The Influence of Irrelevant Anchors on the Judgments and Choices of Doctors and Patients

Noel T. Brewer, PhD, Gretchen B. Chapman, PhD,
Janet A. Schwartz, PhD, George R. Bergus, MD

Background. *Several decades of research have shown that judgments typically and robustly assimilate toward irrelevant anchor numbers. However, little or no research has examined how anchor numbers affect choice.* **Methods**. *In one experiment, HIV-positive patients (*N *= 99) judged the chance that a sexual partner would become HIV infected after sex with a defective condom and then indicated their treatment choices. In a second experiment, Iowa physicians (*N *= 191) rated the chance that a hypothetical patient had a pulmonary embolism and then formulated a treatment plan.* **Results**. *The 2 experiments showed support for a scale distortion hypothesis that predicts that choices following an anchored judgment would actually contrast away from the anchor. Both experiments showed the expected assimilation effect in judgment, but both also showed the predicted contrast effect in choice.* **Discussion**. *Biases that have been demonstrated in judgment should be demonstrated in choice rather that assuming such an effect exists. The practical implications of the anchoring bias are noteworthy for risk judgments but of less concern for treatment choices. The findings also show that the theoretical underpinnings of the anchoring bias are more complex than previously thought.* **Key words:** *anchoring bias; assimilation effect; contrast effect; medical decision making; risk perception.* ***(Med Decis Making 2007;27: 1–9)***

When Tversky and Kahneman (1974) introduced heuristics—anchoring, representativeness, and availability—to the world, they were conceived as simple and powerful strategies that made judgments easier but led to predictable biases. We concern ourselves in this article with the question of whether one of these biases, anchoring, also extends to choices. Irrelevant anchor numbers bias judgments, including personal estimates of health risk. The effect is robust[1] and extremely difficult to debias.[2] Yet, we do not understand what effects anchors might exert on choices, including decisions to treat patients or to pursue medical care. Similarly, powerful anchoring effects in choice could give us cause for concern and they could also offer a vital tool for improving health-related behaviors. Surprisingly, there is at present no ready answer to the question, Should we expect a similarly robust anchoring effect in choice? We present 2 experiments here that address this question. No studies that we are aware of have clearly tested the effect of irrelevant anchor numbers on choice.* Before describing the studies, we first turn to a discussion of theoretical accounts for anchoring that will allow us to make predictions.

Experiments demonstrating the anchoring bias commonly use the standard anchoring design first

Received DATE from the Department of Health Behavior and Health Education, School of Public Health, University of North Carolina, Chapel Hill (NTB); Department of Psychology, Rutgers University, Camden, NJ (GBC); Woodrow Wilson School of International and Public Affairs, Princeton University, Princeton, NJ (JAS); and the Departments of Family Medicine and Psychiatry, College of Medicine, University of Iowa, Iowa City (GRB). The research was supported in part by NSF grant SES-99-75083 to the second author. Please do not cite without permission. Comments welcomed. Revision accepted for publication DATE.

Address correspondence to Noel Brewer, Department of Health Behavior and Health Education, School of Public Health, University of North Carolina at Chapel Hill, 306 Rosenau Hall, CB#7440, Chapel Hill, NC 27599; e-mail: ntb1@unc.edu.

DOI: 10.1177/0272989X06297102

*Wansink, Kent, and Hoch (1998)[25] found that consumers appear to anchor on numbers used in supermarket discount offers. They reported that advertising sale items at, say, 5 for a dollar yielded higher supermarket sales than when the same items were advertised at 20 cents apiece, a finding that may be anchoring but could also be the result of establishing a norm.

introduced by Tversky and Kahneman (1974).[3] Participants first compare a target to an irrelevant anchor number (this article concerns only such uninformative anchors). An example of this *comparative judgment* is the query, "Is the chance you will get the flu more or less than 90%?" in which the anchor is "90%" and the target is "get the flu." Participants next make an *absolute judgment* in which they offer a point estimate of the target. An example is, "What is the chance you will get the flu?"

The design offers a useful, experimentally controlled paradigm for addressing our question of how anchors might influence choice. Consider a simple extension of the standard anchoring design: a comparison to the anchor, followed by a judgment, followed by a choice. Here is an example with stimuli tailored to the medical context. A woman compares her chance of getting the flu to an irrelevant anchor number, she judges her risk of getting the flu, and she chooses whether to have a flu shot.

Hypotheses about an Anchoring Effect in Choice

We elaborate here on several hypotheses of increasing complexity—and decreasing parsimony—for how and why irrelevant anchors may affect choice. This is, How might the woman's decision to be vaccinated be influenced by the anchor number?

Assimilation and the Activation Hypothesis

The simplest hypothesis is that choice will assimilate toward the anchor. For example, considering a large anchor might increase one's perceived risk for the flu and cause one to get a flue shot. The hypothesis is drawn from the selective accessibility model[4–7] and similar theoretical approaches[8,9] that tie the anchoring effect (elicited by irrelevant anchor numbers) to biased activation of knowledge. The comparative judgment (in the standard anchoring design) triggers a process of positive hypothesis testing: People consider whether the anchor number is the correct answer to the comparative question. This triggers a process of confirmatory search[8,9] that causes people to retrieve examples confirming that the anchor may be the correct answer. Although the hypothesis that the anchor is the correct answer is invariably rejected, considering the anchor causes a biased set of information to be activated in memory.[4,5] When asked to make judgments on the anchored topic, the information brought to mind will be biased and can, in turn, bias the judgment in the direction of the anchor. The result in most cases will be an assimilation effect of judgment toward the anchor.

Based on the activation account of the anchoring bias, a reasonable hypothesis is that the anchor-driven assimilation effect in judgment will be passed on to choice. The result would be choice assimilating toward the anchor. It is important to note that the assimilation effect in choice should be mediated by judgment. Thus, if judgment is covaried from the anchor-choice correlation, the correlation would drop from being positive and significant to zero and not significant. For example, a high percentage anchor would cause greater flu vaccination and this effect would be fully mediated by perceived risk for getting the flu. In summary, the activation hypothesis concerns the target and suggests an assimilation effect in choice.

Contrast and the Scale Distortion Hypothesis

A more complex and consequently less parsimonious line of thinking suggests that choice should contrast away from the anchor. Although several arguments yield the contrast prediction, we focus here on the scale distortion hypothesis; we return to the others in the general discussion. The scale distortion hypothesis concerns problems created by mapping the anchor onto the response scale (during the initial comparative judgment). Units of the response scale closest to where the anchor is mapped become magnified, and those farther away collapse. In the case of a 90% anchor, one comes to appreciate the subtle differences between, say, 85% and 90% but cannot make a similar distinction for the same magnitude difference at the bottom of the scale. This argument is somewhat related to Parducci's range-frequency theory.[6]

One implication of the scale distortion hypothesis is that anchors will trigger an assimilation effect in judgments. Because the area of the response scale nearest the anchor is magnified, any judgment will have a greater chance of falling nearer to the anchor than otherwise. Furthermore, making the comparison of "greater than or less than" requires at least an implicit sense of whether the target is small or large, even if one is not asked that explicitly. Stating that one's chance of getting the flu is smaller than 90% probably requires that one have a sense of "a lot less" or "a little less" if not an actual point estimate. The scale distortion hypothesis suggests that the distance between the anchor and the rough or precise target estimate will seem *larger* than it otherwise would. For example, if I think my risk of getting the flu is about 50%, that will seem further away from 90%—and subjectively lower—than if no anchor had been considered. This subjective sense of "big" (or "small") will yield a contrast effect in choices that are based on the risk judgment. To summarize, the scale distortion hypothesis concerns the judgment's response scale and suggests an assimilation effect in judgment and a contrast effect in choice.

No Effect

The last possibility we consider is the least parsimonious: Two separate processes will yield overlapping assimilation effects in judgment and self-canceling assimilation and contrast effects in choice. The result of greatest interest is the absence of a relationship between anchor and choice. This prediction for choice combines the activation and the scale distortion hypotheses described previously. Research has amply supported anchor-elicited assimilation effects in judgment. Much of this research points to a selective activation effect, but it does not eliminate the possibility of an additional assimilation effect of another sort, for example, the one suggested by scale distortion.

None of this research allows us to make empirically supported statements about the effect of anchors on choice. Combining the predictions for choice made by the activation and scale distortion hypotheses suggests no observed effect of anchor on choice because the assimilation and contrast effects would offset each other. However, the activation (assimilation) effects in choice are mediated by judgment and the scale distortion effects in choice are not. This observation leads to the interesting insight that covarying judgment out of the anchor-choice relationship should reveal a suppressed contrast effect. Such a finding would show that a mediated assimilative effect passed through judgment suppressed a direct contrast effect of anchor on choice.

In our initial theorizing, we favored the very parsimonious prediction of assimilation of the choice to anchor. The experiments we report here produce a distinctive pattern of data that call for a more complex explanation. The purpose of the present article is to demonstrate the unexpected influence of an irrelevant anchor number on medical treatment choices. We demonstrate this in a first experiment and then replicate the finding in a second. Because the primary purpose of this article is to demonstrate the effect, we do not conclusively rule out all possible alternative explanations and we return to this issue in the general discussion.

EXPERIMENT 1: HIV+ PATIENTS

The purpose of the first experiment was to demonstrate what effect irrelevant anchor numbers have on health-related choices. The experiment used the standard anchoring design (a comparative judgment followed by an absolute judgment) with the added feature of participants making several treatment choices after their judgment. Participants were HIV-positive patients making judgments and choices related to HIV infection. The experiment was a 2-condition (low and high anchor level) between-subjects design.

Method

Participants

Patients ($N = 99$) diagnosed with human immunodeficiency virus (i.e., who are HIV+) were recruited at the dining facility of the Gay Men's Health Crisis in New York City and paid $5 for participating in the study. The patients had a median age of 42, were mostly men (91%), and were ethnically diverse (37% Latino, 27% White, and 26% Black). They were moderately well educated (40% had a college education), they had low annual incomes (median $7,500 per year), and many had HIV cases severe enough to prevent them from holding a job (40%).

Surveys

Surveys contained a vignette followed by 3 questions requesting comparative and absolute judgments of likelihood and the treatment plan the patient would recommend. Data for female participants ($n = 9$) were dropped as the vignette assumed a sexual activity requiring insertive sexual intercourse. We did not know before conducting the study that Gay Men's Health Crisis served women. Data were also dropped for 10 other participants who did not offer both judgments and a treatment choice, leaving 80 participants in the analyses.

The vignette read, "Please imagine yourself in the following sexual situation. If you are not sexually active or even if you don't have this type of sex, please give us your best answer based on how you think you would react. Imagine you are having insertive anal sex (you are the top) with a person you know to be HIV-negative. You are using a condom and ejaculate (cum) while inside your partner. After you withdraw you realize that the condom broke while you were having sex."

Judgment

Participants were asked to consider the chance that the person they had sex with would become infected with HIV because the condom broke during sex. They first offered a comparative judgment as to whether the chance of infection was more or less than the anchor number, 1% (or 90%). They then offered an absolute estimate of the chance of HIV infection. The anchor number (i.e., 1% or 90%) was uninformative in that it did not give clues to a normatively correct estimate. A high or low anchor value was randomly assigned to each participant, but participants were not explicitly told that the anchor values were randomly selected, a practice that is common in anchoring studies.[10–12]

Treatment Choice

Patients reported the treatment(s) they would recommend to their partner. A checklist of treatments was provided, and patients were instructed to check all that applied. The list contained options that read, Do nothing because they are probably okay; Wait and see if they have symptoms like fever or achiness; Use alternative medicine such as herbs to boost their immune system; Have an HIV test; See a doctor; and Begin "post-exposure prophylaxis," a drug treatment for people who may have been exposed to HIV. A last option asked patients to write in any other treatment options they would recommend.

A composite score for treatment aggressiveness was created for each patient. Patients who recommended no medical attention (i.e., recommended doing nothing, watchfully waiting, or taking an herbal supplement) received a score of "1" ($n = 9$). Patients whose most aggressive recommendation was to see a doctor received a score of "2" ($n = 16$), have an HIV test "3" ($n = 29$), and treat with medication "4" ($n = 26$). The coding scheme reported is one of several logically defensible options. Several alternate scoring systems were also examined. For example, in one arrangement, doing nothing, watchful waiting, and taking an herbal supplement were separated into their own categories, and in another, testing and seeing a doctor were collapsed into a single scoring category. The results obtained using the alternative scoring systems do not differ meaningfully from those reported below.

Results

Patients' judgments assimilated toward the anchors. Patients in the low anchor group estimated the likelihood of HIV infection to be 43% on average, well below the 63% average estimate by the high anchor group. The difference was statistically significant, $t(79) = 2.75$, $P = 0.01$, $\beta = +.28$. This replicates the well-known anchoring bias in a clinically important judgment and population.

As one would expect, higher judgments of HIV infection predicted more aggressive treatment recommendations on the overall measure ($\beta = .42$, $P < 0.001$) as well as the recommendation for testing and taking medication (Ps < 0.05). Covarying out anchor level from the judgment-choice relationship caused it to become larger.

Table 1 Predictors of HIV+ Patients' Treatment Choices (Experiment 1)

	Anchor Level β	Risk Judgment β
Watchfully wait	.18 (.09)	–.28 (–.23)
See a doctor	–.17 (–.11)	.21 (.15)
Test	–.23 (–.08)	.45* (.37*)
Medicate (with antiviral medication)	–.14 (–.04)	.35* (.31*)
Treatment Choice Aggressiveness Index	–.22* (–.10)	.42** (.36*)

Note: Bivariate relationships shown in parentheses; mediated relationships shown outside parentheses, that is, after covarying out the effects of variable in adjacent column.
*$P < 0.05$. **$P < 0.001$.

We turn now to the main outcome of interest, the effect of anchor on treatment choice. There was no observed relation of anchor to choice, but the findings indicate a suppressed contrast effect. Anchor level showed no observed effect on treatment choice aggressiveness in bivariate analyses. Anchor level did not predict overall treatment aggressiveness ($\beta = -.10$, $P > 0.35$), nor did anchor level predict any of the individual treatment actions in several logistic regressions (βs shown in Table 1).

However, covarying judgment—to determine the mediated effect of anchor level on overall treatment choice aggressiveness—revealed a significant contrast effect. Lower anchor level predicted more aggressive treatment ($\beta = -.22$, $P < 0.05$). That is, considering the 1% anchor yielded more aggressive treatment choices than considering the higher anchor. The predicted mean treatment choice score in the low-anchor condition was 3.11, and in the high-anchor condition, it was 2.68. The individual treatment options also suggested a suppressed contrast effect in the form of more waiting, and less testing, hospitalizing, and treating, but none of these individual tests were significant.

The pattern of results represents a suppression effect, a type of mediation that occurs when a previously nonsignificant relationship is brought to significance by holding constant the effect of a third variable. Furthermore, it is a suppressed contrast effect because high-risk anchors were related to less aggressive treatment choices. Paired with the findings for judgment, it appears that high anchors increased judgments and simultaneously inhibited treatment aggressiveness.

Discussion

The results of Experiment 1 show a strong anchor-driven assimilation effect in judgment, with high anchors driving up estimates of risk for HIV infection. The assimilation effect was passed on to treatment choice but opposed by a contrast effect in choice; the net effect was no observed relation of anchor to choice. These opposing forces were revealed when judgment was covaried. The suppression effect was clearly visible in the aggregated treatment choices but only implied in the individual treatment choices. Holding judgment constant, patients were less likely to suggest more aggressive treatment for HIV infection when they received a high rather than a low anchor.

The findings support the most complex explanation we formulated in the introduction: The data are consistent with the simultaneous operation of both the activation hypothesis and the scale distortion hypothesis. Choices contrasted away from the anchor as suggested by the scale distortion hypothesis. The prediction of assimilation in choice that was drawn from the activation account of the anchoring bias was also indirectly supported because this mediated assimilation effect suppressed the contrast effect.

We have argued for a scale distortion hypothesis, but the findings might also be consistent with a deliberate correction process[13–15] (for a review, see Ford and Thompson[16]). There was no "observed bias" insofar as the mediated assimilation effect and the direct contrast effect in choice combined for a net zero effect. Perhaps patients were aware of the assimilation effect in judgment but unable or unwilling to control it and instead revised their choices to be unbiased. This explanation is unsatisfactory, as the anchoring bias in judgment has shown to be outside of conscious awareness[5,8,9,17] and robust against debiasing.[3,18] Wilson and others[2] (studies 4 and 5) attempted to eliminate anchoring by offering a $50 payment to the most accurate participant, warning participants that the anchor would bias their answers, and even warning them of the direction in which the anchor would bias their answer. All of these attempts at debiasing failed. People were unaware of the anchoring bias in judgment and consequently were unable to debias their judgments. If participants are generally unaware of the assimilation effect in judgment, then it seems unlikely that they would be aware of such an effect in choice.

Even so, we wish to give the corrective contrast explanation a fair test. Perhaps people are concerned only with a potential bias in their choices resulting from the anchor numbers (e.g., they have a naive theory that anchors can bias choices but don't affect judgments). Perhaps people attempt to correct for a potential effect of the anchor on their choices but overcorrect, and the result is the direct contrast effect of anchor on choice.

Concern about being evaluated by others (e.g., accountability) could increase people's motivation to be accurate thus eliminating the anchoring bias in choice. Increasing accountability has been shown to motivate participants to reduce or eliminate biases such as the fundamental attribution error.[19] Accountability has also been shown to exaggerate other biases such as the dilution effect (i.e., the tendency of irrelevant information to disrupt the use of diagnostic evidence in prediction tasks) as participants pay added attention to normatively irrelevant information.[20] Because the effect of accountability on anchoring has not yet been tested, we formulated a hypothesis that follows from the corrective contrast account: Accountable participants would take even greater care in making their treatment choices and thus eliminate the contract effect in choice.

EXPERIMENT 2: PHYSICIANS

The second experiment sought to replicate the basic finding of suppressed contrast and to extend the finding to the treatment choices of physicians, a population of experts skilled at medical decision making. Furthermore, the experiment examined whether increasing the physicians' motivation for accuracy (i.e., by increasing accountability) would moderate the anchoring bias. The experiment employed a 2 (Anchor Level) × 2 (Accountability) between-subjects design.

Method

Participants

Surveys were mailed to 461 family practice physicians who were members of the Iowa Medical Association. At 2-week intervals, for the 8 weeks following the initial mailing, nonresponders received either a reminder postcard or a duplicate survey packet. Completed surveys were received from 191 physicians (response rate 41%). Physicians were not paid for their participation.

Surveys

Surveys contained a cover page of instructions and a clinical vignette. (The survey contained an additional 2 vignettes that are reported elsewhere as they examined a topic unrelated to anchoring.[21] After the vignette, participants answered 3 questions. They offered a comparative and absolute judgment of the likelihood of illness, and the treatment plan they would recommend. Data were dropped for 11 participants who did not answer one or more of the questions (comparative judgment, absolute judgment, or treatment choice).

The clinical vignette introduced a case of possible pulmonary embolism, an acute and potentially lethal disease. Pulmonary embolism is notoriously difficult to diagnose, and consequently the diagnostic process is accompanied by a high degree of uncertainty. Elements of the vignette amplified the level of uncertainty by including diagnostic feedback of which some elements tended to suggest pulmonary embolism and some tended to deny it. The vignette read as follows:

> Possible Pulmonary Embolism
>
> The patient is a 32-year-old woman presenting at an acute visit with the following symptoms: cough, pleuritic pain, and a low-grade fever. Over the last two years, the patient has had a substantial weight gain that has caused her BMI to elevate to 34. She is short of breath, has an elevated heart rate (104 bpm) but normal blood pressure (130/80). Her legs show no edema or sensitivity to pressure. A cardiac ausculatory exam shows increased P2. The patient appears anxious about her presenting problems but otherwise healthy with no history of IDU or recent surgery. Her chest x-ray returns near normal and an EKG shows non-specific ST-T wave changes. A blood test shows that arterial P_{O2} is 83 mmHg and P_{CO2} is 36 mmHg, both of which are on the low side of the normal range.

Judgment

After reading the vignette, participants gave a comparative judgment: whether the chance of a pulmonary embolism in the hypothetical patient was greater or less than 1% (or 90%). They then gave an absolute judgment: a point estimate of the chance of pulmonary embolism in the patient.

Treatment Choice

Physicians reported the next step(s) they would take in treating the patient. Physicians were instructed to check all that applied from a list: wait and see whether the patient's condition improves with normal care, order a lung scan, order a pulmonary angiography (i.e., x-ray), hospitalize the patient, and treat with warfarin (or other anticoagulant). A last option asked physicians to write in any other treatment options they would pursue.

A composite score for treatment aggressiveness was created for each physician. Physicians who opted to watchfully wait received a score of 1 (n = 19). Physicians whose most aggressive treatment was to test were scored 2 (n = 112), and those who treated with medication or hospitalized were scored 3 (n = 48). The coding scheme reported is one of several logically defensible options. As in Experiment 1, several alternate scoring systems were also examined. For example, in one arrangement, hospitalization was scored lower, and in another, treating and hospitalization were separated. The results obtained using the alternate coding schemes did not differ meaningfully from those reported below.

Accountability

Physicians were randomly assigned to receive a survey that either did or did not make accountability salient. In the accountable condition, the survey's first inside page indicated that the physicians' responses would be reviewed by "members of a recently formed patient advocacy group, headquartered in Iowa City." They were further asked to indicate times and days of the week that they might be contacted should the reviewers have any questions. After making the judgments and treatment decisions about the scenario, they were also asked to provide a written justification of their answers. In the nonaccountable conditions, the first inside page of the survey was left blank and participants were not asked to justify their answers. All analyses included terms modeling the main and moderating effects of accountability.

Results and Discussion

Physicians' judgments strongly assimilated to the irrelevant anchor numbers. Physicians in the low anchor group estimated the likelihood of pulmonary embolism to be 23% on average, less than half of the 53% average estimate of those in the high anchor group. A 2 (Anchor Level) × 2 (Accountability) ANOVA that predicted judgment revealed a significant main effect of Anchor Level, $F(1, 175) = 51.12$, $P < 0.001$, $\beta = +.49$.

Higher judgments predicted more aggressive treatment choices both in the summary variable ($\beta = +.44$, $P < 0.01$) and in all individual variables (Ps < 0.05)

Table 2 Predictors of Physician's Treatment Choices (Experiment 2)

	Anchor Level β	Risk Judgment β
Watchfully wait	.67** (.30*)	−.98* (−.55*)
Test	−.28* (−.09)	.34*(.18)
Hospitalize	−.24* (−.00)	.45** (.33*)
Medicate (with anticoagulant)	−.33* (−.02)	.56** (.40*)
Treatment Choice Aggressiveness Index	−.54** (−.15)	.75** (.44**)

Note: Bivariate relationships shown in parentheses; mediated relationships shown outside parentheses, that is, after covarying out the effects of variable in adjacent column.
$^*P < 0.05$. $^{**}P < 0.001$.

except testing. Covarying out anchor level from the relation between judgment and choice strengthened all relations such that all became larger.

The most important finding was a replication of the suppressed contrast effect in choice as summarized in Table 2. Anchor level again showed little or no observed relation to treatment choice. Instead, choices showed a contrast effect that was suppressed by a mediated assimilation effect. A 2 (Anchor Level) × 2 (Accountability) ANOVA that predicted the treatment aggressiveness score showed that treatment aggressiveness was unrelated to anchor level ($\beta = -.15$, *n.s.*). When judgment was covaried out of the relation between anchor level and choice, the relationship became significant. Lower anchor level predicted more aggressive overall treatment scores ($\beta = -.54$, $P < 0.001$). The predicted mean overall treatment choice score in the low anchor condition was 2.32, and in the high anchor condition it was 2.01.

In separate 2 (Anchor Level) × 2 (Accountability) logistic regressions predicting each of the 4 treatment options (i.e., wait, test, hospitalize, medicate), only watchful waiting showed a significant bivariate relationship to anchor level (as shown in Table 2). The latter relation showed an observed contrast effect such that higher anchor number predicted greater watchful waiting ($\beta = +.30$, $P < 0.05$). This is a contrast effect because higher probability judgments were related to lower treatment aggressiveness. Analyses of the individual treatment choices that covaried judgment showed significant contrast effects with lower anchors predicting more waiting and less testing, hospitalizing, and treating (Ps < 0.05).

Accountability

Accountability did not moderate the anchoring effects reported above. The average pulmonary embolism likelihood estimates for the nonaccountable and accountable conditions were equivalent in the low (23% v. 20%) and high (52% v. 53%) anchor conditions. The 2 × 2 ANOVA predicting judgment revealed no main effect of or interaction with accountability, Fs(1, 175) < 1, *n.s.* Similarly, accountability did not moderate the relation of anchor to choice. The 2 × 2 ANOVA predicting choice revealed no effects of accountability, Fs(1, 175) < 1, *n.s.* Adding judgment to ANOVA as a covariate revealed no interactions with accountability.

As we have noted elsewhere, the response rate in the accountable condition was about half that in the nonaccountable condition.[21] The absence of a main effect for accountability on risk judgment or choice suggests that there was not a selection bias. This conclusion is bolstered by the equivalent ages, years in practice, and gender[21] between participants and nonresponders, and between accountable and nonaccountable conditions. Although a reasonable concern about the null results is that the accountability manipulation was ineffective, results from the other study conducted concurrently show effects of accountability.[21] The study showed that the physicians' violations of the normative rule of regularity were more exaggerated under accountability. Given Schwartz and other's[21] finding of an effect of the accountability manipulation, we conclude that accountability does not alter anchoring effects of the type we report here. Because increased accuracy motivation did not reduce the anchoring effect in choice, deliberative correction that requires accuracy motivation is an unlikely explanation for the suppressed contrast effect.

GENERAL DISCUSSION

The results of these 2 experiments with patients and physicians demonstrate the surprising conundrum that anchors cause a large bias in health risk judgments but cause no apparent bias in treatment choices. This finding is true for both patients and physician experts. Making sense of this unintuitive pattern of findings requires that we posit 2 opposing processes. The simplest account, that anchors cause the selective activation of information in memory, cannot alone account for the suppressed contrast effect in choice. We discuss the theoretical issues that the experiments raise and then discuss several clinical implications.

We have proposed that the additional contrast effect in choice is due to response scale distortion. Mapping the anchor onto the response scale causes an enhanced sensitivity in the range of the response scale closest to the anchor and diminished sensitivity farthest away. Absolute judgments mapped onto the distorted response scale shift closer to the anchor causing the judgment to resemble the anchor. Yet the subjective understanding of the difference between anchor and judgment seems large and causes choices to contrast away from the anchor. Both experiments support the scale distortion hypothesis. In Experiment 2, physicians treated less aggressively when they received a high anchor when compared to those who received a low anchor (after controlling for judgment). In Experiment 1, HIV-positive patients were less likely to suggest more aggressive treatment for HIV infection when they received a high rather than a low anchor (again controlling for judgment). The prediction of a mediated assimilation effect in choice drawn from the activation account of the anchoring bias was also supported. The mediated assimilation effect offset the contrast effect in choice such that the observed relation of anchor to choice was nil until judgment was held constant.

The 2 experiments support the scale distortion account, and the data from Experiment 2 tend to rule out a deliberative corrective contrast account. Presumably, physicians in the accountable condition should have been additionally motivated to reduce the bias in their choices, but they did not. Because neither this experiment nor any other experiments we are aware of have shown that accountability (or more explicit forms of accuracy motivation) reduced the effect of irrelevant anchors on judgments, the corrective contrast argument must assume that participants were unaware of a bias in judgment. Given this, corrective contrast would then have to assume that participants are somehow motivated to debias choice but not judgment. Furthermore, to account for the present findings, it would have to assume that physicians believed that their correction was just about right (and accounted for an unknown and compensatory bias passed through judgment) and that, even under conditions of accountability, no additional correction was required. We do not find the correction account tenable.

The data do not rule out several alternative theoretical accounts. It is possible that the anchor influences an unmeasured third variable that itself affects choice but not judgment. Another potential explanation is that the contrast effect is triggered by a deliberate comparison of anchor and judgment, in a manner similar to comparison contrast.[7,22] A last possibility is the finding by Mussweiler and Strack that anchors lead to assimilation in objective judgments but lead to contrast effects in subjective judgments.[6] Treatment choices do not seem inherently subjective in nature, but the sense of treatment aggressiveness captured by our summary choice measure may tap into some underlying subjective dimension. Sorting out the additional explanations will require additional research, but the present experiments demonstrate an interesting and potentially important new finding. More important, the studies demonstrate that anchoring effects in judgment are not necessarily passed on to choice, or at least not in a way that is easily observed.

Another limitation of the present studies is that we did not directly assess the perceived influence of the anchors. As we have already noted, some participants in other studies have perceived an influence of anchor numbers, but they are often wrong about the direction; the upshot is that perceived influence is unrelated to observed bias. Furthermore, even when explicitly warned about the presence and direction of the bias, people are unable to debias their judgments. Finally, past research has shown that anchor numbers bias judgments even when they are perceived to be completely uninformative and perceived not to have an influence.[9,23,24] This previous research (and Experiment 2 using the accountability manipulation) strongly suggests that the present findings are not a methodological artifact.

The research has several potential clinical implications. The studies demonstrate the effect in 2 real-world medical populations asked to make judgments and choices within their area of experience. This gives us confidence that we are talking about effects that are general in nature and applicable to clinical judgments and decisions. The most obvious implication of the research is that irrelevant anchor numbers may bias judgments but do not appear to have a large impact on related treatment choices. Similar biases already demonstrated in judgment will need to be demonstrated in choice rather than assuming their generality. The findings leave open the possibility that an adviser with an anchor-biased judgment may give poor advice that is then followed by a patient. Thus, a physician with a biased risk estimate may communicate it to a patient who could then act on it. This series of events seems entirely plausible and would represent an anchoring bias in choice. However, this speculation awaits empirical validation.

The anchoring effects for patients were quite a bit smaller than for physicians (judgment: .28 v. .44; choice: −.22 v. −.54). One conclusion is that physicians are more susceptible to the bias. We think that is a premature and perhaps incorrect conclusion. It is possible that differences in the 2 scenarios themselves influenced the effect sizes. More likely in our opinion is that the reliability of the participants' judgments and choices differed. Patients showed a smaller relationship between their judgments and choices (.42 v. .75, after covarying anchor), one measure of how internally consistent the 2 groups' responses were. Physician responses may have contained less random error, and this higher reliability may have allowed us to obtain larger parameter estimates.

The anchoring bias has presented longstanding fascination for those in the field of judgment and decision making. The present findings suggest that irrelevant anchors may have more complex effects than initially thought, particularly when the bias extends from judgment to choice. Models of the anchoring bias may require refinement to better reflect such findings.

REFERENCES

1. Chapman GB, Johnson EJ. Incorporating the irrelevant: anchors in judgments of belief and value. In: Gilovich T, Griffin DW, Kahneman D, eds. The Psychology of Judgment: Heuristics and Biases. New York: Cambridge University Press; 2002. p 120–38.
2. Wilson TD, Houston CE, Eitling KM, Brekke N. A new look at anchoring effects: basic anchoring and its antecedents. J Exp Psychol General. 1996;125:387–402.
3. Tversky A, Kahneman D. Judgments under uncertainty: heuristics and biases. Science. 1974;185:1124–31.
4. Mussweiler T, Strack F. Hypothesis-consistent testing and semantic priming in the anchoring paradigm: a selective accessibility model. J Exp Soc Psychol. 1999;35:136–64.
5. Mussweiler T, Strack F. Numeric judgments under uncertainty: the role of knowledge in anchoring. J Exp Soc Psychol. 2000a; 36:495–518.
6. Mussweiler T, Strack F. The "relative self": informational and judgmental consequences of comparative self-evaluation. J Person Soc Psychol. 2000b;79:23–38.
7. Strack F, Mussweiler T. Explaining the enigmatic anchoring effect: mechanisms of selective accessibility. J Person Soc Psychol. 1997;73:437–46.
8. Chapman GB, Johnson EJ. The limits of anchoring. J Behav Decis Making. 1994;7:232–42.
9. Chapman GB, Johnson EJ. Anchoring, activation, and the construction of values. Org Behav Human Decis Proc 1999;79:1–39.
10. Chapman GB, Bornstein BH. The more you ask for, the more you get: anchoring in personal injury claims. Appl Cogn Psychol. 1996;10:519–40.
11. Epley N, Gilovich T. Putting adjustment back in the anchoring and adjustment heuristic: divergent processing of self-generated and experimenter provided anchors. Psychol Sci. 2001;12:391–6.
12. Jacowitz KR, Kahneman D. Measures of anchors in estimation tasks. Person Soc Psychol Bull. 1995;21:1161–6.
13. Martin LL. Set/reset: use and disuse of concepts in impression formation. J Person Soc Psychol. 1986;51:493–504.
14. Martin LL, Clark LF. Social cognition: exploring the mental processes involved in human interaction. In: Eysenk MW, ed. Cognitive Psychology: An International Review. Chichester (UK): Wiley; 1990. p 265–310.
15. Petty RE, Wegener DT. Flexible correction processes in social judgment: correcting for context-induced contrast. J Exp Soc Psychol. 1993;29:137–65.
16. Ford TE, Thompson EP. Preconscious and postconscious processes underlying construct accessibility effects: an extended search model. Person Soc Psychol Rev. 2000;4:317–36.
17. Mussweiler T, Strack F, Pfeifer T. Overcoming the inevitable anchoring effect: considering the opposite compensates for selective accessibility. Person Soc Psychol Bull. 2000;26:1142–50.
18. Quattrone GA, Lawrence CP, Finkel SE, Andrus DC. Explorations in anchoring: the effects of prior range, anchor extremity and suggestive hints. Unpublished manuscript, Stanford University, Stanford, CA, 1984.
19. Tetlock PE. Accountability: a social check on the fundamental attribution error. Soc Psychol Q. 1985;48:227–36.
20. Tetlock PE, Boettger R. Accountability amplifies the status quo effect when change creates victims. J Behav Decis Making. 1994;7:1–23.
21. Schwartz J, Chapman GB, Brewer N, Bergus GR. The effect of accountability on physicians' decisions: going from bad to worse. Psychonom Bull Rev. 2004;11:173–8.
22. Schwarz N, Bless H. Constructing reality and its alternatives: an inclusion/exclusion model of assimilation and contrast effects in social judgment. In: Martin LL, Tesser A, eds. The Construction of Social Judgments. Hillsdale (NJ): Lawrence Erlbaum; 1992. p 217–45.
23. Brewer NT, Chapman GB. The fragile basic anchoring effect. J Behav Decis Making. 2001;15:65–77.
24. Russo JE, Shoemaker PJH. Decision Traps. New York: Simon & Schuster; 1989.
25. Wansink B, Kent RJ, Hoch SJ. An anchoring and adjustment model of purchase quantity decisions. J Market Res. 1998;35:71–81.