

# The effects of accountability on bias in physician decision making: Going from bad to worse

JANET SCHWARTZ, GRETCHEN CHAPMAN, and NOEL BREWER  
*Rutgers University, Piscataway, New Jersey*

and

GEORGE BERGUS  
*University of Iowa, Iowa City, Iowa*

Members of the Iowa Academy of Family Physicians participated in a survey study in which they were asked to make hypothetical decisions in either high- or low-conflict treatment conditions. In the low-conflict treatment condition the options were a common medication with a referral or a referral only. In the high-conflict treatment condition the same two options were given plus another attractive medication. In addition, we manipulated accountability by asking half of the participants to provide a written defense of their treatment options, which they would then agree to discuss at a later time. The results showed that physicians in the high-conflict condition chose the referral-only option significantly more than the physicians in the low-conflict condition, thus violating the normative rule of regularity. Moreover, that pattern was significantly amplified for physicians who were held accountable for their treatment decisions. These findings replicate and extend previous research and are discussed within the framework of reason-based choice.

Common sense dictates that holding people accountable for important decisions is an effective way to ensure careful and deliberative choices. Despite this intuition, however, a recent review of the accountability literature (Lerner & Tetlock, 1999) shows that some decision-making biases become stronger when decision makers are held accountable (e.g., Simonson, 1989). By *accountable*, we specifically mean predecisional accountability where decision makers can expect to be called upon to justify their behavior to others. Such accountability has been shown to increase most notably preference reversals that result from simply adding new alternatives to a list of existing options (Simonson & Nye, 1992). The purpose of the present article is to test whether accountability will increase the likelihood of a conflict avoidance bias that results from adding a new attractive option to a set of existing choices (Redelmeier & Shafir, 1995; Tversky & Shafir, 1992) in physician decision makers.

As an example of conflict avoidance, consider the following scenarios taken from Tversky and Shafir (1992):

### Scenario 1, low conflict:

Suppose you are considering buying a compact disc (CD) player, and have not yet decided what model to buy. You

pass by a store that is having a one-day clearance sale. They offer a popular Sony player for just \$99, well below the list price. Do you?

- B. buy the Sony player
- C. wait until you learn more about the various models

### Scenario 2, high conflict:

Suppose you are considering buying a compact disc (CD) player, and have not yet decided what model to buy. You pass by a store that is having a one-day clearance sale. They offer a popular Sony player for just \$99, well below the list price, and a top-of-the-line Aiwa for just \$169. Do you?

- A. buy the Aiwa
- B. buy the Sony
- C. wait until you learn more about the various models

In the low-conflict condition, the results showed that the majority of participants preferred the Sony CD player to waiting. However, in the high-conflict condition, where another attractive CD player was offered, the majority chose to wait to learn more about the various options. This phenomenon represents a decision-making bias because it violates the normative principle of regularity. Regularity states that the preference ordering among alternatives should not be influenced by the addition of a new option, provided that the new alternative offers no additional information. Thus, adding another attractive CD player should not make the option of waiting more attractive than the Sony.

---

This research was supported by NSF Grant SES-99-75083 awarded to the second author. Thanks are due three anonymous reviewers for their insight and helpful comments on a previous draft. We are also grateful to the members of the Medical Decision Making Lab at Rutgers University for their assistance with the mailing and tracking of all questionnaires. Correspondence concerning this article should be addressed to J. Schwartz, Robertson Hall, Woodrow Wilson School, Princeton, NJ 08544-6891 (e-mail: jaschwar@princeton.edu).

Tversky and Shafir (1992) explained the findings in terms of reason-based choice, which is a propensity for decision makers to search for alternatives that are easily defensible to oneself and others. In the low-conflict scenario there are good reasons to choose the Sony: It is a popular model and is being offered at a very fair price. In the high-conflict scenario, however, *both* CD players are attractive, though they differ somewhat in quality and price. Rather than confront those tradeoffs between quality and price, and lacking a good reason to buy one player over the other, one simply avoids choosing. The authors supported their claim that conflict avoidance was driving this effect by showing that adding a dominated alternative (e.g., an inferior Aiwa CD player offered at \$105), or an alternative that did not create conflict, did not produce greater preference for the deferral option. Thus, it is only when choosing creates conflict that decision makers simply opt out and avoid the decision.

In a later study, Redelmeier and Shafir (1995) replicated these findings with physician decision makers. Physicians were asked to make hypothetical treatment decisions about medications and surgical care. For example, consider the following scenarios:

**Scenario 1, low conflict**

The patient is a 67-year-old farmer with chronic right hip pain. The diagnosis is osteoarthritis. You have tried several nonsteroidal anti-inflammatory agents (e.g., aspirin, Naproxen, and Ketoprofen) and have stopped them because of either adverse effects or lack of efficacy. You decide to refer him to an orthopedic consultant for consideration for hip replacement. The patient agrees to this plan. Before sending him away, however, you check the drug formulary and find that there is one nonsteroidal medication that this patient has not tried, Piroxicam. Do you?

- B. start Piroxicam and refer to orthopedics
- C. refer to orthopedics and don't start any new medications

**Scenario 2, high conflict**

The patient is a 67-year-old farmer with chronic right hip pain. The diagnosis is osteoarthritis. You have tried several nonsteroidal anti-inflammatory agents (e.g., aspirin, Naproxen, and Ketoprofen) and have stopped them because of either adverse effects or lack of efficacy. You decide to refer him to an orthopedic consultant for consideration for hip replacement. The patient agrees to this plan. Before sending him away, however, you check the drug formulary and find that there are two nonsteroidal medications that this patient has not tried, Ibuprofen and Piroxicam. Do you?

- A. start Ibuprofen and refer to orthopedics
- B. start Piroxicam and refer to orthopedics
- C. refer to orthopedics and don't start any new medications

Similar to the pattern of results with the CD player scenario in the Tversky and Shafir (1992) study, the majority of physicians in the low-conflict condition chose to

start Piroxicam and give a referral. However, in the high-conflict condition, the majority chose only to write a referral.

The conflict avoidance seen with experts in this study is not the only bias to result from adding new alternatives. Schwartz and Chapman (1999) showed that adding an inferior medication choice to a set of existing options increased the likelihood that one of the original medications would be chosen. This phenomenon, called the *attraction effect* (Huber, Payne, & Puto, 1982), is also a violation of the regularity principle. The attraction effect can also be explained by the reason-based choice account (Shafir, Simonson, & Tversky, 1993; Simonson, 1989), because people may rely on the dominance of one option over another as a means of making decisions that can be easily justified. In fact, empirical evidence has shown that the strength of the attraction effect increases when decision makers anticipate having to defend their responses (Simonson, 1989).

The primary goal of this study is to demonstrate that accountability amplifies the conflict avoidance bias, as it does the attraction effect, and thus show a common psychological mechanism among regularity rule violations. It also provides a direct test of the reason-based choice hypothesis. We then take our examination of the relationship between accountability and the conflict avoidance bias a step further to demonstrate that even experts fall prey to context effects. Given the frequency with which physicians are held accountable for their decisions—indeed it is part of their training—and their susceptibility to bias when new options become available, they are both an interesting and an important population on which to test the effects of accountability. Consequently, another goal of this study was to replicate and extend the findings of Redelmeier and Shafir (1995) and determine whether accountable physicians, relative to nonaccountable physicians, are more likely to show the conflict avoidance bias.

## METHOD

### Participants

Surveys were mailed to 461 members of the Iowa Academy of Family Physicians. 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%).

### Materials

Paper and pencil surveys were mailed to all of the physician participants. The osteoarthritis scenario was one of three hypothetical treatment scenarios about which the participants were asked to make judgments. The two other clinical scenarios were unrelated to the present study. Scenario order was varied so that half of the subjects reviewed the osteoarthritis scenario first and the other half reviewed it last.

### Procedure

This experiment employed a 2 (accountable vs. nonaccountable) × 2 (low conflict vs. high conflict) between-subjects design. Participants were randomly assigned first to the accountability condition

and then to the conflict condition. In the accountable condition, the physicians were told that a patient advocacy group in Iowa City, whose purpose was to increase the quality of physician-patient communication, would review their responses. The physicians were asked to read the scenario, choose a treatment option, and then provide a written justification of their decision to the hypothetical patient. The accountable physicians also agreed to be contacted by a member of the advocacy group, listed their availability to discuss their responses over the telephone, and signed their names to the survey before returning it. Physicians in the nonaccountable condition were mailed an anonymous survey and were simply asked to read the clinical scenarios and choose the most appropriate treatment option. Physicians in both accountability conditions were asked to choose among the treatment options and then rate the appropriateness of each option on a scale of 1 to 5 (where 1 was *the least appropriate* and 5 was *the most appropriate*).

In the low-conflict condition, the treatment options were either Diclofenac with a referral to an orthopedic specialist or simply the referral. In the high-conflict condition, another medication, Piroxicam, was added to the list of options. Note that there are some changes to the medications used by Redelmeier and Shafir (1995). Specifically, we removed Ibuprofen from the medication choices because we felt it was so ubiquitous that it would be considered unlikely that the patient had never tried it either by prescription or over the counter. Another difference between the Redelmeier and Shafir study and ours was that we used Diclofenac, not Piroxicam, as the medication presented in both the high- and low-conflict conditions.

## RESULTS

Table 1 shows, for the accountable and nonaccountable conditions, choice for Diclofenac and Piroxicam with a referral versus the referral-only option by high- and low-conflict condition. The table shows the percentage of preference for each treatment option but for the purposes of analysis, the choice outcomes were dichotomized so that responses were coded as “medication choice” or “referral-only choice.” In the high-conflict condition, that is, a selection of either Diclofenac or Piroxicam was considered a medication choice. Note that the response rate was considerably higher in the nonaccountable condition than the accountable condition. We address this issue later in the Results section.

The results of a chi-square analysis showed a significant conflict avoidance bias when pooled across accountability conditions [ $\chi^2(1, N = 191) = 10.03, p < .01$ ]. The physicians in the accountable and nonaccountable conditions combined were significantly more likely to choose the referral-only option when two attractive medications were available than when there was just one medication. This result replicates Redelmeier and Shafir (1995). We

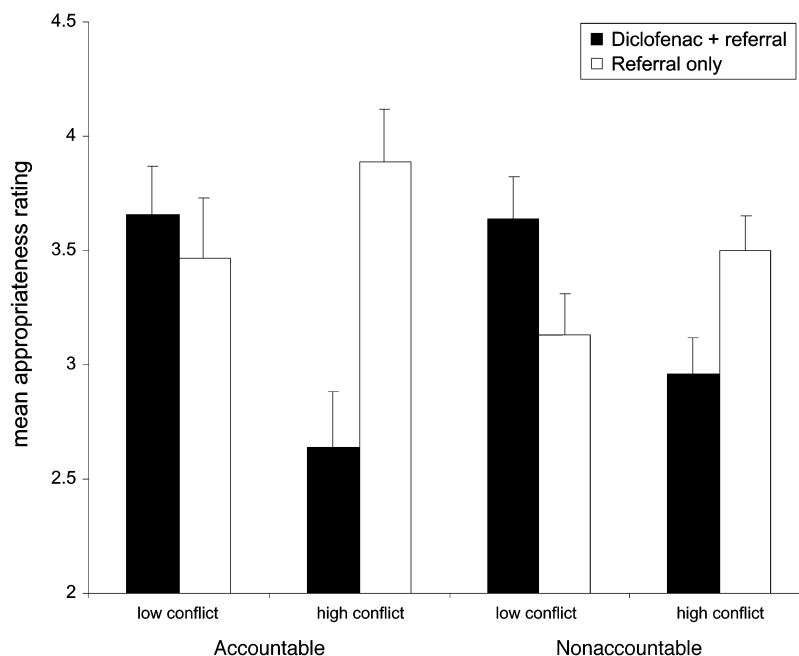
also examined choice  $\times$  conflict condition separately for both the accountable and nonaccountable groups. For the nonaccountable group, there was no significant conflict avoidance bias [ $\chi^2(1, N = 127) = 0.99, n.s.$ ], meaning that the physicians did not choose the referral-only option significantly more often in the high-conflict condition. There was, however, a significant conflict avoidance bias in the accountable condition [ $\chi^2(1, N = 64) = 10.51, p < .01$ ]. Finally, a logistic regression analysis revealed a conflict  $\times$  accountability interaction [ $\chi^2(1, N = 191) = 4.27, p < .05$ ], indicating that the conflict avoidance bias was stronger in the accountable condition than it was in the nonaccountable condition. This finding was consistent with our predictions.

Figure 1 shows the mean appropriateness rating for the Diclofenac and referral-only options  $\times$  conflict and accountability conditions. Four and 6 participants in the accountable and nonaccountable conditions, respectively, were excluded because of missing ratings. The ratings for Piroxicam were not included because that medication was rated only in the high-conflict condition. A 2 (treatment option)  $\times$  2 (conflict condition)  $\times$  2 (accountability) analysis of variance showed clear evidence of the conflict avoidance bias in the treatment appropriateness ratings. That is, there was a significant treatment option  $\times$  conflict condition interaction where the rating of the referral-only option was higher (and the medication option rating lower) in the high-conflict condition than in the low-conflict condition [ $F(1,177) = 11.0, p < .001$ ]. However, there was no treatment option  $\times$  accountability  $\times$  conflict interaction [ $F(1,177) = 0.28, n.s.$ ], meaning that the effect of the conflict manipulation was not significantly greater for the accountable physicians than for the nonaccountable physicians. This result shows a discrepancy between the outcomes of choice and rating response modes.

The response rate in the accountable condition was about half that in the nonaccountable condition. Consequently, one might worry that self-selection factors were stronger in the accountable condition than in the nonaccountable condition and that these factors, rather than the accountability manipulation, were responsible for the pattern of results. Thus, we did some post hoc analyses to determine whether there were any differences between the physicians who responded and those who did not. We examined three variables: age, number of years in practice, and gender. Four physicians from the “responder” group were excluded because no demographic information was available on them. The mean ages of the physi-

**Table 1**  
Percentage of Accountable and Nonaccountable Participants Choosing Medications (Diclofenac or Piroxicam) With a Referral Versus Referral Only by Conflict Condition

Treatment Option	Accountable ( $n = 64$ )		Nonaccountable ( $n = 127$ )	
	Low Conflict ( $n = 34$ )	High Conflict ( $n = 30$ )	Low Conflict ( $n = 58$ )	High Conflict ( $n = 69$ )
Piroxicam + referral	–	14	–	18
Diclofenac + referral	70	17	65	38
Referral only	30	69	35	44



**Figure 1.** Mean appropriateness ratings for medications with a referral versus a referral only by accountability and conflict conditions.

cians who responded ( $n = 187$ ) and did not respond ( $n = 204$ ) were 46.9 ( $SD = 8.67$ ) and 46.3 ( $SD = 9.00$ ) years old, respectively. A between-subjects  $t$  test revealed no significant difference between the mean ages of these groups [ $t(389) = .64$ , n.s.]. The responders had a mean number of years in practice of 19.1 ( $SD = 9.1$ ) versus 18.1 ( $SD = 9.3$ ) for the nonresponders [ $t(389) = .95$ , n.s.]. Finally, of the 386 (84% of the original sample) physicians for whom we had gender information, 294 were male and 92 were female. In the responder group, 22% of the participants were female versus 25% in the nonresponder group [ $\chi^2(1, N = 386) = .39$ , n.s.]. Taken together, the results showed no significant differences between the responders and nonresponders with regard to age, gender, or years of expertise.

In addition to the analyses above, we further broke down the “responder” group to examine any differences between the accountable and nonaccountable physicians. The mean ages of the accountable ( $n = 63$ ) and nonaccountable ( $n = 124$ ) physicians were 45.61 ( $SD = 8.35$ ) and 46.3 ( $SD = 8.41$ ), respectively. A between-subjects  $t$  test revealed no significant difference between the mean ages of the groups [ $t(185) = .12$ , n.s.]. The accountable physicians had a mean number of years in practice of 18.08 ( $SD = 9.1$ ) versus 18.65 ( $SD = 9.8$ ) for the nonaccountable group [ $t(185) = .22$ , n.s.]. Finally, in terms of gender, 25% of the accountable participants were female versus 17% in the nonaccountable group [ $\chi^2(1, N = 183) = 1.54$ , n.s.]. Identical to the pattern of results for comparing responders and nonresponders, there were no significant differences between account-

able and nonaccountable physicians with regard to age, gender, or years of expertise. Thus, these results suggest that responders did not differ from nonresponders and that, consequently, selection bias is an unlikely explanation of the effects of the accountability manipulation.

## DISCUSSION

The results of this study were consistent with those of previous research (Redelmeier & Shafir, 1995; Tversky & Shafir, 1992) in showing that physician decision makers were significantly more likely to choose to defer choice when two attractive options, rather than just one, were available. The conflict avoidance bias was seen with both choice and treatment appropriateness ratings. However, a significant conflict avoidance bias in choice was seen only in the accountable condition whereas the previous demonstration of this bias with physicians by Redelmeier and Shafir showed the effect in presumably nonaccountable conditions. More important, the results also showed that physicians in the accountable condition were significantly more likely to show the conflict avoidance bias than were nonaccountable physicians. The effect of accountability was apparent only in choice outcomes and not in treatment appropriateness ratings.

### Accountability

The fact that accountability amplified the conflict avoidance bias is an important finding. First, it demonstrates a common thread among regularity rule violations and suggests that the same psychological mecha-

nism underlies a certain class of biases such as conflict avoidance and the attraction effect. Second, that the effect is shown in experts suggests that critical decisions in patient care may not always result from judgments about what is best for the patient, but rather from what is most defensible for the practitioner.

We do note, however, that a potential limitation of this study was that the response rate (and hence, the number of participants) in the accountable condition was half that in the nonaccountable condition. Thus, there may have been something unique about the physicians who chose to participate in the accountable condition that made them more susceptible to the conflict manipulation. We can, however, provide some arguments against that explanation. One is that our own data revealed no differences in age, level of expertise, or gender between responders and nonresponders that would suggest a sampling problem. Another is that previous research (e.g., Simonson, 1989) has shown robust effects of accountability on other regularity principle violations when participants were not self-selected. Physicians should consequently be aware that the practice of being held accountable to patients, insurance companies, or other authoritative sources might negatively influence their decisions. They should be conscious of clever pharmaceutical marketers and advertisers, in particular, who might exploit decision makers' tendency to search for easily defensible options.

Although the effects of accountability are in no way limited to experts, there is a question as to whether the effects are stronger for experts or for naive decision makers. Physicians are one group of experts who may be especially susceptible to effects of accountability because the practice of having to justify their decisions is so routine. Consequently, they may be particularly prone to seek out easily defensible options. Redelmeier and Shafir (1995) tested for evidence of the conflict avoidance bias with legislators and medical specialists in addition to family practitioners. Each group showed the bias. It would be interesting to know whether such experts are also influenced by accountability, especially when they are compared with experts who might not incorporate justification to others in their decision-making processes.

Accountability does have some benefits beyond the obvious for those in positions of responsibility. In their review of the literature, for instance, Lerner and Tetlock (1999) highlighted many examples where accountability attenuated decision-making biases. In addition, accountability effects are useful to an understanding of decision processes. From an information-processing perspective, that is, one of the most interesting findings of the present study is that accountability had a stronger effect on choice outcomes than on treatment appropriateness ratings. An explanation is that choice relies more on qualitative heuristics than do other elicitation methods (Slovic, 1995). People are more likely to seek qualitative reasons for choosing, or not choosing in this case, than when they are making judgment ratings.

### Choice Versus Rating Differences

In the present study, the interaction between accountability and high- versus low-conflict condition was apparent for choice but not for appropriateness ratings. In analyses secondary to the main research question, we examined possible reasons for this discrepancy. No physicians gave ratings that were inconsistent with choice (e.g., choosing to write a referral but giving a higher appropriateness rating to Diclofenac). However, 6 and 13 subjects in the accountable and nonaccountable conditions, respectively, gave tied ratings to the referral and Diclofenac, whereas when asked to make a choice, they expressed a preference for one option. If we were to infer choice from ratings, these 19 physicians would have no defined choice. In order to determine whether the presence of tied ratings contributed to the discrepancy between the choice and rating results, we performed an additional analysis where we removed subjects who gave tied ratings for Diclofenac and the referral-only option from the choice analysis. The pattern of results was very similar to those from the entire sample (data not shown). The conflict avoidance bias was again significant for the accountable condition but not for the nonaccountable condition. The interaction pattern, however, was somewhat weaker and the power lower, because of the smaller sample size. As a result, the accountability  $\times$  conflict condition interaction was no longer significant. In addition, although the 19 physicians with tied ratings constituted too small a group for analysis, their choices demonstrated the predicted interaction pattern. We therefore conclude that the data from the 19 physicians with tied ratings indeed contributed to the interaction but did not explain it entirely.

If the presence of some ties in the data does not explain the different results for ratings and choice, what then explains the discrepancy? Although the present study yields no definitive answer, we suggest the following explanation. Choices require that decision makers indicate the best option. Ratings also express the best option (note that because in our study ratings followed choices, it is unsurprising that the chosen option always received the top or tied rating); however, ratings may also express how much better the best option is than the second best. If a conflict avoidance mechanism drives selection of the best option but is not used in determining how much better that option is over the other alternatives, we would expect that choices (or choices inferred from ratings) would provide the clearest picture of the relationship between accountability and the conflict avoidance bias. Ratings, alternatively, would be driven not only by a conflict avoidance mechanism but also by mechanisms used to determine the size of relative advantage of one option over another. Consequently, ratings, as compared with choices, would incorporate additional variance that diluted the conflict avoidance pattern. This is exactly the pattern observed in the present study.

This argument is consistent with the reason-based choice idea that decision makers avoid conflict while they are choosing and select options that can be most

easily defended. They are especially likely to select easily defensible options if they are made to feel accountable. Thus, people may feel that choices are in need of greater justification than ratings because the outcome is committal. Alternatively, a determination of how much better one option is than another need not be justified. As a result, accountability has less of an effect on ratings than on choice.

Judgment ratings, because they are often more sensitive than dichotomous choice outcomes, are frequently used as a proxy for choice in decision-making studies. However, researchers might want to rethink that practice in light of the present results, which suggest that the processes underlying choice and judgment ratings are different. (See Goldstein and Einhorn, 1987, for additional results on choice vs. rating differences.)

In conclusion, this research has demonstrated two important findings. The first is that accountability increases the likelihood that decision makers, including experts, will make conflict-avoiding choices. The second is that the effects of accountability vary with response mode. Taken together, these results support the reason-based choice hypothesis (Shafir et al., 1993) that people assign greater value to options they feel can be easily justified. This qualitative strategy is more evident in choice than in other response modes.

## REFERENCES

- GOLDSTEIN, W. M., & EINHORN, H. J. (1987). Expression theory and the preference reversal phenomena. *Psychological Review*, **94**, 236-254.
- HUBER, J., PAYNE, J. W., & PUTO, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, **9**, 90-98.
- LERNER, J. S., & TETLOCK, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, **125**, 255-275.
- REDELMEIER, D. A., & SHAFIR, E. (1995). Medical decision making in situations that offer multiple alternatives. *Journal of the American Medical Association*, **273**, 302-305.
- SCHWARTZ, J. A., & CHAPMAN, G. B. (1999). Are more options always better? The attraction effect in physicians' decisions about medications. *Medical Decision Making*, **19**, 315-323.
- SHAFIR, E., SIMONSON, I., & TVERSKY, A. (1993). Reason-based choice. *Cognition*, **49**, 11-36.
- SIMONSON, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research*, **16**, 158-174.
- SIMONSON, I., & NYE, P. (1992). The effect of accountability on susceptibility to decision errors. *Organizational Behavior & Human Decision Processes*, **51**, 416-446.
- SLOVIC, P. (1995). The construction of preference. *American Psychologist*, **50**, 364-371.
- TVERSKY, A., & SHAFIR, E. (1992). Choice under conflict: The dynamics of deferred decision. *Psychological Science*, **3**, 358-361.

(Manuscript received April 8, 2002;  
revision accepted for publication December 4, 2002.)