anchoring effect. The results are in accord with the scale distortion theory of anchoring, but other theories may also account for the observed findings with additional adjustments.

Keywords: anchoring, judgment, scale distortion, selective accessibility, debiasing

1 Introduction

The anchoring effect refers to the assimilation of judgment toward a previously considered value — the anchor (Bahník, Englich & Strack, 2017; Tversky & Kahneman, 1974). Anchoring has been usually studied using the so-called standard anchoring paradigm. In this paradigm, people are first asked to compare a target value with an anchor and then make an absolute judgment about the target value. For example, in a comparison question, people may be asked whether the average annual temperature in New York City is lower or higher than 102 °F. Then, they estimate the average annual temperature in New York City. The absolute judgment is assimilated toward the anchor value, and people thus provide higher estimates if they compared the temperature with 102 °F than if they had compared the temperature with a lower value. Anchoring has been shown to be robust (Klein et al., 2014) as well as relevant to various applied domains (e.g., Englich, Mussweiler & Strack, 2006; Galinsky & Mussweiler, 2001). Even though it has been studied for more than 40 years, its underlying mechanism is still debated (e.g., Bahník & Strack, 2016; Frederick & Mochon, 2012; Harris & Speekenbrink, 2016; Lewis, Gaertig, & Simmons, 2019; Mochon & Frederick, 2013). In particular, the recently introduced scale distortion theory of anchoring (Frederick & Mochon, 2012; Mochon & Frederick, 2013) is supposed to represent an alternative to the previously favored explanation of the anchoring effect—the selective accessibility model (Strack, Bahník & Mussweiler, 2016; Strack & Mussweiler, 1997).

According to the selective accessibility model, anchoring is a result of increased accessibility of information consistent with the anchor, which results from positive hypothesis testing. This means that in the standard anchoring paradigm people test, in a biased way, the hypothesis that the correct target value is equal to the anchor value when they answer the comparison question. In the example of the judgment of the average annual temperature in New York City, they would answer the comparison question by recalling information that is compatible with the two values (the average annual temperature in New York City and 102 °F) being equal (Klayman & Ha, 1987). Even when the recalled information is not consistent with the equivalence of the two values, it is subsequently more easily accessible in mind. Such information is therefore more likely to be used when answering the absolute judgment question, and thus bias the judgment in the direction of the anchor.

The scale distortion theory argues that the anchor value influences the perception of the scale on which the judgment is made (Frederick & Mochon, 2012; Mochon & Frederick, 2013). The comparison question introducing a high anchor of 102 °F would thus influence the subsequent judgment of the average annual temperature in New York City through the subjective change of perception of the Fahrenheit scale.

We would like to thank Jan Jolič for helping with development of the software used for data collection, Vojtěch Zíka, Kateřina Jiřinová and Markéta Sýkorová for research assistance and Marek Vranka for assistance with recruiting the participants.

The research was supported by Internal Grant Agency of Faculty of Business Administration, University of Economics, Prague (VŠE IP300040).

Materials, data, analysis scripts, and pre-registration of the study can be found on https://osf.io/qzsnf/ and https://osf.io/mnz45/.

*Faculty of Business Administration, University of Economics, Prague, náměstí Winstona Churchilla 4, Prague, 130 67, Czech Republic. Email: bahniks@seznam.cz.

†Faculty of Business Administration, University of Economics, Prague, náměstí Winstona Churchilla 4, Prague, 130 67, Czech Republic.

The same moderate temperature would seem relatively lower when the high anchor value was previously considered. Consequently, the perceived correct value of the average annual temperature in New York City would be mapped to a higher value on the scale. Among other evidence, Frederick and Mochon support their theory using a sequential anchoring paradigm, in which participants make two subsequent absolute judgments. The first judgment (e.g., of wolf's weight) serves as an anchor and influences the second judgment (e.g., of giraffe's weight), but only if they are on the same scale.

The two theories of anchoring give different predictions in various cases (see Bahník et al., 2017, for a summary of evidence for and against both theories). For example, the scale distortion theory predicts that the comparison question is not sufficient to elicit anchoring when the target is compared with an object rather than a numerical value because scale distortion requires consideration of a numerical value on the scale. On the other hand, such comparison question might influence the absolute judgment according to the selective accessibility model because people should activate information that is compatible with the target value being the same as the anchor value, even if it is not explicitly provided. Mochon and Frederick (2013, Study 1) tested the two opposing predictions and found that the comparison question without an explicit anchor value does not elicit the anchoring effect. However, Mochon and Frederick tested the effect using a single item in an experiment with limited statistical power. Chapman and Johnson (1994) also tested the effect of a comparison question and found a marginally significant effect of the comparison question, which could be consistent with the existence of the effect as well as with a null effect. Therefore, a replication of the finding could help further strengthen the objection against selective accessibility model or alleviate it if the anchoring effect was found even without the presence of an explicit anchor value. We thus use a larger sample of scales and a higher number of participants to replicate the apparent null effect obtained by Mochon and Frederick.

While the comparison of the target with another object may not elicit the anchoring effect, it could still increase the anchoring effect in the sequential anchoring paradigm. Consistent with this hypothesis, Mochon and Frederick (2013, Study 4) showed that a sequential anchor does not influence the target judgment if the two absolute judgments refer to dissimilar targets, but the anchoring effect reappears when a comparison question is introduced between them. Presumably, people are more likely to automatically compare similar targets and the comparison has to be externally elicited if the targets are dissimilar. Similarly, Harris and Speekenbrink (2016) showed that the comparison question may induce anchoring in the sequential anchoring paradigm even in the case that the anchor related to a different dimension (e.g., weight instead of height as the target judgment), in which case no effect of the anchor is otherwise observed. It is thus also possible that the comparison question may generally increase the anchoring effect because it makes people focus on similarities between the two targets as the selective accessibility model argues. The comparison question may thus increase the anchoring effect not only for dissimilar targets as Mochon and Frederick showed, but for targets belonging to the same category as well. We test this possibility in the present study.

Apart from the effect of a single sequential anchor, Mochon and Frederick (2013, Study 2) also showed that using two low anchors results in a stronger anchoring effect than just using one of the anchors. We follow up on this result by using both high and low sequential anchors before the target judgment. We manipulate the order of the two anchors to find out which of the two anchors influences the target judgment more. The inclusion of the second anchor could also be considered as a way of debiasing the sequential anchoring effect, similar to the strategy of considering the opposite, which was shown to reduce the anchoring effect in the standard anchoring paradigm (Mussweiler, Strack & Pfeiffer, 2000).

Given that the scale distortion theory argues that the anchoring effect is a result of inappropriate mapping of a judgment to a response scale, we developed a debiasing technique specifically targeting the assumed mechanism. The technique aimed to provide participants with reference points which could be used in mapping the value of the target judgment to the correct value on the scale, thus reducing the possibility of subjective distortion of the scale. To make participants think about the reference points, we asked them to estimate two values on the scale before introducing a sequential anchor. Notably, this manipulation should prevent the distortion of the scale and it should not therefore influence the estimate of the anchor value, only its subsequent effect on the absolute judgment of the target value.

In sum, we test several hypotheses related to the scale distortion theory and sequential anchoring paradigm. First, we examine whether a comparison of two objects on a scale is sufficient to influence a subsequent judgment of one of the objects on that dimension. Second, we test whether a comparison question increases the anchoring effect in a sequential anchoring paradigm. Third, we examine whether the first or the second sequential anchor influences the target judgment more. Fourth, we test a technique that may debias anchoring in the sequential anchoring paradigm. Finally, while previous anchoring research used mostly single items, we employ a larger set of items, which enables us to treat them properly as a random effect, and thus check the robustness and generalizability of some previously observed effects (Bahník & Vranka, 2017; Judd, Westfall & Kenny, 2012).

2 Methods

2.1 Participants

We recruited 424 Czech-speaking participants for a set of studies conducted on a computer in a laboratory. Participants were compensated with 250 CZK (~12 USD) for their participation in the whole set of studies, which took about two hours to complete. The study was conducted in groups of up to 17 participants and participants were invited for sessions some time in advance, so the sample size differed slightly from 400, which we planned to collect.[1] The set of studies included an instructional manipulation check (Oppenheimer, Meyvis & Davidenko, 2009) and three items in various scales that instructed participants to pick a specific response. According to a pre-registered exclusion criterion, we excluded from analysis 20 participants who failed to correctly answer the instructional manipulation check or at least two of the control items in the scales. We also excluded data from 16 additional participants from a single session where participants answered each question three times due to a software bug. Out of the remaining 388 participants, 70% were women. In terms of employment, 71% were students, 22% employed, and the remaining 7% had some other employment status. The median age of the participants was 23.5 years (IQR = 5).

2.2 Procedure and materials

Participants were presented with 22 trials which always ended with an absolute judgment question asking for a numerical judgment. This *target* judgment was used as the dependent variable in analysis. Before the target judgment, participants were given various questions depending on the condition in a given trial. All questions within a trial related to the same scale (kg, km/h, °C, . . . ; see Appendix for the full list). Each of the 22 trials used a different scale, the order of which was randomly determined for each participant. Only after each question was answered was a subsequent question displayed on the same screen. Participants were thus able to reply only to a single question at any given time. After making the target judgment, participants confirmed their response by pressing a button, which led them to the next trial on a different screen.

In total, there were 11 conditions and each participant received 2 trials for each condition (see Table 1 for an overview of all conditions). Assignment of the conditions to the scales was randomized for each participant. In the *control* condition, participants were given only the question asking for the target judgment (e.g., "What is the weight of a donkey in kilograms?"). In the *high* and *low sequential anchor* conditions, participants answered before the target judgment the same question about an object for which the correct answer was higher or lower than the correct value of the target (e.g., "What is the weight of an elephant in kilograms?" or "What is the weight of a fox in kilograms?").[2] In the *high* and *low sequential anchor with comparison* conditions, participants were additionally asked to compare the target to the high or low anchor between providing judgments about the anchor and the target (e.g., "Does an elephant weigh more or less than a donkey?" or "Does a fox weigh more or less than a donkey?"). The *high* and *low comparison anchor* conditions consisted only of the comparison question preceding the target judgment. The *high-low* and *low-high sequential anchor* conditions asked absolute judgment questions relating to both anchors before asking for the target judgment and they differed only in the order of presentation of the two anchors. Finally, *high* and *low sequential anchor with debiasing* conditions were the same as the sequential anchor conditions but included an additional instruction to mentally map two reference points on a scale before presenting the anchor. The participants were asked to estimate values of both anchors and imagine them on a numerical scale (e.g., "Estimate how many kilograms does an elephant and fox weigh and imagine these estimates on a numerical scale." or "Estimate how many kilograms does a fox and elephant weigh and imagine these estimates on a numerical scale.").

3 Results

3.1 Answers to the comparison question

In most of the scales, participants answered the comparison question correctly in more than half of the cases. For one scale (precipitation in mm per month), the majority of participants answered the comparison question incorrectly for both low and high anchor, and this scale was therefore excluded from analysis.[3] For all other scales, the answers

[1]The planned sample size was determined by the availability of funds. Power analysis is complicated for the design of the present study, but a study with 400 participants has a power .80 to detect an effect with a small-to-moderate size $d = 0.20$ for a two-sample t-test.

[2]We selected anchors for which we expected that most of the participants would be able to answer the comparison question correctly and for which we expected that participants' estimates would not differ by several degrees of magnitude from the correct value.

[3]This exclusion criterion was not pre-registered, but the effect of the anchors for this scale could not be interpreted as intended given that the low anchor was actually subjectively perceived as a high anchor and vice versa. Moreover, the results were virtually the same when the scale was included in analysis. For the voltage scale, only 56% of participants answered the comparison question for a high anchor correctly. However, 99% of participants answered the comparison question for a low anchor correctly, and 98% of participants who estimated both low and high anchors estimated the value of the high anchor as higher than that of the low anchor. Because the low anchor was perceived as lower than the high anchor, we kept the scale in analysis despite the relatively low number of participants answering the comparison question for the low anchor correctly. For all other anchors,

Table 1: An overview of experimental conditions with an example. The column on the right shows an example of the materials used for the weight scale (kg). See Appendix for the list of all the scales.

Condition	*Example of materials*
Control	What is the weight of a donkey in kilograms?
Low sequential anchor	What is the weight of a fox in kilograms?
	What is the weight of a donkey in kilograms?
High sequential anchor	What is the weight of an elephant in kilograms?
	What is the weight of a donkey in kilograms?
Low sequential anchor with comparison	What is the weight of a fox in kilograms?
	Does a fox weigh more or less than a donkey?
	What is the weight of a donkey in kilograms?
High sequential anchor with comparison	What is the weight of an elephant in kilograms?
	Does an elephant weigh more or less than a donkey?
	What is the weight of a donkey in kilograms?
Low comparison anchor	Does a fox weigh more or less than a donkey?
	What is the weight of a donkey in kilograms?
High comparison anchor	Does an elephant weigh more or less than a donkey?
	What is the weight of a donkey in kilograms?
Low-high sequential anchor	What is the weight of a fox in kilograms?
	What is the weight of an elephant in kilograms?
	What is the weight of a donkey in kilograms?
High-low sequential anchor	What is the weight of an elephant in kilograms?
	What is the weight of a fox in kilograms?
	What is the weight of a donkey in kilograms?
Low sequential anchor with debiasing	Estimate how many kilograms does a fox and elephant weigh and imagine these estimates on a numerical scale.
	What is the weight of a fox in kilograms?
	What is the weight of a donkey in kilograms?
High sequential anchor with debiasing	Estimate how many kilograms does an elephant and fox weigh and imagine these estimates on a numerical scale.
	What is the weight of an elephant in kilograms?
	What is the weight of a donkey in kilograms?

were correct in the majority of cases (56%-100% for high anchors and 88%-100% for low anchors, with averages of 89% and 96%). The comparison question was presented to participants in two conditions — comparison anchor and sequential anchor with comparison. It was therefore possible to test whether the initial absolute judgment of the anchor value influenced the answer to the comparison question. The results of a mixed-effect logistic regression with the correctness of the answer to the comparison question as the dependent variable and direction of the anchor (low X high) and condition (comparison anchor X sequential anchor with comparison) as predictors showed that there was no systematic effect of condition ($z = -1.57$, $p = .12$, $OR = 0.64$, 95% CI = [0.37, 1.11]). However, the interaction between condition and direction of the anchor showed that the effect of condition differed for high and low anchors ($z = -3.24$, $p = .001$, ratio of OR = 0.22, 95% CI = [0.09, 0.55]). While there was no effect for low anchors ($z = 0.82$, $p = .41$, $OR = 1.49$, 95% CI = [0.57, 3.89]) participants were less likely to answer the comparison question correctly for high anchors when it was preceded by an estimate of the anchor value ($z = -3.47$, $p < .001$, $OR = 0.31$, 95% CI = [0.16, 0.60]).

at least 76% of participants answered the comparison question correctly.

Table 2: Comparison of the anchoring conditions with the control condition.

Condition	*Estimate*	*CI*	*p*
Low sequential anchor	−0.100	−0.230 to 0.030	.147
High sequential anchor	0.150	0.037 to 0.264	**.016**
Low sequential anchor with comparison	−0.177	−0.314 to −0.040	**.019**
High sequential anchor with comparison	0.200	0.074 to 0.326	**.005**
Low comparison anchor	−0.058	−0.154 to 0.037	.241
High comparison anchor	0.014	−0.095 to 0.123	.804
Low-high sequential anchor	0.117	−0.006 to 0.240	.075
High-low sequential anchor	0.021	−0.101 to 0.143	.742
Low sequential anchor with debiasing	−0.038	−0.160 to 0.083	.544
High sequential anchor with debiasing	0.011	−0.106 to 0.128	.854

3.2 Differences between the control and anchoring conditions

Next, we performed pre-registered analyses using a mixed-effect linear regression. To allow inclusion of all the scales in one model, the target judgment was transformed to z-scores using the distribution of the target judgment in the control condition for each scale. That is, we computed in which percentile would a given value of the target judgment be if it was an answer in the control condition. We then used this percentile to compute the z-score. The first step served to alleviate the effect of outliers and to make the values comparable between scales. The second step made distributions of values close to normal and it reduced the impact of large differences in percentiles resulting from small differences around average values, where the estimates are more condensed. The z-scores were then used as the dependent variable in a mixed-effect model, in which dummy variables for conditions served as predictors (apart from the control condition) and which did not include an intercept. The estimated coefficient for each condition therefore shows the difference of the condition from the control condition. We included random intercepts for participants and scales as well as random slopes for scales. The results of the analysis are displayed in Table 2. It is possible to see that sequential anchors led to the anchoring effect, but the comparison question itself did not elicit the anchoring effect.

3.3 Sequential anchoring

For all subsequent analyses, we conducted similar mixed-effect regressions including only selected conditions to test specific hypotheses. All models included a full random structure for scales and random intercepts for participants. For models including four conditions (i.e., testing differences of effects of high and low anchors between two conditions), we report the interaction effect. All the results pertaining to the absolute judgment question of the target value are displayed in Figure 1.

First, the high and low sequential anchor conditions differed from each other, replicating the sequential anchoring effect ($t(20.3) = 3.44$, $p = .003$, $b = 0.247$, 95% CI = [0.106, 0.389]). Next, we assessed the effect of the two anchors in the high-low and low-high sequential anchor conditions. The answer to the second sequential anchor was influenced by the first anchor. That is, participants gave higher absolute judgments for the low sequential anchor when it was preceded by the high anchor ($t(1481.0) = 3.33$, $p < .001$, $b = 0.173$, 95% CI = [0.071, 0.274]), and to the high sequential anchor when it was *not* preceded by the low anchor ($t(19.4) = 1.74$, $p = .10$, $b = 0.118$, 95% CI = [−0.015, 0.251]), even though only the former effect was significant. While the effect was not significant, the low-high condition led to somewhat higher target judgments than the high-low condition ($t(17.9) = 1.78$, $p = .09$, $b = 0.098$, 95% CI = [−0.010, 0.206]), suggesting that, if there is any difference, the second anchor might influence the target judgment more. Correspondingly, presentation of the second anchor decreased the anchoring effect of the preceding, opposite anchor ($t(23.4) = -3.75$, $p = .001$, $b = -0.348$, 95% CI = [−0.529, −0.166]), but we did not find evidence for the influence of the first anchor on the anchoring effect of the subsequent, opposite anchor ($t(19.2) = -1.62$, $p = .12$, $b = -0.153$, 95% CI = [−0.338, 0.032]).

3.4 Debiasing

To test the effect of the debiasing procedure, we compared the sequential anchor with debiasing conditions with sequential anchor conditions. The interaction between condition and direction of an anchor showed that debiasing reduced the sequential anchoring effect ($t(45.5) = -2.57$, $p = .01$, $b = -0.202$, 95% CI = [−0.355, −0.048]). When low and

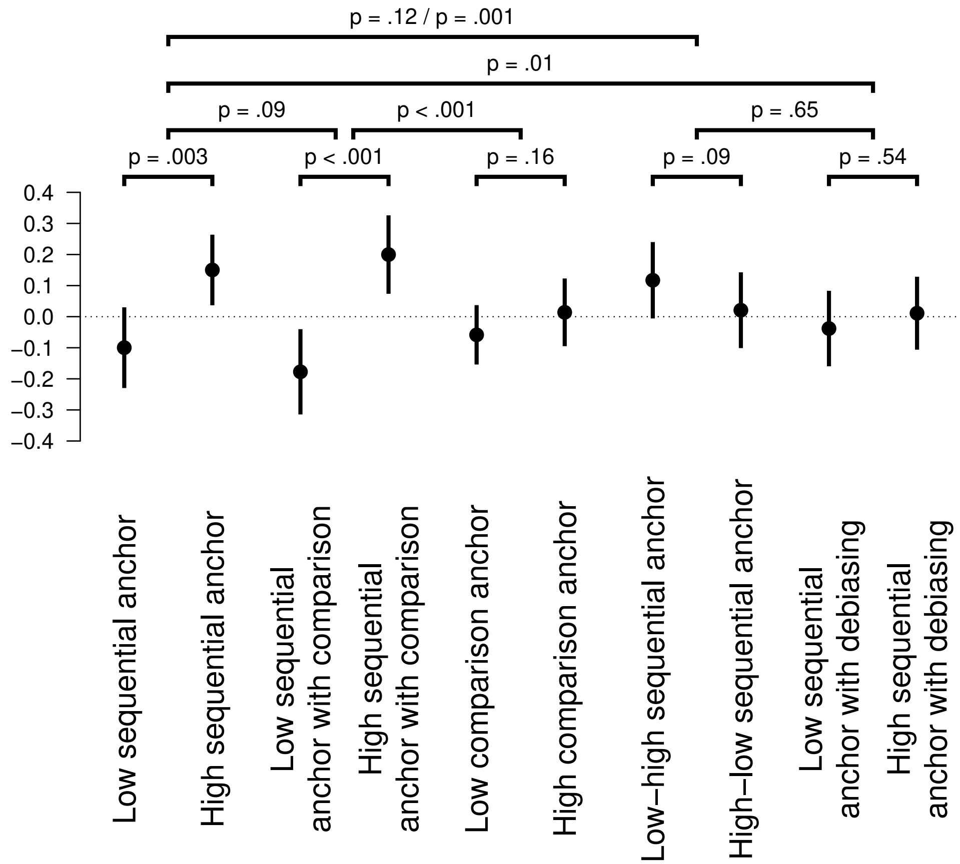

Figure 1: Results of analyses comparing absolute judgments of the target value. The points and error bars represent estimates of the effects and their 95% confidence intervals for comparison with the control condition, which are also reported in Table 2. At the upper part of the graph are results of all comparisons of absolute judgments of the target value reported in the text. The bottom line of results shows p-values for comparisons of pairs of conditions and the remaining results pertain to 2×2 interaction effects. A comparison of sequential anchors and two-anchors conditions can be performed in two ways, both of which are reported in the top line of results. The first analysis tests the effect of the first anchor and the second tests the effect of the second anchor.

high anchors were analyzed separately, anchoring effect was weaker for both the low sequential anchor condition with debiasing ($t(18.7) = -1.06$, $p = .30$, $b = -0.062$, 95% CI = [−0.177, 0.053]), and for the high sequential anchor with debiasing ($t(351.2) = 2.68$, $p = .008$, $b = 0.141$, 95% CI = [0.038, 0.244]) in comparison to the sequential anchor conditions even though only the latter effect was significant. Consistent with the lack of the anchoring effect after debiasing, the two debiasing conditions also did not differ significantly from each other ($t(19.5) = -0.63$, $p = .54$, $b = -0.044$, 95% CI = [−0.179, 0.092]).

The debiasing condition can be compared to the sequential anchor condition with two anchors since both conditions introduce both anchors, although in a different way, before the target judgment. We found no significant difference between the low-high sequential anchor condition and high sequential anchor condition with debiasing ($t(18.0) = 1.70$, $p = .11$, $b = 0.108$, 95% CI = [−0.017, 0.232]), as well as no significant difference between the high-low sequential anchor condition and low sequential anchor condition with debiasing ($t(19.2) = 0.94$, $p = .36$, $b = 0.058$, 95% CI = [−0.063, 0.179]). When the four conditions were analyzed together, we found no difference between the differences of the debiasing conditions and two-anchors conditions ($t(19.7) = -0.47$, $p = .65$, $b = -0.047$, 95% CI = [−0.246, 0.151]), suggesting that the effects of the debiasing manipulation and opposite anchor on subsequent sequential anchoring effect did not differ. However, unlike the two-anchor conditions, the debiasing conditions did not influence estimated values of the anchors. That is, the estimate of high anchors did not differ between the high sequential anchor with debiasing and high sequential anchor conditions ($t(19.2) = 0.01$, $p = 1$, $b = 0.000$, 95% CI = [−0.110, 0.110]). Similarly, the estimates of low anchors did not differ between the low sequential anchor with debiasing and low sequential anchor conditions ($t(1115.6) = -0.68$, $p = .50$, $b = -0.034$, 95% CI = [−0.134, 0.065]). On the other hand, the estimates of high anchors for the low-high sequential anchor condition were lower than for the high sequential anchor with debiasing condition ($t(18.1) = -1.93$, $p = .07$, $b = -0.121$, 95% CI = [−0.243, 0.002]), and the estimates of low anchors for the high-low sequential anchor condition were higher than for low sequential anchor with debiasing condition ($t(21.1) = 2.42$, $p = .02$, $b = 0.136$, 95% CI = [0.026, 0.247]). These results suggest that the debiasing manipulation worked through moderation of the effect of the anchor on the absolute judgment of the target value rather than mediation through its effect on the anchor.

3.5 Effect of the comparison question

The two sequential anchor with comparison conditions clearly differed between each other ($t(17.9) = 7.12$, $p < .001$, $b = 0.376$, 95% CI = [0.273, 0.480]). To find out whether the comparison question increases the anchoring effect in the sequential anchoring paradigm, we compared the effect of high and low anchors between the sequential anchor with comparison conditions and the sequential anchor conditions. Sequential anchors with the comparison question led to a somewhat stronger anchoring effect than sequential anchor conditions without the comparison question ($t(106.5) = 1.71$, $p = .09$, $b = 0.128$, 95% CI = [−0.019, 0.274]), but the effect was not significant. Separate analyses of high and low anchors showed that the inclusion of a comparison question led to a stronger anchoring effect both for the high anchor ($t(219.1) = 1.02$, $p = .31$, $b = 0.053$, 95% CI = [−0.049, 0.156]), and for the low anchor ($t(17.8) = -1.40$, $p = .18$, $b = -0.078$, 95% CI = [−0.187, 0.031]); however, neither effect was significant.

The comparison question was not sufficient to produce the effect by itself — the two comparison anchor conditions did not significantly differ ($t(566.5) = -1.42$, $p = .16$, $b = -0.070$, 95% CI = [−0.167, 0.027]). While the effect of the comparison question in sequential anchoring was not supported, the anchor itself clearly played a role, as seen by the difference of effects of high and low anchors between the comparison anchor condition and the sequential anchor with comparison condition ($t(75.9) = 4.13$, $p < .001$, $b = 0.304$, 95% CI = [0.160, 0.449]). Separate analyses for low and high anchors showed that the sequential anchor with comparison condition led to stronger anchoring effect both for low ($t(17.2) = -2.15$, $p = .05$, $b = -0.128$, 95% CI = [−0.245, −0.011]), and high anchors ($t(141.2) = 3.61$, $p < .001$, $b = 0.182$, 95% CI = [0.083, 0.281]).

4 Discussion

Using a set of scales, we replicated the sequential anchoring effect demonstrated by Mochon and Frederick (2013). We thus corroborated their findings and showed that they generalize to different stimuli as well. When two sequential anchors were used, the second anchor seemed to influence the target judgment somewhat more. The effect was, however, not significant and should be replicated. While the second anchor clearly influenced the target judgment, we did not find strong evidence for the effect of the first anchor. The data therefore did not fully corroborate the finding of Mochon and Frederick that two low anchors influence the target judgment more than one.

Similarly as Mochon and Frederick (2013), we did not find an effect of the comparison question alone when the target was compared with an object rather than with a numeric value as in the standard anchoring paradigm. Mochon and Frederick argued that the effect should have been observed even in this case according to the selective accessibility model. The selective accessibility model claims that the comparison question makes people activate information that is compatible with the anchor value being the true value of the target. When an anchor is an object, the mechanism should be presumably the same and the present finding would therefore pose a problem for the selective accessibility model. Yet, it is still possible that an object is compared differently to another object than to a numeric value. In the case of two objects, the comparison is symmetrical and the underlying mechanism could be therefore different than in the case of a comparison with a numeric value, where activation of information about an object might be much more natural — the numeric value is precise and might not need activation of additional information to be assessed. Nevertheless, the present study further corroborates the importance of a specific numeric value in the anchoring effect (but see also Oppenheimer, LeBoeuf & Brewer, 2008).

While Mochon and Frederick (2013) showed that a comparison question might make anchoring reappear for an anchor dissimilar from a target, they did not test the influence of the comparison question on the effect of an anchor similar to the target. Even though we found an effect consistent with the possibility that the comparison question increases the anchoring effect, the effect was not significant. Therefore, as Harris and Speekenbrink (2016, Experiment 1), we did not find evidence that inclusion of a comparison question increases the anchoring effect even when the anchoring judgment and target judgment are made on the same scale. Nevertheless, given the equivocal results, a future study should further test the possibility that the comparison question might increase the anchoring effect generally. It is possible that the comparison question makes people test the hypothesis that the target value is equal to the anchor value, as argued by the selective accessibility model. The initial absolute judgment of the anchor value would make the comparison question asymmetrical because then the anchor value would have been already estimated when the comparison was made, which could explain the difference in the effect of the comparison question with a preceding estimate of the anchor value and without it. It is also possible that the explanation used by Mochon and Frederick for the effect of the comparison question may apply even to an anchor that is similar to a target. Spontaneous comparison of the anchor to the target may not always occur even when they belong to the same category. The comparison question may therefore make this comparison more likely or its effect stronger and thus increase the effect of the anchor. These possible processes would be compatible with the effect of the comparison question which was somewhat indicated by previous studies as well as the present results.

Mussweiler et al. (2000) used the selective accessibility model to design a strategy for debiasing the anchoring effect. By asking people to consider arguments against an anchor value, they were able to reduce the anchoring effect in the standard anchoring paradigm. Here, we tested a procedure aimed at debiasing the anchoring effect in the sequential anchoring paradigm. Given that according to the scale distortion theory, people are influenced by an anchor because the anchor distorts the numerical scale on which the judgment is made, we asked participants to first imagine two numeric values on opposite ends of the scale to reduce scale distortion. We did not observe any effect of anchors after the manipulation. The debiasing manipulation included judgment about both low and high anchors, so the lack of the anchoring effect could have been caused by opposing anchoring effects of the two anchors, which was observed in the two conditions presenting the two anchors without the additional instructions used in the debiasing condition. However, the lack of an effect of debiasing instructions on the estimate of the anchor suggests that the debiasing manipulation worked through a different mechanism than presentation of an opposite anchor, which influenced the estimate. Yet, it is not clear whether the imagination of a numerical scale or consideration of two values on opposite ends of the scale debiased the anchoring effect. Even though the precise process which led to the effect of the debiasing instructions is a topic for future studies, the instructions could be used for debiasing the anchoring effect in the sequential anchoring paradigm and situations where the same underlying process operates.

While the present study focused on selective accessibility and scale distortion, there are other explanations of the anchoring effect as well. In fact, it is likely that the anchoring effect can occur through various mechanisms (Bahník et al., 2017; Turner & Schley, 2016). However, these explanations do not easily account for the effects observed in the sequential anchoring paradigm. According to one alternative view, an anchor can serve as a source of information and is used in subsequent judgment because it is seen as informative. Given that the anchor value is not provided externally in the sequential anchoring paradigm (unlike in the standard anchoring paradigm), it does not provide any new information to the participant. If the mere mention of a specific object was informative, we would also expect an effect in the comparison anchor conditions in our study, where we found none. A numeric priming account of anchoring (Wilson, Houston, Etling & Brekke, 1996; Wong & Kwong, 2000) could account for most of the results observed in the present study. However, it is not clear why the debiasing manipulation would prevent the anchoring effect if it was a result of numeric priming. If the effect of an anchor was debiased because participants were primed in the opposite direction by an estimate of the object on the opposite side of the scale, this estimate would have influenced the estimate of the anchor as well. Furthermore, numeric priming cannot explain other findings in studies using the sequential anchoring paradigm; for example, that the anchoring effect is observed only when the anchor and target judgments are made on the same scale and that the comparison question between the two judgments might influence the anchoring effect (Frederick & Mochon, 2012; Harris & Speekenbrink, 2016). Finally, anchoring could be a result of insufficient adjustment of judgment from the anchor value (Epley & Gilovich, 2001; Simmons, LeBoeuf & Nelson, 2010). However, it is not clear why two sequential anchors of the same direction, as used by Mochon and Frederick (2013), should have a stronger effect on the target judgment than one, when people can adjust only from one of the anchors. The anchoring-and-adjustment account also does not readily explain the effect of the debiasing manipulation in our study. These alternative accounts therefore do not present a parsimonious explanation for the anchoring effect in the sequential anchoring paradigm.

In this study, we replicated and extended findings of Frederick and Mochon (2012; Mochon & Frederick, 2013). As in their studies, we found the basic sequential anchoring effect. When two anchors were presented, the second seemed to influence the target judgment somewhat more even though it was itself affected by the first anchor. We replicated the finding that a comparison of a target with a different object is not sufficient to elicit the anchoring effect. A comparison introduced between two estimates on the same scale also did not reliably increase the anchoring effect of the first estimate. Finally, we developed a debiasing procedure, which was shown to prevent the anchoring effect in the sequential anchoring paradigm. Future studies could further elucidate the process underlying the effect of the debiasing procedure. The findings are potentially compatible with both the selective accessibility model and scale distortion theory; however, the selective accessibility model would need some revisions to account for all the observed effects. Namely, the selective accessibility model does not currently explain why a comparison of an object with a target does not elicit the anchoring effect. It is possible that an anchor in the form of a numerical value is required for positive hypothesis testing, which is argued to cause the anchoring effect by the selective accessibility model. The efficacy of the debiasing manipulation is also not easy to explain from the view of the selective accessibility model. Even though the scale distortion theory is most compatible with the findings of the present study, it is possible that selective accessibility or other processes used to explain the anchoring effect operate together or that they operate separately in different people or situations.

5 References

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., & 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.

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.

Harris, A. J. L., & Speekenbrink, M. (2016). Semantic cross-scale numerical anchoring. *Judgment and Decision Making*, *11*(6), 572–581.

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

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology, 103*(1), 54–69.

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 Jr., R. B., Bahník, Š., Bernstein, M. J., ... & Nosek, B. (2014). Investigating variation in replicability. *Social Psychology, 45*(3), 142–152.

Lewis, J., Gaertig, C., & Simmons, J. P. (2019). Extremeness aversion is a cause of anchoring. *Psychological Science*, *30*(2), 159–173.

Mochon, D., & Frederick, S. (2013). Anchoring in sequential judgments. *Organizational Behavior and Human Decision Processes*, *122*(1), 69–79.

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.

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, *45*(4), 867–872.

Oppenheimer, D. M., LeBoeuf, R. A., & Brewer, N. T. (2008). Anchors aweigh: A demonstration of cross-modality anchoring and magnitude priming. *Cognition*, *106*(1), 13–26.

Simmons, J. P., LeBoeuf, R. A., & Nelson, L. D. (2010). The effect of accuracy motivation on anchoring and adjustment: Do people adjust from provided anchors? *Journal of Personality and Social Psychology*, *99*(6), 917–932.

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.

Turner, B. M., & Schley, D. R. (2016). The anchor integration model: A descriptive model of anchoring effects. *Cognitive Psychology*, *90*, 1–47.

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, 125*(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.

6 Appendix

A list of scales used in the present study in a format "**scale (unit)**: low anchor, target, high anchor":

age (years): Jiří Ovčáček, Michal Horáček, Karel Schwarzenberg

altitude (metres above sea level): Amsterdam (Netherlands), Berlin (Germany), Lhasa (Tibet)

annual deaths in the Czech Republic (people/year): drowning, car accident, heart failure

area (ha): soccer field, Prague zoo, Ostrava

area (km^2): Austria, France, Canada

average annual temperature (°C): Oslo, Rome, Dubai

average precipitation (mm per month): Prague in March, Prague in September, Prague in July

calories (kcal): carrot, strudel piece, normal package of butter

decibels (db): whispering, normal speech, shouting

distance (km): Prague - Munich, Prague - Brussel, Prague - Athens

duration of a flight (h.): Prague - London, Prague - Tel Aviv, Prague - New York

fuel consumption (l/100km): average car, fully loaded truck, Boeing 737

GDP (billions of CZK): Somalia, Canada, Germany

length (cm): table tennis paddle, tennis racquet, hockey stick

movement speed (km/h): turtle, pig, tiger

population (): Belgium, Poland, Brasil

speed (rpm): gramophone record, car wheels on a highway, PC hard drive

unemployment (%): Germany, Italy, Greece

voltage (V): AA battery, electric chair, railway lines

volume (m^3): average car, subway carriage, airship

weight (kg): fox, donkey, elephant

year of birth (year of common era): Otakar II. of Bohemia, George of Poděbrady, Maria Theresa