Check for updates

the british psychological society
promoting excellence in psychology

Legal and Criminological Psychology (2021), 26, 215–227
© 2021 The Authors. Legal and Criminological Psychology published by John Wiley & Sons Ltd on behalf of British Psychological Society
www.wileyonlinelibrary.com

Combined Anchoring: Prosecution and defense claims as sequential anchors in the courtroom

Roland Imhoff* and Christoph Nickolaus

Social and Legal Psychology, Johannes Gutenberg University, Mainz, Germany, Binger Str. 14-16, Mainz, 55122, Germany

Purpose. When making judgements under uncertainty not only lay people but also professional judges often rely on heuristics like a numerical anchor (e.g., a numerical sentencing demand) to generate a numerical response. As the prosecution has the privilege to present its demand first, some scholars have speculated about an anchoring-based unfair disadvantage for the defence (who has the last albeit less effective word in court). Despite the plausibility of this reasoning, it is based on a hitherto untested assumption that the first of two sequential anchors exerts a greater influence on a later judgement (a primacy effect). We argue that it is also conceivable that the last word in court has a recency advantage (a recency effect) or that order does not matter as both demands even each other out (a combined anchor).

Methods. We report a pre-registered experiment with German law students ($N = 475$) who were randomly assigned to six experimental conditions in a study on legal decision-making order to test these three possibilities.

Results. Results indicate an influence of both the prosecution and the defence recommendation, but no effect of order.

Conclusion. This provides strong support for combined anchoring even for knowledgeable participants and rich case material. Specifically, the data are best compatible with the notion that both anchors exert an influence but each on different individuals. The implications of this finding for theory and legal decision-making are discussed.

When making judgements under uncertainty people often rely on heuristics to reach an accurate response. As one of the many instantiations of this phenomenon, a plethora of research has firmly established that a seemingly irrelevant previous number influences people's numerical estimates of a given quantity. In a seminal study, a previous comparative question including an arbitrary irrelevant anchor (determined by a wheel of fortune) biased subsequent judgements on an absolute question (Tversky & Kahneman, 1974). Ever since this initial demonstration, the anchor effect has received robust empirical support (for an overview Furnham & Boo, 2011). In fact, within the replication crisis in social psychology, the anchoring effect (Jacowitz & Kahneman, 1995) proved to be one of the few effects that did not only replicate well, but with effect sizes two to three times larger than in the original studies (Klein et al., 2014). In the present research, we

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

**Correspondence should be addressed to Roland Imhoff, Social and Legal Psychology, Johannes Gutenberg University, Mainz, Germany (email: roland.imhoff@uni-mainz.de).*

DOI:10.1111/lcrp.12192

20448333, 2021, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/lcrp.12192 by Cochrane Germany, Wiley Online Library on [17/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

seek to gain a better understanding of the anchoring process in a situation with two sequential anchors, as is typically the case in legal decisions involving sentencing demands (i.e. anchors) by the prosecution as well as the defence.

Such legal decisions are not the only examples of more realistic and less mundane decisions made by experts that are subject to anchoring. Real estate agents can be influenced in their estimates of what a house is worth (Northcraft & Neale, 1987) as can be car dealers in their evaluation of a used car (Mussweiler, Strack, & Pfeiffer, 2000). Making the first offer in negotiations – and thus setting an anchor – is an advantage as it increases (decreases) the final deal for sellers (buyers), respectively (Galinsky & Mussweiler, 2001). Numerical anchors drastically influence both patients and doctors in their likelihood judgements of becoming infected with or a patient having a certain condition (Brewer, Chapman, Schwartz, & Bergus, 2007).

Legal decision-making is no exception here and this is all the more noteworthy as the very basis of legal systems is rooted in an ideal of decisions uninfluenced by extra-legal influences. In simulated civil cases (with undergraduate students as participants), higher demands on side of the plaintiff were rewarded with higher compensations (Chapman & Bornstein, 1996). Anchors influence sentences for criminal cases even if the participants are legal experts like trainee lawyers completing their mandatory clerkship (Englich, 2005; Englich & Soder, 2009), or experienced professional judges or prosecutors (Englich, Mussweiler, & Strack, 2005, 2006; Guthrie, Rachlinski, & Wistrich, 2000), even if they have average professional experience of fifteen years (Englich & Mussweiler, 2001; Study 3).

Although the recruitment of such legal professionals is a major asset of these studies, other aspects of such experimental studies are less structurally comparable to the real life. For instance, many studies introduced the anchor in a highly artificial way. In one study, the anchor was delivered by a journalists' phone call in a break (e.g. Englich, Mussweiler, & Strack, 2006). In another study, a partisan heckler allegedly shouted into the courtroom (Englich, 2005). Arguably, such instances are comparatively rare in real court proceedings. Further, these studies then asked participants to first judge whether a given anchor (e.g. the sentencing demand by the prosecution or the partisan heckler) was too high or too low before openly asking about the sentence (e.g. Englich & Mussweiler, 2001; Englich, Mussweiler, & Strack, 2005; Englich et al., 2006; Englich & Soder, 2009). This is important as not only may consistency with one's own response to a comparative question alone explain anchoring effects (Grau & Bohner, 2014), but empirically, not demanding such a cognitive evaluation of a demand can make anchoring effects disappear (Englich, 2005). In a realistic courtroom, however, no judge is obliged to engage in this cognitive comparative process of scrutinizing whether an irrelevant anchor is too high or too low. A more naturalistic form of anchoring could be concluded from an effect of the sentencing demands made by one of the parties (in previous research almost exclusively the prosecution) without further instructions to evaluate it. As an important caveat, it seems fair to say that assimilation to such demands might also be interpreted as (rational) cue integration rather than (irrational) anchoring effects (as is also true for other more naturalistic 'anchoring' effects in applied setting, for example Mussweiler et al., 2000). This distinction is not at the centre of our work, but we will revisit it in the discussion of our findings.

Another limitation of the existing literature is that judges (i.e., participants) are typically confronted with only one demand – either by the prosecution or a clearly irrelevant source. In reality, however, judges hear the prosecution as well as the defence, thus being subject to (at least) two anchors. Conceivably, the first anchor given carries

20448333, 2021, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bjcp.12192 by Cochrane Germany, Wiley Online Library on [17/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

more weight on the final sentence, thus constituting a structural disadvantage to the defence. Englich et al. (2006) cite an unpublished manuscript in which the order in which the sentencing demands were presented (first prosecution, then defence vs. first defence, then prosecution) indeed had an effect on the outcome in form of higher sentences if the prosecution (i.e., the higher anchor) came first. Although not explicated, this reasoning implies a primacy effect of a greater weight given to the first anchor. Such a primacy effect is also in line with early work on order effects in persuasion (Lund, 1925). Supporting this notion, real sentencing decisions at an adult felony court in California were heavily influenced by the suggestion of a probation officer, which was the first information to the judge (Ebbesen & Konečni, 1981).

Although this seems plausible, another early empirical investigation observed an opposite recency effect, whereby the last piece of evidence had greater weight in final verdicts than prior one (Furnham, 1986; Weld & Roff, 1938). A meta-analytic integration of order effects on judgements concluded that 'we have today an assortment of miscellaneous variables, some of which tend to produce primacy, [. . .] others of which, to produce recency' (Rosnow & Robinson, 1967; p. 89). Other authors have indeed also identified conditions that more likely produce an advantage for information presented early vs. late in a sequence (Hogarth & Einhorn, 1992; Kerstholt & Jackson, 1998). Translating the recency effect to the role of sentencing demands in the court, it is at least conceivable that the sentence proposed by the defence does not bear less but more weight on the final verdict, as it is still vividly in mind, the advantage of the 'last word in court'.

In summary, despite court decisions being frequently cited real-life examples of anchoring, hitherto no empirical study in this context has sufficiently taken into account that the decision maker (the judge) is always exposed to two sequential anchors, not just one. It is thus unclear whether indeed the first anchor has a stronger influence than the second or vice versa. In principle, it is also conceivable that both anchors exert a combined influence, rather than one of the two. As an example of such combined anchoring, individuals might engage in cue integration and spontaneously generate the average of both demands as an anchor. Alternatively, some people may be influenced strongly by the first anchor, others by the second – a process that would yield similar responses in the aggregated mean, but not in the pattern of individual responses which would disperse more. The present research sought to address these possibilities empirically.

The present research

In the present research, we sought to test the influence of both prosecution and defence pleas as potential anchors. That means, we focused on numerical sentencing demands rather than the order of presented evidence. Specifically, we were interested in whether both recommendations may exert an influence on final sentences and whether the order in which the demands are presented matters. In addition to the possibility of a primacy effect (greater weight of the first anchor), we also entertained the possibility of a recency effect and an effect of the combined anchor of both without any relevance of order. Advanced law students knowledgeable in criminal law were asked to take the perspective of a judge and read a case file with sufficient detail to rest the sentencing verdict on other information than the respective anchors. Two anchors were given in the form of demanded sentences of prosecution and defence. Both the exact demands and their order of presentation were manipulated. According to the primacy account, the final verdict

20448333, 2021, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bjcp.12192 by Cochrane Germany, Wiley Online Library on [17/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

should be a (linear) function of the first anchor provided. In contrast, the recency account would predict the verdict to be a (linear) function of the second anchor provided. The combined anchoring account predicts the verdict to be function of the average of both anchors. We pre-registered patterns of results compatible with each of these accounts as well as our analytical strategy (https://aspredicted.org/5f465.pdf). We report all measures, observations and exclusions. All materials, raw data, and analysis script for the current study as well as two pilot studies are available at https://osf.io/4um8q/.

Method

We gave advanced students of law (recruited in university lectures) a case file with relevant information (e.g., witness report, police report, relevant laws, prosecution and defence speeches) and asked them to decide what verdict they would give, were they the responsible judge. Importantly, we refrained from introducing constraints to the ecological validity of the task (e.g., forcing participants to answer a comparative question about the anchors before giving their open-ended sentence; enforcing an ecologically invalid constraint to revisit earlier information).

Design and hypotheses

Our experiment followed a 2 (order of final speeches) by 3 (specific combinations of prosecution and defence demands) between-subjects design with participants randomly assigned to one of these six conditions. The experiment was conducted to critically test a) the existence of any order effect and b) more specifically the respective plausibility of a primacy effect (stronger influence by the first anchor), a recency effect (stronger influence by the later anchor), and a combined anchor effect (strong influence by the combined anchor, independent of order). To do so, we created a total of six conditions on a case of manslaughter. This specific delict was chosen as the German Criminal Code (StGB) has a particularly broad window on sentencing recommendation for this felony (1 year to 10 years, in severe cases up to 15 years, Section 212 German Criminal Code and Section 213 German Criminal Code). Thus, we could create both high and low anchors for the prosecution demand (12 years, 9 month; 8 years, 9 months) as well as the defence demand (6 years, 3 months; 2 years, 3 months) without including implausible demands. These demands were combined to create three conditions (Table 1).

Conditions A and B both included the high defence anchor, but differed in the prosecution anchor, and thus allowed to establish an anchoring effect under standard conditions. A third condition (Condition C) was created by combining the high prosecution anchor with a particularly low defence anchor so that their average would be identical to that of Condition A. The comparison of this condition with the one with an identical prosecution anchor but a higher defence anchor (Condition B) would speak to the relevance of the second anchor. Three additional conditions were identical in demands but in reversed order with the defence coming first and the prosecution second. Importantly, all three theoretical accounts (primacy, recency, and combination) predicted a specific pattern of results that we pre-registered together with our exclusion and data analysis plan (http://aspredicted.org/5f465.pdf):

A primacy account would predict the anchor to be a function of whatever sentencing demand is made first, thus predicting highest sentences for Conditions B and C. A recency account, in contrast, would attribute changes in the sentence to whatever

20448333, 2021, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bcp.12192 by Cochrane Germany, Wiley Online Library on [17/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1. All experimental conditions (including the realized average of demands, the combined anchor)

Condition		Demands			Theoretical Predictions		
		1st	2nd	Ø	Primacy	Recency	Combined
A	1st: Prosecution, 2nd: Defence	105	75	90	2nd highest	3rd Highest	Lowest
B	1st: Prosecution, 2nd: Defence	153	75	114	Highest	3rd Highest	Highest
C	1st: Prosecution, 2nd: Defence	153	27	90	Highest	Lowest	Lowest
D	1st: Defence, 2nd: Prosecution	75	105	90	3rd highest	2nd Highest	Lowest
E	1st: Defence, 2nd: Prosecution	75	153	114	3rd highest	Highest	Highest
F	1st: Defence, 2nd: Prosecution	27	153	90	Lowest	Highest	Lowest

Note. All experimental conditions, the order of final speeches, the expressed demands (in months), the average of demands (in months) and the theoretical prediction from three diverging accounts (from highest to lowest). For the primacy account, the verdict follows linearly from the first demand, for the recency account it is a function of the second demands, for the combined account a function of the average demand.

recommendation was heard last, thus predicting the highest sentences for Conditions E and F. A combined anchor account (an influence of the arithmetic mean of the two sentencing recommendations) would suggest the highest sentences for conditions B and E, with no differences between the other conditions.

The order was manipulated by placing the demands in a specific order in the test leaflet. Specifically, the first demand was printed on top of one page, while the second demand was printed on the back of the same page. Moreover, participants were instructed to read the leaflet in sequential order. Both this instruction and natural reading habit would make it unlikely that participants were initially exposed to the anchors in an order different from the intended one. In addition, the layout guaranteed that they could not see both anchors at the same time. This change in layout was the only difference from the first to the second pilot study. We did not constrain participants' ability to revisit the respective anchors at a later time. While this may have attenuated strict experimental control overexposure to information, it provides an ecologically more valid test of the hypothesis that sequential order of final speeches constitutes an unfair advantage in court. In German courts, judges are expected to take notes during the trial as they are responsible for an accurate recording of the trial. Any relevant information (including the exact demands by the parties as a central one in this regard) will thus be taken note of and judges will have this information in front of them until they decide on the verdict. Any ecologically meaningful anchor effect will thus have to bias the way a judge mentally construes a case, which will then give rise to confirmatory information processing in line with this initially biased representation. Order effects that only show under strict experimental control of sequential exposure (e.g. in a computer-aided experiment) may be of interest for basic research, but they would not constitute a fair test of order effects in an actual courtroom.

Sample

As explained above, a critical examination of a realistic anchoring effect requires that the power of an anchor is observable even if there are other factors possibly influencing a decision. Specifically, what is needed is sufficient and detailed case information (see below) and participants that bring the background knowledge with them that allows

20448333, 2021, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/lcrp.12192 by Cochrane Germany, Wiley Online Library on [17/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

them to base their decision on something else than then numerical anchor. We therefore recruited participants in an advanced lecture of criminal law and pre-registered to sample at least 300 participants to have at least 50 participants per condition. This equipped us with 80% power to detect effect sizes of $d = 0.50$ or larger for each of the planned comparison between two conditions. As we recruited law students in lectures of criminal law, however, the numbers were subject to influences outside of our control (number of students present and willing to participate). As pre-registered, we recruited as many participants as seemed feasible and we collected data in four different classes at different universities in different cities (*blinded for peer review*), which resulted in an initial sample of 507 students. As pre-registered, we excluded all participants who had no strong expertise in criminal law (i.e. students from other disciplines; $n = 24$) and participants who indicated to have prior knowledge of anchoring effects (when probed about the goal of the study in an open-ended manner, they spontaneously generated a suspicion about anchoring effects; $n = 9$), leaving a final sample of $N = 475$. We refrained from collecting any demographic data to maximize anonymity (given the classroom setting, age or a non-binary gender identification might be sufficient to identify a participant) and thus increase willingness to participate.

Material

All participants received a printed four-page booklet in German language. They were instructed to take the perspective of a judge who has to issue a decision (i.e., pass a sentence) in a case of manslaughter. It was stressed that there is no strictly right or wrong answer, but that all parties agreed that the defendant was responsible for his actions and that the case at hand constituted one of manslaughter. Further, it was made clear to the participants that they should issue a prison sentence, not suspended on probation. In the following, they received a file consisting of a police report, a doctor's report, the bill of indictment, two short witness reports, and a statement by the defendant as well as his criminal record. In this case, the defendant made a surprise visit to his partner's town where he entered a bar and saw her kissing another man, the victim. The defendant then hit a glass bottle over the head of the victim, who died soon thereafter from cerebral haemorrhage in the hospital. The next two pages presented the closing arguments of the prosecution and the defence, including the manipulation of the independent variable. Finally, participants received excerpts from the legal code including the three relevant paragraphs (Section 212 German Criminal Code (Manslaughter), section 213 German Criminal Code (Manslaughter under mitigating circumstances), and section 46 (Principles of sentencing)), before they could enter their (open-ended) response to the question 'On what sentence do you decide?' at the bottom of that page. Prior to collecting data, several professional lawyers provided guidance and comments. All agreed that the case and the material were highly realistic.

This study met all criteria to be exempted from ethics committee approval (no vulnerable population, no harm, no deception, no conflict of interest, no personal information, or physiological data recorded) according to local regulations.

Results

All responses were recoded to months as numerical values. We had not pre-registered how to deal with responses that spanned a range rather than giving a point value (e.g., '9 to

20448333, 2021, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bjcp.12192 by Cochrane Germany, Wiley Online Library on [17/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 2. All experimental conditions (including the realized average of demands, the combined anchor), and resulting average final sentences

Condition		Ø	*n*	*M*	*SD*
A	1st: Prosecution 8 years, 9 months (105 months) 2nd: Defence 6 years, 3 months (75 months)	90	78	89.46	14.11
B	1st: Prosecution 12 years, 9 months (153 months) 2nd: Defence 6 years, 3 months (75 months)	114	76	102.86	26.73
C	1st: Prosecution 12 years, 9 months (153 months) 2nd: Defence 2 years, 3 months (27 months)	90	79	84.65	30.93
D	1st: Defence 6 years, 3 months (75 months) 2nd: Prosecution 8 years, 9 months (105 months)	90	80	87.93	21.13
E	1st: Defence 6 years, 3 months (75 months) 2nd: Prosecution 12 years, 9 months (153 months)	114	81	102.63	26.32
F	1st: Defence 2 years, 3 months (27 months) 2nd: Prosecution 12 years, 9 months (153 months)	90	79	85.77	28.36

10 years'). We thus conducted analyses with the average value of the upper and lower bound of these ranges as well as without these participants. As the results do not differ, we present analyses for the full data set (with calculated averages for range responses). An ANOVA revealed an omnibus effect of difference between the groups, $F(5, 467) = 8.57$, $p < .001$, $\eta_p^{2^2} = .08$ (Table 2 for mean sentences per condition).

To follow-up on this, we conducted simple *t*-tests for each pair of the conditions. As each of our three explanations predicted a specific pattern, these were treated as conjunctive rather than disjunctive tests. That means, that each theory made a specific predictions which of the fifteen possible contrasts comparing six condition would have to be significant and which not (see Table 1). Any failure of a single contrast being significant or an unpredicted contrast being significant would violate the theoretical predictions in total (sometimes also referred to as intersection union test, e.g., Neuhäuser, 2006). Thus, the corroboration of the prediction by chance would happen at a likelihood that results from the conjunctive likelihoods of all fifteen tests (alpha for predicted contrast, beta for non-predicted contrast). Given that alpha and beta will always be below 1, this conjunctive product will have to be lower than each of the entered factors. This yields a correction for multiple tests unnecessary (as it is essentially one for a pattern, not multiple independently interpreted tests). Were we to interpret each significant contrast on its own, independent of all other, this would constitute a case of disjunctive hypotheses or tests, which can indeed inflate alpha error. This was not the case for the current study, so we pre-registered not adjusting the alpha. Note, however, that the results were so clear that even conservative correction for multiple testing would leave all interpretations unchanged.

Fully in line with the notion of a combined anchor, participants' responses were largely a function of the average of both anchors (Figure 1). As predicted by this account, Conditions B and E were not different from each other, $t(155) = 0.05$, $p = .958$, Hedges' $g = 0.01$, but they were significantly larger than all other conditions, all *p*s <.001. Specifically, Condition B led to higher sentences than Condition A, C, D, and F (Table 3). Likewise, Condition E (which differed from Condition B only in the order the closing arguments were presented) led to higher sentences than Condition A, C, D, and F. The responses in the four conditions with the low average value (Conditions A, C, D, F) were

20448333, 2021, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bjcp.12192 by Cochrane Germany, Wiley Online Library on [17/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

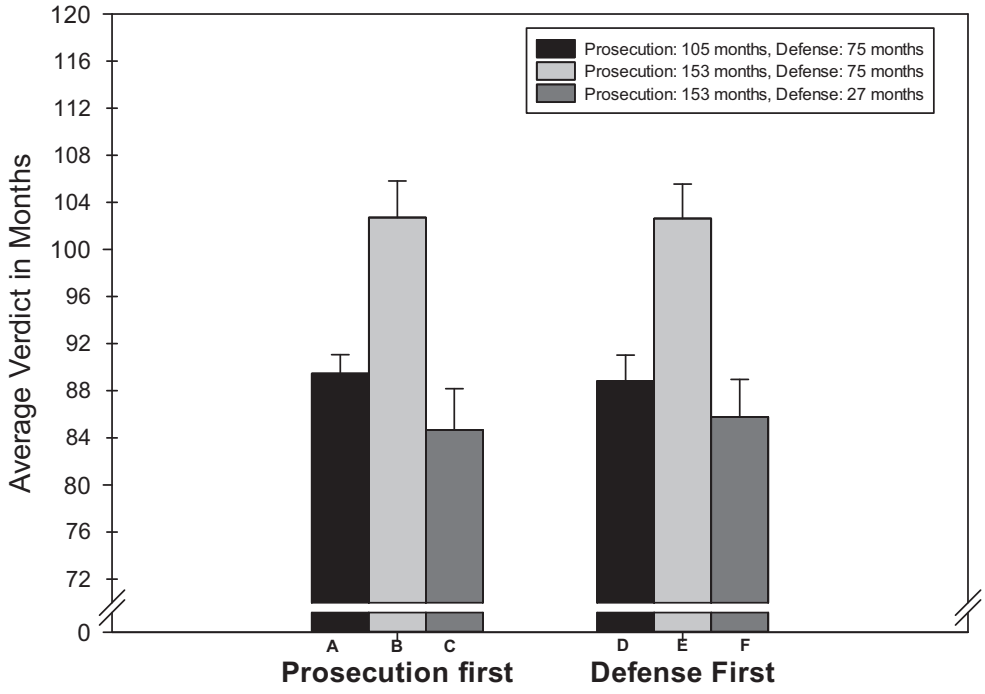

Figure 1. Average sentences (+ *SE*) as a function of experimental condition in Experiment 3.

Table 3. Simple contrasts pre-registered for combined anchoring accounts. Degrees of freedom adapted to accommodate unequal variances

	t	*df*	Hedges' *g*
Condition B vs. A	3.87	113.109	0.629
Condition B vs. C	3.92	153	0.629
Condition B vs. D	3.86	142.725	0.621
Condition B vs. F	3.86	153	0.620
Condition E vs. A	3.95	123.418	0.620
Condition E vs. C	3.96	158	0.627
Condition E vs. D	3.91	152.640	0.615
Condition E vs. F	3.90	158	0.616

all not significantly different from each other, $ps > .21$, Hedges' $gs < .20$. This pattern of results is exactly the one pre-registered for the combined anchor account.

We also conducted control analyses to account for data dependency due to data collection in four separate classrooms and hence treated individual data as nested in classrooms in a multi-level analysis. Specifically, we ran a mixed linear model to predict the sentence with the respective prosecution recommendation and defence recommendation (in months) and the dummy-coded order as well as the interaction of the two recommendations and order as fixed effects, including random slopes and intercepts for classes. The results revealed a significant contribution of the prosecution request,

20448333, 2021, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bjcp.12192 by Cochrane Germany, Wiley Online Library on [17/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

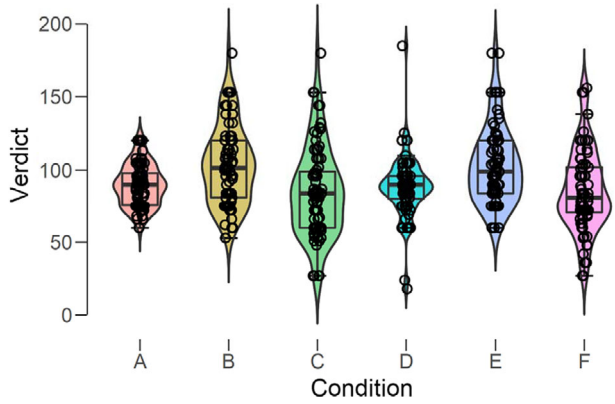

Figure 2. Violin plots for sentences as a function of experimental condition in Experiment 3. Thick horizontal lines indicate the median, the surrounding box the 1st and 3rd quartile. Dots indicate individual responses.

$B = 0.293$, $SE = .059$, $p < .001$, and defence request, $B = 0.365$, $SE = .059$, $p < .001$, but neither a main effect nor an interaction with order, $ps > .81$. Intercepts and slopes were virtually identical for all four classrooms.

As argued above, the combined anchoring account leaves two possibilities through which both anchors can exert an influence. The first one presupposes that each participant averages both anchors and uses this combined number as an anchor, whereas in the second possibility, this combination takes place at an aggregate level. Some people align with the high anchor, others with the low one. If the former is true, only the average of the two demands should play a role, independent of how far apart they are. If the second proposition is valid, we would not expect the distance between the anchors to influence the mean, but the dispersion of final sentences. Conditions A/D and C/F do not differ in their arithmetic mean of demands (90 months in each case), but in the distance between the demands (30 months in A/D vs. 126 in C/F). The data show that they are not significantly different at the mean level, but that the standard deviation was substantially larger for the high distance conditions (Table 2). Figure 2 depicts a violin plot that provides a detailed depiction of response distribution. Formally testing for homogeneity of variances yielded significant Levene tests comparing condition A to condition C, $F(78, 77) = 28.24, p < .001$, as well as condition D to condition F, $F(78, 79) = 10.35, p = .002$. Across all six conditions, the standard deviation within the conditions increased with the distance between both anchors, $r(4) = .89$, $p = .018$. Thus, it seems that indeed both anchors pulled in the respective direction, independent of the order in which they were presented, speaking to the notion that the combined effect happens at the aggregate level and not within each participant.

Discussion

A large-scale experiment with legally knowledgeable participants and rich case materials that provides the opportunity to make an informed decision provides strong evidence for a combined anchoring account. The final sentence was highest when the average of both demands was higher. When their arithmetic means were the same, the average sentences did not differ. This speaks strongly against an unduly high influence of one of the two

20448333, 2021, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bcp.12192 by Cochrane Germany, Wiley Online Library on [17/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

anchors consistent across participants (Englich et al., 2006). Our results are best reconcilable with a view that the two sequential anchors both exert an antagonistic effect in opposite directions. Rather than cancelling each other out, it seems they pulled different participants in different directions. A strict cue integration account would assume individuals to integrate the relevant information and come up with a judgement based on the outcome of this integration. In other words, judges could take the middle between the two demands and adjust from there. Our observation of a strong correspondence of distance between the demands and the dispersion of sentencing decisions seems to suggest that the combined anchoring effects rest on integration across not within participants: Some adjusted from the low anchor, others from the high anchor. Translating this possibility of combined anchoring to the courtroom would suggest that very disparate demands make the judge's final verdict more volatile.

The fact that the observed combined anchor effect was not moderated by the order of the closing arguments (and included sentencing recommendations) speaks against an unfair advantage that has frequently been raised (Englich et al., 2005). Quite on the contrary, the defence can in principle adjust its sentencing recommendation strategically to counter an anchor set by the prosecution. This is of course only true to the extent that defence requests are not already anchored by the prosecution, but the experience and preparation of professional lawyers speak against the notion that they will easily fall prey to unintended adjustment to an anchor set by the prosecution.

One important difference that differentiates our study from some previous studies on anchoring is the fact that the anchors were not presented as random numbers (Englich et al., 2005), but as actually relevant information (the official sentencing request by one of the parties). In classical anchoring studies, great care is taken to rule out that the use of the anchor can be seen as a valid strategy (e.g., by having dice or a wheel of fortune determine the exact anchor). The same is not true for our experiment and the empirical reality. To maximize ecological validity, we intentionally refrained from introducing such artificial conditions. Contrary to earlier claims that the relevance of an anchor does not matter (Englich & Mussweiler, 2001), a prosecution or defence sentencing demand as a highly relevant anchor produces larger anchoring effects than clearly irrelevant anchors (which produced no anchoring at all; Glöckner & Englich, 2015). This, however, introduces the possibility that the observed effect is not an (irrational) insufficient adjustment, but judges might (rationally) accept both prosecution and defence as other legal experts and see their requests as valid cues.

Further, the prosecution request might serve as an upper bound of plausible sentences and judges may perceive prosecutors as coming from the 'same side'. In contrast, adjusting proportionally in the direction of the defence request might be seen as instrumental to achieve 'peace under the law', a decision that both sides can accept. In fact, archival data on court proceedings show that both prosecution and defence sentencing requests have an influence on the final sentences (Schünemann, 1988). Although potential confounds with the severity of the incriminated act rule out a causal interpretation, the data provide real-life patterns similar to the ones experimentally observed here.

Of basic theoretical relevance, our study is – to our best knowledge – the first to explore a situation of two sequentially given anchors empirically. Although studies in the context of the sequential anchoring paradigm (Mochon & Frederick, 2013) operate with more than one anchor, these are typically not externally given, but self-generated by prior judgements. In addition, the majority of studies in this paradigm have provided sequential stimuli that pull judgements in identical direction, thus creating no conflict between two

20448333, 2021, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bjcp.12192 by Cochrane Germany, Wiley Online Library on [17/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

sequential anchors (for an exception see Bahník, Houdek, Vrbová, & Hájek, 2019). Apart from this paradigm, research using the more traditional external presentation of a numerical anchor so far has mainly focused on the less complex and clearer situation of a single anchor. In reality, however, it seems implausible that humans only encounter isolated single numerical anchors that can bias their decisions. Extrapolating from our findings it seems plausible to assume that numerical estimates in an information ecology rich with potential numerical anchors are more likely a function of their overall average than any specific single anchor – at least at the aggregate level.

Limitations and future directions

There are some limitations that are important to bear in mind. Importantly, all conclusions were based on hypothetical, not actual, sentences. This highly likely leads to less elaborated processing of the case material thus strengthening the impact of heuristic information processing. As another limitation, as our participants read the materials in a self-paced manner (and some might have flipped back and forth between the two). Arguably, this might have attenuated order effects that would actually show. If that were the case, however, we would argue that this is also a very likely constellation in the field where judges are free and actually highly likely to have the demands of both parties on a written note in front of them. Ruling this out experimentally might increase experimental control, but undermine ecological validity.

As the greater dispersion of responses in conditions with far apart anchors suggested, participants did not integrate these two anchors into one average before reaching a decision, but individuals differed in which anchors had an influence on them. There is only sparse literature on individual differences in the susceptibility to anchoring effects per se (McElroy & Dowd, 2007), but no prior research that can help determine which of two sequentially presented anchors will exert a (greater) influence. Future research may elucidate this interesting and relevant question.

From an applied perspective, our research points to the important question whether it may be more instrumental for defence lawyers to provide a particularly low numerical anchor (that can bias a combined anchor) or demand an acquittal (unlike in common law, lawyers in Germany have one final statement where they can demand either). Although the latter may seem more desirable as the ultimate outcome, it may have the downside of not having the same potential to serve as an antagonistic anchor to the prosecution demand.

Conclusion

In summary, our experiment adds to existing literature of anchoring effects in legal context by adding some nuance. Although anchoring effects seem to be a robust phenomenon even in the presence of expertise and detailed case information, their effects are more nuanced than an exclusive one-way influence from prosecution to final sentences. On the contrary, both parties can make use of this effect to exert leverage on the final sentence.

Acknowledgements

Thanks to Prof. Dr. Dr. Hauke Brettel, Prof. Dr. Sven Henseler, Prof. Dr. Jörg Scheinfeld, and Prof. Dr. Tobias Reinbacher for allowing data collection in their lectures, to Marcus Traut for

20448333, 2021, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bcp.12192 by Cochrane Germany, Wiley Online Library on [17/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

generous support. Experiments 1 and 2 reported here were conducted as part of the second authors' dissertation under the first authors' co-supervision. During the final stage of preparing this manuscript, our dear colleague Birte Englich suddenly passed away († 2 September 2019) – too early and too sudden to continue the conversation we had begun about this project. Birte will be greatly missed, her enthusiasm for research that matters, her support and mentoring, her laugh.

Conflicts of interest

All authors declare no conflict of interest.

Author contributions

Roland Imhoff, Ph.D. (Conceptualization; Formal analysis; Methodology; Project administration; Supervision; Writing – original draft) Christoph Nickolaus (Conceptualization; Data curation; Project administration; Writing – review & editing).

Data Availability Statement

All materials, raw data, and analysis script for the current study as well as two pilot studies are available at https://osf.io/4um8q.

References

Bahník, Š., Houdek, P., Vrbová, L., & Hájek, J. (2019). Variations on anchoring: Sequential anchoring revisited. *Judgment and Decision Making*, *14*(6), 711–720. https://doi.org/10.31234/osf.io/skmqg

Brewer, N. T., Chapman, G. B., Schwartz, J. A., & Bergus, G. R. (2007). The influence of irrelevant anchors on the judgments and choices of doctors and patients. *Medical Decision Making*, *27*, 203–211. https://doi.org/10.1177/0272989X06298595

Chapman, G. B., & Bornstein, B. H. (1996). The more you ask for, the more you get: Anchoring in personal injury verdicts. *Applied Cognitive Psychology*, *10*, 519–540. http://doi.org/10.1002/(SICI)1099-0720(199612)10:6<519::AID-ACP417>3.0.CO;2-5.

Ebbesen, E. B., & Konečni, V. J. (1981). The process of sentencing adult felons. In B. D. Sales (Ed.), *The trial process* (pp. 413–458). Boston, MA: Springer. https://doi.org/10.1007/978-1-4684-3767-6_12

Englich, B. (2005). Geben Sie ihm doch einfach fünf Jahre!. *Zeitschrift Für Sozialpsychologie*, *36*, 215–225. https://doi.org/10.1024/0044-3514.36.4.215

Englich, B., & Mussweiler, T. (2001). Sentencing under uncertainty: Anchoring effects in the courtroom. *Journal of Applied Social Psychology*, *31*, 1535–1551. https://doi.org/10.1111/j.1559-1816.2001.tb02687.x

Englich, B., Mussweiler, T., & Strack, F. (2005). The last word in court—A hidden disadvantage for the defense. *Law and Human Behavior*, *29*, 705–722. https://doi.org/10.1007/s10979-005-8380-7

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*, 188–200. https://doi.org/10.1177/0146167205282152

Englich, B., & Soder, K. (2009). Moody experts—How mood and expertise influence judgmental anchoring. *Judgment and Decision Making*, *4*, 41–50.

20448333, 2021, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bjcp.12192 by Cochrane Germany, Wiley Online Library on [17/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Furnham, A. (1986). The robustness of the recency effect: Studies using legal evidence. *The Journal of General Psychology*, *113*, 351–357. https://doi.org/10.1080/00221309.1986.9711045

Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics*, *40*, 35–42. https://doi.org/10.1016/j.socec.2010.10.008

Galinsky, A. D., & Mussweiler, T. (2001). First offers as anchors: The role of perspective-taking and negotiator focus. *Journal of Personality and Social Psychology*, *81*, 657–669. https://doi.org/10.1037/0022-3514.81.4.657

Glöckner, A., & Englich, B. (2015). When relevance matters: Anchoring effects can be larger for relevant than for irrelevant anchors. *Social Psychology*, *46*, 4–12. https://doi.org/10.1027/1864-9335/a000214

Grau, I., & Bohner, G. (2014). Anchoring revisited: The role of the comparative question. *PLoS One*, *9*, e86056. https://doi.org/10.1371/journal.pone.0086056

Guthrie, C., Rachlinski, J. J., & Wistrich, A. J. (2000). Inside the judicial mind. *Cornell Law Review*, *86*, 777–830. https://doi.org/10.2139/ssrn.257634

Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology*, *24*, 1–55. https://doi.org/10.1016/0010-0285(92)90002-J

Jacowitz, K. E., & Kahneman, D. (1995). Measures of anchoring in estimation tasks. *Personality and Social Psychology Bulletin*, *21*, 1161–1166. https://doi.org/10.1177/01461672952111004

Kerstholt, J. H., & Jackson, J. L. (1998). Judicial decision making: Order of evidence presentation and availability of background information. *Applied Cognitive Psychology*, *12*, 445–454. https://doi.org/10.1002/(SICI)1099-0720(199810)12:5<445::AID-ACP518>3.0.CO;2-8.

Klein, R. A., Ratliff, K. A., Vianello, M., Adams, R. B., Bahník, Š., Bernstein, M. J., … Nosek, B. A. (2014). Investigating variation in replicability: A "many labs" replication project. *Social Psychology*, *45*, 142–152. https://doi.org/10.1027/1864-9335/a000178

Lund, F. (1925). The psychology of belief IV. The law of primacy in persuasion. *Journal of Abnormal and Social Psychology*, *20*, 183–191. https://doi.org/10.1037/h0066996

McElroy, T., & Dowd, K. (2007). Susceptibility to anchoring effects: How openness-to-experience influences responses to anchoring cues. *Judgment and Decision Making*, *2*(1), 48–53.

Mochon, D., & Frederick, S. (2013). Anchoring in sequential judgments. *Organizational Behavior and Human Decision Processes*, *122*(1), 69–79. https://doi.org/10.1016/j.obhdp.2013.04.002

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*, 1142–1150. https://doi.org/10.1177/01461672002611010

Neuhäuser, M. (2006). How to deal with multiple endpoints in clinical trials. *Fundamental & Clinical Pharmacology*, *20*, 515–523. https://doi.org/10.1016/0749-5978(87)90046-X

Northcraft, G. B., & Neale, M. A. (1987). Experts, amateurs, and real estate: An anchoring-and-adjustment perspective on property pricing decisions. *Organizational Behavior and Human Decision Processes*, *39*, 84–97. https://doi.org/10.1016/0749-5978(87)90046-X.

Rosnow, R., & E. Robinson (Eds.) (1967). *Experiments in persuasion*. New York, NY: Academic Press.

Schünemann, B. (1988). Daten und Hypothesen zum Rollenspiel zwischen Richter und Staatsanwalt bei der Strafzumessung [Data and hypotheses on the role play between judge and prosecutor in sentencing decisions]. In G. Kaiser, H. Kury & H. J. Albrecht (Eds.), *Kriminologische Forschung in den 80er Jahren: Projektberichte aus der Bundesrepublik Deutschland* (pp. 265–280). Freiburg, Germany: Max-Planck-Institut für ausländisches und internationales Strafrecht.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Weld, H., & Roff, M. (1938). A study in the formation of opinion based upon legal evidence. *American Journal of Psychology*, *51*, 609–628. https://doi.org/10.2307/1415696

Received 19 November 2020; revised version received 21 April 2021

20448333, 2021, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bjcp.12292 by Cochrane Germany, Wiley Online Library on [17/10/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License