Playing Dice With Criminal Sentences: The Influence of Irrelevant Anchors on Experts’ Judicial Decision Making

Birte Englich
University of Würzburg

Thomas Mussweiler
University of Cologne

Fritz Strack
University of Würzburg

Judicial sentencing decisions should be guided by facts, not by chance. The present research however demonstrates that the sentencing decisions of experienced legal professionals are influenced by irrelevant sentencing demands even if they are blatantly determined at random. Participating legal experts anchored their sentencing decisions on a given sentencing demand and assimilated toward it even if this demand came from an irrelevant source (Study 1), they were informed that this demand was randomly determined (Study 2), or they randomly determined this demand themselves by throwing dice (Study 3). Expertise and experience did not reduce this effect. This sentencing bias appears to be produced by a selective increase in the accessibility of arguments that are consistent with the random sentencing demand: The accessibility of incriminating arguments was higher if participants were confronted with a high rather than a low anchor (Study 4). Practical and theoretical implications of this research are discussed.

Keywords: anchoring; sentencing decisions; experts; irrelevant anchors; decision making; selective accessibility

“God does not play dice with the universe.”
—Albert Einstein

Human judgment is often shaped by irrelevant influences. How we judge other people for example depends on the social category to which they belong (e.g., Bodenhausen, 1990), their physical attractiveness (Nisbett & Wilson, 1977), and whether or not we encounter them on a scary bridge (Dutton & Aron, 1974). Such influences from normatively irrelevant factors appear to be a fact of our mental lives (for an overview, see Wilson & Brekke, 1994).

One of the most intriguing instances of irrelevant influences on human judgment is that of a preceding judgment. Research on the so-called anchoring effect has demonstrated that a randomly chosen standard in a comparative judgment task may dramatically influence a subsequent absolute judgment of the same target. Indicating whether the percentage of African nations in the United Nations is higher or lower than an arbitrary number (the anchor) that has been determined by spinning a wheel of fortune (showing 65% or 10%), for example, influences subsequent estimates of this percentage (Tversky & Kahneman, 1974). Specifically, absolute judgments are assimilated toward the randomly chosen “anchor values.” Such anchoring effects pervade a plethora of judgments (e.g., Epley & Gilovich, 2001; Northcraft & Neale, 1987) and are remarkably robust. Many findings indicate that clearly irrelevant numbers—even if they are blatantly determined at random—may guide numeric judgments that are generated under uncertainty (for an overview, see Chapman & Johnson, 2002; Epley, 2004; Mussweiler, Englich, & Strack, 2004; Mussweiler & Strack, 1999a).

Authors’ Note: This research was supported by the German Research Foundation (DFG). We would like to thank all participating legal experts, the German Postgraduate School of Administrative Sciences, and the Ministry of Justice of Baden-Württemberg. Thanks to Kirsten Soder, Viola Rost, Simone Dohle, and Thomas Dunlap for their indispensable practical support. Correspondence concerning this article should be sent to englich@psychologie.uni-wuerzburg.de.
ANCHORING IN COURT

This insight has important implications for many domains of human life. Specifically, these mechanisms seem to apply to decisions made by social institutions, decisions whose consequences may severely affect our lives. One example is the domain of legal decision making. Judges and juries have to decide the length of a defendant’s prison term on the basis of uncertain and partially contradictory evidence. Therefore, judicial decisions are often judgments under uncertainty (see, e.g., Diamond, 1981; Ebbesen & Konecni, 1981; Hogarth, 1971; Partridge & Eldridge, 1974). Is it possible that such important judgments with far-ranging implications could be influenced by random numbers? Specifically, would a judge impose a longer prison term after being exposed to a higher number? Would this be the case even if the number is clearly irrelevant for the sentencing decision because it is randomly determined?

At first sight, this scenario seems highly unlikely. In fact, there is reason to believe that the laboratory findings that demonstrate influences of irrelevant and random anchors do not extend to the realm of judicial decision making. This is the case because legal decisions are distinct in a number of important ways. First of all, in the legal domain, a number of rules and prescriptions exist to minimize irrelevant influences on decision making. Typically, the penal code defines a set of criteria that needs to be met before a given deed qualifies as a crime. For example, the killing of another person is only seen as murder if the defendant killed the victim intentionally. Furthermore, the penal code and sentencing guidelines specify a range within which sentences for a crime have to fall. Finally, a large body of procedural rules prescribes how evidence must be gathered, presented, and processed. Hence, legal decision making appears to be more structured than those judgments for which influences from random anchors have been demonstrated in the laboratory. In addition, legal decision makers have been thoroughly trained in their specific domain of judgment and typically have considerable experience with related cases. A criminal judge who has spent several years studying criminal law and also has several years of professional experience may thus seem much better prepared to generate a sentence in a murder case than students are prepared to estimate the percentage of African nations in the United Nations. Finally, legal decision makers are likely to be more deeply involved in the decision process and to care more about the accuracy of the decision outcome. Making a judgment about the number of years a defendant has to spend in prison is clearly more consequential than making a judgment about the percentage of African nations in the United Nations. In combination, all of these factors may work against a potential influence of random numbers on legal decisions.

At the same time however, some empirical findings suggest that even legal decisions may be open to anchoring influences under certain conditions. This seems to be the case for some specific types of anchors, namely, those that can be seen as providing relevant information about the legal decision at hand. Specifically, research in the civil context of damage awards shows clear effects of potentially relevant anchors: The higher a plaintiff’s request in court, the higher the award that is obtained (Hastie, Schkade, & Payne, 1999; Malouff & Schutte, 1989; Marti & Wissler, 2000). In personal injury verdicts, the requested compensation systematically influences the compensation awarded by the jury as well as the judged probability that the defendant caused the plaintiff’s injuries (Chapman & Bornstein, 1996).

Similar effects have been demonstrated in the criminal context. More specifically, it has been demonstrated that legal professionals who take the role of a judge in a realistic rape case are strongly influenced by the prosecutor’s sentencing demand (Englich & Mussweiler, 2001). In addition, correlational evidence stemming from analyses of actual court files is suggestive of the same data pattern: Judges heavily weigh prosecution requests in their decisions (Englich, Mussweiler, & Strack, 2005). Furthermore, actual bail decisions were found to depend on whether the prosecution requested conditional bail or opposed bail (Dhami, 2003).

In all of these cases, a numeric value that is requested or suggested in court influences legal decisions. Notably, all of these anchors can be construed as providing valuable information for the legal decision at hand. For example, the compensation a party requests is likely to correspond to the actual damage that has been made. Thus, it cannot be ruled out that the anchors provided in these studies had some informational relevance.

Taken together, this research demonstrates that potentially relevant anchors may influence legal decisions. In combination with the basic research demonstrating that even completely irrelevant anchors influence judgments (Tversky & Kahneman, 1974), this applied work raises the question whether important judicial decisions may also be influenced by completely irrelevant numbers.

THE PRESENT RESEARCH

It thus remains unclear whether the influence of random numbers is limited to unstructured judgments for which people have little expertise, information, and motivation for accuracy or whether random numbers also influence more structured decisions that have important consequences and are made by experienced experts. The present research was designed to examine
this question in the domain of legal decision making: Would the sentencing decisions of experienced legal professionals be influenced by irrelevant anchors even if the latter are determined at random? To find out, in four experiments we provided experienced legal professionals with realistic legal case materials and asked them to determine a sentence for the defendant. As is the case in actual trials, participating judges were exposed to a potential sentence (e.g., the sentencing demand of the prosecutor) before making their sentencing decision. In Study 1, the potential sentence stemmed from a source that—on normative grounds—should be irrelevant for the judge’s decision, namely, from a journalist. In Study 2, the potential sentence was even more clearly irrelevant. Here, participants were informed that the given prosecutor’s sentencing demand was determined at random. In Study 3, participants randomly determined this demand themselves by throwing a set of dice. Would their final sentences be influenced by these clearly irrelevant anchors?

In addition to examining this question, we wanted to explore the underlying psychological mechanisms of anchoring in the courtroom. In our past research, we have suggested that anchoring effects are produced by mechanisms of selective accessibility (Mussweiler et al., 2004; Mussweiler & Strack, 1999a, 1999b, 2000b; Strack & Mussweiler, 1997; for a related account, see Chapman & Johnson, 1999). More specifically, we assume that comparing the judgmental target to a provided anchor value increases the accessibility of anchor-consistent target knowledge. For example, considering a high sentencing demand as an anchor would selectively render accessible information that speaks for a high sentence (e.g., the defendant used force). Conceivably, using such easily accessible incriminating arguments as a basis for subsequent sentencing decisions leads to high sentences. From this perspective, higher sentences that follow from exposure to a high anchor are produced by a selectively increased accessibility of incriminating arguments. Past anchoring research outside the legal domain has provided substantial support for the selective accessibility notion (e.g., Mussweiler & Strack, 1999b, 2000b). Because all of this evidence was obtained with students who answered trivia questions about which they had little knowledge however, it remains unclear whether selective accessibility also plays a role in anchoring effects that are obtained in applied settings where experts have ample information about the judgment they make. In the present research, we set out to examine whether selective accessibility contributes to anchoring effects in sentencing decisions. Specifically, Study 4 exposed legal experts to randomly determined sentencing demands. Subsequently, we assessed the accessibility of incriminating and exculpatory arguments. If the selective accessibility mechanism is involved in how experts process these random sentencing anchors, incriminating arguments should be more accessible after exposure to a high anchor than after exposure to a low anchor.

In sum, the present research was designed to advance research on judgmental anchoring and research on judicial decision making in a number of important ways. First, we want to examine whether legal decision making is shaped by irrelevant influences. Second, we want to examine whether anchor values that are clearly irrelevant also influence judgments that are (a) structured by extensive norms and procedures, (b) made by experienced experts, and (c) pertain to a decision for which judges are professionally motivated to be accurate. Finally, we want to demonstrate that anchoring in information-rich settings involves a selective increase in the accessibility of anchor-consistent information.

STUDY 1

The goal of our first experiment is to examine whether a clearly irrelevant anchor influences the sentencing decisions of legal professionals. To ensure the social validity of the procedure, we focused on the potential influence of an irrelevant anchor to which judges may realistically be exposed during actual sentencing decisions.

In particular, we examined whether a sentencing anchor that is suggested by the media may influence judges’ sentencing decisions. Clearly, the media often confront judges with potential sentencing anchors that—on normative grounds—should not influence their sentencing decisions. However, to the extent that judges process such normatively irrelevant anchors—so the selective accessibility model suggests—they may still have an effect on their sentencing decisions. Study 1 was designed to examine whether this is indeed the case. Specifically, we examined whether sentencing decisions would be influenced by an inquiry from a journalist that includes a potential sentencing anchor. Legal professionals were exposed to either a high or a low potential sentence in a case of alleged rape.

Method

Participants. We recruited 42 experienced legal professionals (28 men) at educational conferences for judges and prosecutors. Of these participants, 23 were judges and 19 were prosecutors. Note that in the German system of legal education, judges and prosecutors receive identical training and alternate between both positions in the first years of professional practice. On average, the legal professionals had 129.90 (SD = 105.87) months—more than 10 years—of professional experience in the courtroom. Participants’ age ranged from 27 to 60 years, with a mean of 41.78 years (SD = 8.86). Men
and women were equally distributed across experimental conditions.

**Materials.** Participants received realistic case material about an alleged rape. The materials were designed to include all the relevant information that is typically provided in actual court cases. To ensure that the material seemed realistic, it was designed in close collaboration with experienced trial judges. These judges worked through the material and supplemented it with information they believed was necessary to determine a sentence. This material was pretested in previous studies (Englich & Mussweiler, 2001), where it was judged to be complete and realistic by the participating legal experts.

The case material covered about four pages and consisted of brief descriptions of the incidence, the victim (“Sabine K.”), and the defendant (“Peter F.”). Furthermore, advisory opinions from a medico-legal and a psycho-legal expert and statements by the victim, the defendant, and two witnesses were provided. Participants took about 15 minutes to work through this material. The material included all the information (e.g., psychological consequences for the victim, resistance of the victim, threats by the assailant) that previous research has demonstrated to be important for an ascription of guilt in cases of rape (Krahé, 1991). For example, the consumption of alcohol by the victim and perpetrator was described as moderate (Schuller & Stewart, 2000), and further details—like the fact that the perpetrator used a condom (Hynie, Schuller, & Couperthwaite, 2003)—were revealed. In addition, participants had the relevant passages from the penal code at their disposal.

**Procedure.** Legal professionals participated in groups of up to 15. Participants were first handed the case materials, asked to work through them, and to put themselves in the role of the criminal judge in this specific case. Subsequently, they received the crucial questionnaire while keeping all the materials. In this questionnaire, participants were first instructed to imagine the following situation: During a court recess they receive a telephone call from a journalist who directly asks them, “Do you think that the sentence for the defendant in this case will be higher or lower than 1/3 year(s)?” (low/high anchor). About half of the participants were exposed to the high anchor, the other half to the low anchor. Participants were further instructed to imagine that to remain unbiased, they refuse to answer this question and instead bring the telephone call to a quick end. At a subsequent coffee break however, they meet a colleague with whom they start talking about the case. In the course of this conversation, they tell their colleague about the journalist’s call. Embedded in this scenario participants were asked, “What point of view would you represent to your colleague: Do you think that the sentence suggested by the journalist was too high, too low, or just right?” Participants indicated whether this potential sentence was too high, too low, or just right. Congruent with the standard anchoring procedure (Strack & Mussweiler, 1997; Tversky & Kahneman, 1974), this comparative anchoring question was included to ensure that participants do indeed process and consider the given anchor value. Subsequently, participants were asked to give their own sentencing decision. The sentencing decision was followed by a question about how certain the participant felt about the decision (1 = not at all certain, 9 = very certain) and by a rating of the quality of the presented case material. Specifically, participants indicated how realistic the provided case material was (1 = not at all realistic, 9 = absolutely realistic). Finally, participants provided some demographic details.

**Results and Discussion**

**Preliminary analyses.** As in our previous research in which we used similar case materials (Englich & Mussweiler, 2001), participants judged the materials to be realistic, \( M = 7.38, SD = 1.40 \). This judgment did not depend on the anchoring condition, \( t < 1 \). The 19 prosecutors and 23 judges in our sample did not differ in the extent to which they judged the case material as realistic, \( t < 1 \). More important, the prosecutors and judges did not differ in the overall length of their sentences, \( t < 1 \).

Furthermore, prosecutors and judges were similarly susceptible to anchoring influences in their sentencing decisions: There is no interaction between the anchoring condition and participants’ legal profession, \( F < 1 \). Therefore, responses were collapsed across prosecutors and judges.

**Sentencing decisions.** Legal professionals’ sentencing decisions for the identical legal case ranged from acquittal to 5 years in prison. An analysis of the mean sentencing decisions indicated that judges were clearly influenced by the potential sentence suggested by the journalist. Participants who had been exposed to the high sentencing anchor gave considerably higher sentences, \( M = 33.38 \text{ months}, SD = 9.65 \), than participants who were confronted with a low anchor, \( M = 25.43, SD = 10.49, t(40) = 2.56, p < .02 \). Furthermore, participants felt fairly certain about their sentencing decision, \( M = 6.02, SD = 2.04 \), and certainty did not depend on the anchoring condition, \( t < 1 \).

**Comparing the effects of relevant and irrelevant anchors.** These findings indicate that sentencing decisions are influenced by clearly irrelevant sentencing anchors. A remaining open question is whether the magnitude of this influence differs from that of relevant sentencing anchors. To find out, we compared sentencing decisions in the present study with those obtained in one of our
previous studies (Englich & Mussweiler, 2001, Study 2) in which legal professionals were confronted with the same materials and anchor values with the exception that these anchors represented the prosecutor’s sentencing demand. A combined analysis of both studies revealed that sentencing decisions in both anchor conditions differed no matter whether this anchor was suggested by a journalist (M = 25.43 vs. M = 33.38) or the prosecutor (M = 19.09 vs. M = 25.91). In a 2 (anchor: high vs. low) × 2 (source: journalist vs. prosecutor) ANOVA, only the main effects for anchor, F(1, 60) = 9.38, p < .01, and source, F(1, 60) = 8.19, p < .01, reach significance, F < 1 for the interaction. This indicates that the magnitude of the resulting anchoring effect is independent of anchor relevance.

Taken together, these results demonstrate that the sentencing decisions of experienced legal professionals may indeed be influenced by clearly irrelevant sentencing anchors. Although on normative grounds a potential sentence that is suggested by a journalist should not influence a final sentence in court, participants gave substantially higher sentences if they were exposed to a high rather than a low sentencing anchor. In fact, final sentences differed by about 8 months. Identical crimes were thus punished with strikingly divergent prison sentences depending on the sentencing anchor to which judges were exposed by the journalist.

In Study 2, we explored potential boundaries of this influence of irrelevant anchors. Would sentencing decisions by legal professionals be influenced even by anchors that were supposedly determined at random?

STUDY 2

Method

Participants. We recruited 39 legal professionals (18 women) during an educational conference; 37 were judges and 2 were prosecutors. Careful inspection of the 2 prosecutors’ data in our sample reveals that neither their sentences nor their evaluations of the case materials differed from those of the judges in our sample. Therefore, we collapsed data for judges and prosecutors. On average, participants had 13.38 years (SD = 8.58) of professional experience in court. Participants’ age ranged from 29 to 61 with a mean of 42.59 years (SD = 9.44).

Materials and procedure. As in Study 1, participants were asked to put themselves in the role of a trial judge. This time, they had to find a sentence in a fictitious shoplifting case concerning a woman who had stolen some items from a supermarket for the 12th time. Again, the case material was compiled in close collaboration with legal professionals and consisted of brief descriptions of the incidence and the defendant (“Lena M.”), an advisory opinion from a psycho-legal expert, and statements by the defendant and a witness. An independent pretest using a different group of experienced legal professionals as participants (N = 13) demonstrated that these materials were judged to be complete (M = 6.77, SD = 1.48) on a 9-point rating scale with 1 = not at all complete and 9 = absolutely complete) and realistic (M = 7.39, SD = 1.33) on a 9-point rating scale with 1 = not at all realistic and 9 = absolutely realistic). Pretest participants were also asked to put themselves in the position of the trial judge in this case and to report their sentencing decision. Their mean sentence was M = 5.62 months (SD = 2.57).

After reading the case materials and the corresponding passages from the penal code, participants were handed the critical questionnaire in which they were asked to report their sentencing decision. As is the case in actual sentencing decisions in court, after working through the case material participants were asked to consider the sentencing demands of the prosecutor and the defense attorney before reporting their final decision. Participants were first confronted with a prosecutor’s sentencing demand that was either high (9 months on probation) or low (3 months on probation). Instructions clearly pointed out that this demand had been determined at random, thus it did not represent any judicial expertise. Specifically, the instructions read,

For experimental purposes, the following prosecutor’s sentencing demand was randomly determined, therefore, it does not reflect any judicial expertise: The prosecutor demands as a sentence for the accused Lena M. 3/9 months on probation. Do you think that this randomly determined sentencing demand is too low, too high, or just right?

Participants indicated whether they considered the randomly determined prosecutor’s demand to be too low, too high, or just right. Subsequently, participants were confronted with the defense attorney’s demand, which was always 1 month on probation, and again indicated whether they considered this demand to be too low, too high, or just right. Finally, participants reported their sentencing decision, indicated how certain they felt about their judgment (1 = not at all certain, 9 = very certain), answered the same question about the materials used in Study 1, and provided some demographic data.

Results and Discussion

Again, the case material was judged to be realistic, M = 6.74, SD = 1.94. Judges’ sentencing decisions for the given shoplifting case varied between acquittal and 12 months on probation, with a mean sentence of 5.05 months (SD = 3.18). Further inspection of the given sentences reveals that they clearly depended on the prosecutor’s sentencing demand, which was obviously deter-
mined at random. Specifically, judges who were exposed to the high demand gave higher sentences, $M = 6.05$ months, $SD = 3.07$, than judges who were exposed to the low demand, $M = 4.00$ months, $SD = 3.02$, $t(37) = 2.10, p < .05$. Again, participating legal professionals felt fairly certain about their sentencing decision, $M = 5.28, SD = 2.62$, and this certainty was independent of the anchoring condition, $t < 1.1$.

Comparing the effects of relevant and irrelevant anchors. To examine whether the magnitude of the obtained effect depends on the relevance of the given anchor, we compared the present findings with those of another study in which legal professionals ($N = 80$) were confronted with the same materials and anchor values with the exception that these anchors were relevant because they were suggested by the prosecutor (Englich, 2006). A combined analysis of both studies revealed that sentencing decisions in both anchor conditions differed no matter whether this anchor was blatantly selected at random ($M = 4.00$ vs. $M = 6.05$) or suggested by the prosecutor ($M = 4.10$ vs. $M = 6.98$). In a $2$ (anchor: high vs. low) $\times 2$ (source: random vs. prosecutor) ANOVA, only the main effect for anchor, $F(1, 115) = 26.82, p < .001$, reaches significance, $F(1, 115) = 1.16, p < .3$, for the main effect of source, $F < 1$ for the interaction. Thus, the magnitude of the resulting anchoring effect did not depend on anchor relevance.

These findings demonstrate that sentencing decisions of legal professionals may be influenced by a sentencing anchor even if this anchor is obviously determined at random.

The Influence of Expertise and Experience in Studies 1 and 2

In principle, one may expect that such random influences would be evident primarily in sentencing decisions by nonexpert judges. To the extent that professional expertise provides judges with alternative information that compensates for anchoring influences, experts may well show less bias from having processed a random anchor. The anchoring literature however suggests otherwise. Research on anchoring effects in the legal domain (e.g., Englich & Mussweiler, 2001) and beyond (e.g., Mussweiler, Strack, & Pfeiffer, 2000; Northcraft & Neale, 1987) has demonstrated that the judgments of experts and nonexperts could be construed as relevant. Thus, it remains unclear whether experts may also be influenced by patently irrelevant anchors. A combined analysis of Studies 1 and 2 allows us to examine this question and to investigate whether the magnitude of anchoring differs for experts and nonexperts.

All participants in Studies 1 and 2 were experienced legal professionals, but they differed with respect to the specific fields of law in which their primary expertise lay. In Study 1, 24 participants were specialized in criminal law, whereas 18 were experts in other fields, such as civil law, administrative law, social law, and so on. Similarly, in Study 2, 17 participants were experts in criminal law and 22 were experts in other fields. The composition of our participant population thus allows us to differentiate between legal professionals who have specific expertise and long professional experience in the specific domain to which our legal cases pertained and those who had little specific expertise and professional experience in this domain.

In combination, Studies 1 and 2 included 41 experts in criminal law and 40 nonexperts. Although the sample sizes are too small to examine how expertise influenced sentencing decisions separately for the two studies, such an analysis is possible if the two studies are combined. To compare sentences across the different cases, we $z$-transformed the sentencing decisions.

As an examination of Figure 1 reveals, the sentencing decisions of experts and nonexperts in criminal law depended on the irrelevant sentencing anchors to similar degrees. In fact, a $2$ (expert vs. nonexpert) $\times 2$ (high vs. low anchor) ANOVA using the $z$-transformed sentencing decisions as the dependent variable only found a significant main effect of anchor, $F(1, 77) = 10.90, p < .001$, but no main effect of expertise and no interaction, all $F < 1$. Furthermore, simple effect analyses demonstrated that a significant anchoring effect occurs for experts, $t(39) = 2.45, p < .02$, and nonexperts, $t(38) = 2.23, p < .04$, alike. The only notable difference between the experts and nonexperts is that the experts felt more certain about their sentencing decision, $M = 6.88, SD = 1.60$, than the nonexperts, $M = 4.45, SD = 2.37, t(68.47) = 5.35, p < .001$. The certainty experienced by the judges
however was unrelated to their susceptibility to the anchoring bias: In an additional analysis, the degree of bias is indicated by the distance between the prosecutor’s initial demand and the judge’s sentence and thus reflects the extent to which judges’ sentences were assimilated toward the initial anchor. In fact, certainty and bias were uncorrelated, r = .08, p > .5. In combination with the higher certainty ratings of experts, this suggests that experts may mistakenly see themselves as less susceptible to biasing influences on their sentencing decisions.

STUDY 3

The results of Study 2 demonstrate that sentencing anchors that were blatantly determined at random influenced the sentencing decisions of legal professionals. There may however still remain some doubt whether the random nature of these anchors was fully accepted by our participants. To make absolutely sure that participants were fully aware of the fact that anchor values were randomly determined, we put randomization into the hands of our participants. In Study 3, participants randomly determined the sentencing anchors themselves by throwing a pair of dice. Would such anchors still influence sentencing decisions by legal experts?

Method

Participants. We recruited 52 legal experts (28 men) from a supplemental national postgraduate training program at the German University of Administrative Sciences in Speyer. Participants were junior lawyers from different German courts who had recently received their law degree and had acquired their first experiences as judges in court. Their ages ranged from 24 to 33 years with a mean of 27.5, SD = 1.79.

Materials and procedure. With the exception of the randomization procedure, materials and procedures were identical to those used in Study 2. This time, after working on the experimental materials, participants were told to randomly determine the prosecutor’s sentencing demand themselves by throwing a pair of dice. The dice were loaded so that participants in Study 3 were confronted with exactly the same sentencing demands as participants in Study 2 (3 vs. 9 months on probation). Participants were informed that the experiment was a pretest for a study on optimal questioning sequences in the courtroom, hence they should strictly follow the order of questions in the questionnaire. In addition, it was explained that the prosecutor’s demand was determined at random to ensure that it did not influence participants’ answers to the subsequent questions.

About half of the participants were handed a pair of dice that was loaded so that the dice always indicated the numbers 3 and 6. After the dice had been thrown, participants were instructed to calculate the sum of the two dice and to fill in this sum as the prosecutor’s sentencing demand in the questionnaire. Participants then worked on the sentencing questionnaire, which consisted of the same questions that were used in Study 2. This time, we did not ask for additional ratings of the case materials because these materials had been extensively pretested in our previous research and in Study 2.

Results and Discussion

As in the previous studies, sentencing decisions for the identical crimes varied substantially, ranging from 1 month on probation to 12 months on probation with a mean sentence of 6.6 months, SD = 3.57. Overall, the sentencing decisions of the junior lawyers in Study 3 are thus similar to those of their more experienced colleagues who participated in Study 2.

Further analysis again revealed that sentencing decisions were influenced by the random sentencing anchors. Judges who were exposed to the high anchor gave higher final sentences, M = 7.81, SD = 3.51, than those who were confronted with a low anchor, M = 5.28, SD = 3.21, t(50) = 2.71, p < .01. Participants felt moderately certain about their sentencing decisions, M = 5.87, SD = 1.86, and perceived certainty was independent of the anchoring condition, t < 1.

Comparing the effects of relevant and irrelevant anchors. We again compared the present findings with those of the other study with identical materials but relevant anchors (Englich, 2006). Again, sentencing decisions in both anchor conditions differed for anchors determined at random (M = 5.28 vs. M = 7.81) and suggested by the prosecutor (M = 4.10 vs. M = 6.98). In a 2 (anchor: high vs. low) × 2 (source: random vs. prosecutor) ANOVA, only the main effects for anchor, F(1, 128) = 32.56, p < .001, and source, F(1, 128) = 4.54, p < .04, reached significance, F < 1 for the interaction. No matter whether the anchor was determined by throwing dice or suggested by the prosecutor did judges assimilate their sentencing decisions to it to a similar degree.

Taken together, these findings demonstrate that even if legal experts randomly determined a sentencing anchor themselves by throwing a pair of dice, they were influenced by it. Judges assimilated their sentencing decisions toward these clearly irrelevant sentencing demands.

Comparing the Effects of High and Low Anchors to an Unanchored Control Group

Anchoring research typically compares the effects of high and low anchors without including an unanchored control group. This is done for at least three reasons.
First, eliminating the preceding judgment in which judges consider the anchor makes it unclear whether potential differences in judgment are attributable to the absence of an anchor or to the absence of the preceding judgment task. Second, the fact that no anchor is provided in the experimental materials does not ensure that no anchor is used. It has been demonstrated that any number that is sufficiently accessible can serve as an anchor for numerical judgments (Mussweiler & Englich, 2005; Wilson, Houston, Etling, & Brekke, 1996). Judges who are not provided with an anchor value are thus likely to bring their own anchor into the experiment so that a no anchor control group is merely a self-anchored control group. Finally, most anchoring research is primarily interested in influences on the absolute magnitude of anchoring. Whether the low or the high anchor has a stronger effect is usually of minor importance. For all of these reasons, we have also followed the typical procedure of comparing judgments in a low and a high anchor condition in the present research.

At the same time, our pretesting data allow us to examine whether both anchors influence judges’ sentencing decisions in Studies 2 and 3. To do so, we combined the data from both studies and examined whether the mean sentence that was given in the high versus low anchoring condition across both studies differed from the norm sentence that was given in our pretest. This analysis revealed that the combined mean sentence in the low anchor condition \((M = 4.73, SD = 3.16)\) differed from the sentence given in the high versus low anchoring condition across both studies differed from the norm sentence that was given in our pretest. This analysis revealed that the combined mean sentence in the low anchor condition \((M = 4.73, SD = 3.16)\) differed from the sentence given in the high anchor condition \((M = 5.62, SD = 2.57)\), \(t(43) = 1.88, p < .07\). The combined mean estimate in the high anchor condition \((M = 7.06, SD = 3.41)\) also differed from the unanchored sentence, \(t(46) = 2.9, p < .01\). This supplemental analysis thus suggests that both anchors influenced sentencing decisions.

In combination, Studies 1 through 3 demonstrate that irrelevant and random numbers have powerful and robust effects on the sentencing decisions of legal professionals. In our final study, we attempt to shed light on the psychological mechanisms that contribute to this sentencing bias.

**STUDY 4**

Our previous research suggests that anchoring effects in other judgmental domains are produced by mechanisms of selective accessibility. Specifically, considering an anchor value selectively increases the accessibility of knowledge indicating that the target quantity may be similar to this anchor. Considering a high sentencing demand for example may selectively render those arguments accessible that imply a high sentence (e.g., the defendant used force). Study 4 was designed to examine whether such a selective accessibility effect would also be apparent for anchoring in sentencing decisions.

To do so, we developed a novel task as a measure of the expected selective accessibility consequences of judgmental anchoring. Specifically, we asked participants to categorize a series of arguments that were relevant for the present case as either incriminating or exculpatory as fast as possible.

**Method**

**Participants.** We recruited 57 legal experts (30 women) from a postgraduate training program for junior lawyers who had recently received their law degree and had acquired their first experiences as judges in court. Age ranged from 24 to 36, \(M = 27.32\) years, \(SD = 2.17\).

**Procedures and materials.** The case materials were identical to those used in Studies 2 and 3, and the procedures were largely similar to those of Study 3. Importantly, participants again determined the prosecutor’s sentencing demand themselves by throwing a pair of dice. After working through all the materials and after indicating whether the prosecutor’s and the defense attorney’s demands were too high, too low, or just right, participants did not however report a sentencing decision. Instead, they worked on a categorization task that was designed to measure the accessibility of incriminating and exculpatory arguments.

In this categorization task, participants were exposed to a series of brief statements denoting incriminating and exculpatory arguments for the shoplifting case. For each of these statements, they had to indicate as fast as possible whether it corresponded to an incriminating or an exculpatory argument in the context of the present case. Participants thus categorized the given statements as incriminating or exculpatory by pressing either the right or the left Ctrl key on the computer board. The material consisted of 7 incriminating and 7 exculpatory arguments. These 14 arguments were pretested with a group of legal professionals (\(N = 48\)) who rated the relevance of each argument for the specific case. A selective accessibility effect should be most apparent for relevant arguments. Therefore, we selected the 4 incriminating arguments and 4 exculpatory arguments with the highest mean relevance ratings as our critical stimuli for the categorization task. The 4 critical incriminating arguments were previous convictions, violation of probation, persistent offender, and rapid subsequent offenses. The 4 critical exculpatory arguments were diminished responsibility, insignificance, kleptomania, and willingness to undergo therapy. The remaining 6 arguments that were of low relevance for the given case were used as filler items.
In the instructions to this task, participants were first informed about the ostensible purpose of the task, namely, to assess their current ability to concentrate by measuring their response latencies to material related to the case they had just worked on. Specifically, participants would be exposed to a series of brief statements that they should categorize as incriminating or exculpatory as fast and as accurately as possible. To that end, they should keep their left and right index fingers on the corresponding response keys throughout the task and should focus on a fixation point that was presented in the center of the computer screen. The fixation point was presented for 1,500 milliseconds and was overwritten by the target statement, which remained on the screen until participants had made their categorization decision. This sequence was repeated for all 14 statements. Presentation order was determined at random.

After completion of the categorization task, participants provided some demographic information, were debriefed and thanked for their participation. In the context of Study 4, we thus did not ask participants to provide sentencing decisions. The reason for this is that the categorization task we used to assess the accessibility of incriminating and exculpatory arguments simultaneously manipulated the accessibility of these arguments. Specifically, by presenting participants with anchor-consistent and anchor-inconsistent arguments, the categorization task provided them with judgment-relevant information that is likely to directly influence subsequent judgments. In fact, previous research has demonstrated that manipulations of knowledge accessibility that are independent of the anchor manipulation influence target judgments (e.g., Chapman & Johnson, 1979; Mussweiler et al., 2000). In light of these findings, anchoring effects that were obtained subsequent to the categorization task would be difficult to interpret. Because of these ambiguities and because Studies 1 through 3 clearly demonstrated the judgmental influences that irrelevant anchors have in the judicial context, we focused exclusively on the selective accessibility consequences of irrelevant anchors in Study 4.

Results and Discussion

As suggested by Fazio (1990), we excluded response latencies that deviated by more than 3 standard deviations from the argument mean as outliers. The means given in Figure 2 provide substantial support for our hypothesis. As expected, incriminating arguments were categorized faster by participants who were exposed to a high rather than to a low sentencing demand, \( t(55) = 2.03, p < .05 \). In contrast, response latencies to exculpatory arguments did not depend on the anchoring condition, \( t(55) < 1 \). In a 2 (exculpatory vs. incriminating argument) × 2 (high vs. low anchor) mixed-model ANOVA with argument valence as a within factor and anchor as a between factor, this pattern produced a significant interaction effect, \( F(1, 55) = 5.23, p < .03 \). In this analysis, none of the main effects reached significance, with \( F < 1.5, p > .25 \) for the main effect of the anchor on response times, and \( F < 2.2, p > .1 \) for the main effect of the argument valence showing slightly shorter reaction times for incriminating arguments.

These findings indicate that processing a random sentencing anchor leads to a selective increase in the accessibility of anchor-consistent arguments. Participants who were exposed to the high anchor were subsequently able to categorize incriminating arguments faster than participants who were exposed to a low sentencing anchor. Notably, categorizations of exculpatory arguments did not depend on the anchoring condition. Although this lack of an effect for exculpatory arguments is surprising at first sight, it is understandable in the light of research demonstrating that information with negative valence often carries more weight than information with positive valence (see Rozin & Royzman, 2001). In general, negative information receives more attention and seems to obtain priority in processing. This is apparent for example in the fact that negative words are detected more easily than positive words (Dijkstra & Aarts, 2003). The priority of negative information is also a potent influence on person judgment. For example, judgments about the moral qualities of a person are more strongly influenced by negative than positive deeds (Reeder & Brewer, 1979). In light of this research, it makes sense that our legal experts focused primarily on the incriminating arguments when processing the anchor value. This tendency is likely to be further strengthened by the fact that in the legal domain, judges’ task is to determine whether the defendant is guilty beyond reasonable
doubt. Because the defendant’s guilt is primarily determined by the incriminating arguments, judges may focus more on incriminating than on exculpatory arguments. The extent to which judges focus on incriminating arguments however further depends on how compatible these arguments are with judges’ processing objectives. Just as negative information is particularly attention grabbing if it is compatible with judges’ processing objectives (Neumann & Strack, 2000), incriminating arguments will receive more attention if they are compatible with the hypothesis judges test while processing a given anchor. Incriminating arguments thus receive the most attention and are consequently most accessible if a high sentencing anchor was processed. As a consequence, the selective accessibility effect becomes apparent for these incriminating arguments.

**GENERAL DISCUSSION**

Judicial sentencing decisions should be guided by facts, not by chance. On normative grounds, the sentences that criminal judges impose should be immune to random influences. In the present research we have examined whether one particular random influence, namely, exposure to random numbers, has an effect on legal decisions. More specifically, we have investigated whether irrelevant anchor values that were obviously determined at random may influence sentencing decisions of legal professionals. Our results demonstrate that this is indeed the case. Using two different sets of case materials and different populations of legal experts as participants, our results demonstrate that blatantly irrelevant sentencing anchors influenced the sentencing decisions of legal experts. In Study 1, legal professionals were influenced by potential sentences that were suggested by a journalist’s question during a telephone call—a source that clearly should be irrelevant on normative grounds. In the remaining three studies, the prosecutor’s sentencing demands were even more clearly irrelevant. Specifically, in Study 2 participants were explicitly told that the provided sentencing demand was determined at random. Still, these potential sentences served as judgmental anchors and influenced the subsequent sentencing decisions of experienced legal professionals. Our final two studies went even further to ensure that sentencing demands were clearly irrelevant. Using a set of loaded dice, in Studies 3 and 4 our participants randomly determined the sentencing demands of the prosecutor themselves. Even though this procedure ensured that our participants were aware of the irrelevance of the sentencing demands, their sentencing decisions were dramatically influenced by them.

Notably, this influence of irrelevant anchors on sentencing decisions did not depend on judges’ experience and expertise. Our analyses indicate that legal professionals who were experts in criminal law and had considerable experience in similar legal cases were influenced by irrelevant sentencing demands in much the same way as legal professionals who were experts in other aspects of law and had no actual experience with similar cases.

Furthermore, our analyses reveal that irrelevant sentencing anchors produce effects that are comparable in magnitude to those of relevant sentencing anchors. Thus, an anchor that participants randomly determined themselves by throwing a set of dice influences their sentencing decisions to a similar extent as the sentencing demand of a prosecutor.

Finally, the results of Study 4 suggest that these random influences on sentencing decisions involve mechanisms of selective accessibility. Considering a high irrelevant sentencing demand selectively makes incriminating arguments accessible. Because the final sentencing decision is then strongly influenced by those arguments that come to mind easily, this ultimately leads to higher sentencing decisions.

This research has a number of important implications for research on legal decision making as well as for research on judgmental anchoring. Research on legal decision making has repeatedly demonstrated that sentencing decisions are influenced by factors that—on normative grounds—are irrelevant and should thus not have any effect (e.g., Blair, Judd, & Chapleau, 2004; Lieberman, 2002). The present research demonstrates that sentencing decisions are even open to completely random influences. Random numbers may serve as anchors to which sentencing decisions are assimilated.

Previous research has shown that anchoring effects constitute a strong influence on legal decisions in general (e.g., Chapman & Bornstein, 1996; Hastie et al., 1999) and on criminal sentencing decisions in particular (Englich & Mussweiler, 2001). In this previous work however, the anchors that were given could be seen as providing relevant information about the judgment to be made. The present studies extend this work in at least two ways. First, they demonstrate that anchors that are clearly irrelevant because they were randomly determined still influence legal decisions. Second, they show that experienced experts in the specific sentencing domain are as susceptible to this influence as nonexperts. Thus, it is not only legal laypeople serving as jury members (e.g., Chapman & Bornstein, 1996) who are influenced by a given anchor. Experienced criminal judges who have worked on many related cases and have made many related sentencing decisions were still influenced by a sentencing demand that was determined by throwing a set of dice.
What are the implications of these findings for real-world court cases? Even though judges typically do not throw dice before making sentencing decisions, they are still constantly exposed to potential sentences and anchors during sentencing decisions. The mass media, visitors to the court hearings, the private opinion of the judge’s partner, family, or neighbors are all possible sources of sentencing demands that should not influence a given sentencing decision. As the results of Study 1 demonstrate however, such authentic but normatively irrelevant anchors influence sentencing decisions as well as random anchors. Furthermore, random numbers that have been made accessible in an unrelated context may also be used as anchors for sentencing decisions. It has been demonstrated that numbers that were made accessible by extensive use in a prior task (Wilson et al., 1996) may influence judgments in much the same way as anchors that arise in the judgmental context itself. This suggests that sentencing decisions may also be influenced by irrelevant anchors that simply happen to be uppermost in a judge’s mind when making a sentencing decision. The fact that random numbers may influence sentencing decisions—as our research demonstrates—suggests that irrelevant influences on sentencing decisions may be a widespread phenomenon.

In addition to these implications for legal decision making, the present studies also provide a number of important novel insights into the anchoring phenomenon. First, they demonstrate that the influence of random anchors is not limited to the psychological laboratory. To date, the influence of random anchors has only been demonstrated with judgments for which participants had little background knowledge, little experience, and little motivation to provide an accurate answer (Mussweiler & Strack, 2000a; Tversky & Kahneman, 1974). The judgments we examined in the present studies markedly differ from these earlier demonstrations. Specifically, the sentencing decisions we have focused on are more clearly structured in that decision criteria, procedural norms, and sentencing ranges are prescribed by the law. Furthermore, as legal professionals, our participants had received extensive training in the critical judgment domain, had considerable experience in making similar sentencing decisions, and were motivated to provide an accurate judgment. Still, they were influenced by random numbers even if they determined these numbers themselves by throwing dice.

Second, the present findings extend previous work examining anchoring influences on expert judgment in important ways. These previous studies (e.g., Joyce & Biddle, 1981; Mussweiler et al., 2000; Northcraft & Neale, 1987; Wright & Anderson, 1989) have all examined the influence of anchors that may be seen as providing relevant information, such as the listing price of a house (Northcraft & Neale, 1987), a number provided by the experimenter (Joyce & Biddle, 1981), or a suggested selling price for a car (Mussweiler et al., 2000). As a consequence, these demonstrations of anchoring in expert judgment may simply be the result of experts’ ability to make efficient use of relevant information. To our knowledge, evidence demonstrating that experts are influenced by truly irrelevant anchors has not been reported to date. In fact, one may well argue that experts are particularly likely to remain uninfluenced by irrelevant anchors. After all, experts have ample knowledge about the target domain that they could use to retrieve or construct a more relevant anchor. An experienced legal expert who has to determine a sentence in a case of rape for example may easily think back to similar cases and use the given sentences as anchors for the present sentencing decision. In this situation, a randomly determined number that is clearly irrelevant seems unlikely to have an effect. The present research however demonstrates that despite their experience and knowledge, expert judges are influenced by randomly determined anchors. In this respect, the present studies are the first to demonstrate that expert judgments are influenced by clearly irrelevant anchors. In addition, the present findings demonstrate that whereas experts are as susceptible to anchoring influences as novices, they feel more certain about their judgments. Expertise thus does make a difference. Ironically however, this difference is only apparent in the subjective not the objective quality of the judgment.

Furthermore, this research allows us to directly compare the magnitude of anchoring effects that are produced by relevant and irrelevant anchors. Previous research has demonstrated that anchors that are clearly irrelevant because they were selected at random influence judgments (e.g., Cervone & Peake, 1986; Mussweiler & Strack, 2000a; Tversky & Kahneman, 1974). The present studies demonstrate that such clearly irrelevant anchors produce anchoring effects of similar magnitude as clearly relevant anchors.

Finally, the present research provides important insights into the psychological mechanisms that underlie anchoring. We have suggested that anchoring effects are produced by a selective increase in the accessibility of anchor-consistent information about the judgmental target. Our previous research has supported this notion (for an overview, see Mussweiler & Strack, 2001). The results of Study 4 supplement this former work in two important ways. First, they suggest that selective accessibility also plays a role in anchoring effects in information-rich contexts where judges have ample knowledge about the judgmental target. One may well expect that exposure to an anchor value may only change the accessibility of target knowledge if judges...
rarely activate such knowledge. Knowledge that is at the core of judges’ expertise however and that is consequently used on an almost daily basis may be so chronically accessible that it is difficult to further increase its accessibility by exposure to an anchor. The present findings however demonstrate that selective accessibility effects are also obtained for experienced judges who have ample background knowledge about the judgment to be made. This further emphasizes the ubiquity of the selective accessibility mechanism. Second, the results of Study 4 demonstrate that selective accessibility effects also result from exposure to randomly determined anchors. In this respect, the present findings suggest that selective accessibility may contribute to the effects of relevant and irrelevant anchors in a variety of judgmental settings.

Within and beyond the legal domain, irrelevant anchors may stem from different sources. They may be explicitly provided, subtly suggested, self-generated, simply coming to mind, or determined by throwing dice. As the present findings suggest, sentencing decisions may be influenced in all of these cases. God may not play dice with the universe—as Albert Einstein reassured us. But judges may unintentionally play dice with criminal sentences.

NOTES

1. Congruent with the typical methodological approach to the study of anchoring effects, our analyses focused on judges’ sentencing decisions and mostly ignored responses to the comparative anchoring question. Recent research has demonstrated that oftentimes anchoring effects occur no matter whether a comparative question is or is not asked. For example, subliminal presentation of an anchor value yields similar effects as including this value in a comparative question (Mussweiler & Englich, 2005). A supplemental analysis of our data indicates that answers to the comparative question and sentencing decisions were correlated, $r = .35$, $p < .03$. The more judges’ answers to the comparative question imply a high sentence, the higher their actual sentencing decision.

2. Answers to the comparative anchoring question and sentencing decisions were again correlated, $r = .5$, $p < .001$.

3. As in Studies 1 and 2, answers to the comparative anchoring question and sentencing decisions were correlated, $r = .44$, $p < .001$.

REFERENCES


Received December 16, 2004
Revision accepted June 5, 2005