Are Medical Treatments for Individuals and Groups Like Single-Play and Multiple-Play Gambles?

Michael L. DeKay,¹ John C. Hershey,², ³ Mark D. Spranca,⁴ Peter A. Ubel,⁵–⁸ and David A. Asch³,⁹

¹Department of Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, PA
²The Wharton School, University of Pennsylvania, Philadelphia, PA
³Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, PA
⁴Abt Associates, Cambridge, MA
⁵VA Health Services Research and Development Center of Excellence, VA Ann Arbor Healthcare System, Ann Arbor, MI
⁶Division of General Internal Medicine, University of Michigan, Ann Arbor, MI
⁷Center for Behavioral and Decision Sciences in Medicine, Ann Arbor, MI
⁸Department of Psychology, University of Michigan, Ann Arbor, MI
⁹Center for Health Equity Research and Promotion, Philadelphia Veterans Affairs Medical Center, Philadelphia, PA

Correspondence and reprint requests should be addressed to:

 Michael L. DeKay
 Department of Engineering and Public Policy
 Carnegie Mellon University
 Pittsburgh, PA 15213-3890
 Telephone: 412-268-1877
 Facsimile: 412-268-3757
 E-Mail: dekay@andrew.cmu.edu

Running Head: Medical Treatments for Individuals and Groups

Manuscript under review. Please do not quote without permission.
ABSTRACT

People are often more likely to accept risky gambles with positive expected values when the gambles will be played more than once. We investigated whether this distinction between single-play and multiple-play gambles extends to medical treatments for individual patients and groups of patients. Resident physicians and medical students \((n = 69)\) and undergraduates \((n = 99)\) ranked 9 different flu shots and a no-flu-shot option in 1 of 4 combinations of perspective (individual patient vs. group of 1000 patients) and uncertainty frame (probability vs. frequency). The rank of the no-flu-shot option (a measure of preference for treatment vs. no treatment) was not significantly related to perspective or participant population. The main effect of uncertainty frame and the interaction between perspective and uncertainty frame approached significance \((0.1 < p < 0.05)\), with the no-flu-shot option faring worst (treatment faring best) when decisions about many patients were based on frequency information. Undergraduate participants believed that the no-flu-shot option would be less attractive (treatment would be more attractive) in decisions about many patients, but these intuitions were inconsistent with the actual ranks. These results and those of other studies suggest that medical treatments for individuals and groups are not analogous to single-play and multiple-play monetary gambles, perhaps because people are unwilling to aggregate treatment outcomes over multiple patients.
Are medical treatments for individuals and groups like single-play and multiple-play gambles?

SINGLE-PLAY AND MULTIPLE-PLAY GAMBLES

A convincing body of research demonstrates that people often make different choices when making *multiple-play* decisions than when making *single-play* decisions. Samuelson (1963) initiated this literature with a revealing anecdote about a lunch colleague who would reject a single gamble with an even chance of winning $200 or losing $100, but who would accept a series of 100 such gambles. Subsequently, several studies have indicated that people are more likely to accept mixed gambles (i.e., gambles involving a possible gain and a possible loss) with positive expected values (EVs) when the gambles will be played more than once (Benartzi & Thaler, 1999; DeKay & Kim, 2005; Keren, 1991; Klos, Weber, & Weber, 2005; Langer & Weber, 2001; Li, 2003; Redelmeier & Tversky, 1992; Wedell & Böckenholt, 1994), although the opposite result has been also been observed (Benartzi & Thaler, 1999; Langer & Weber, 2001).

Multiple plays of Samuelson-type gambles are particularly attractive when participants are shown the distribution of possible outcomes resulting from repeated plays (Benartzi & Thaler, 1999; DeKay & Kim, 2005; Langer & Weber, 2001; Redelmeier & Tversky, 1992).

Although the rationality of making different choices for single-play and multiple-play gambles has been debated (Lopes, 1981, 1996; Nielsen, 1985; Ross, 1999; Samuelson, 1963; Schoemaker & Hershey, 1996; Tversky & Bar-Hillel, 1983), this article is concerned primarily with the empirical distinction. Related research shows that multiple plays may also increase the attractiveness of higher-EV unmixed gambles (Montgomery & Adelbratt, 1982; but see Chen & Corter, 2006, for conflicting evidence); reduce the incidence of certainty and possibility effects.
Medical Treatments for Individuals and Groups

(Barron & Erev, 2003, Experiment 5; Keren, 1991; Keren & Wagenaar, 1987); reduce choosing/pricing preference reversals (Wedell & Böckenholt, 1990); reduce the “illusion of control” (Budescu & Bruderman, 1995; Koehler, Gibbs, & Hogarth, 1994); and facilitate the multiplicative combination of probabilities and outcomes (Joag, Mowen, & Gentry, 1990).

Taken together, these results indicate that choices and preferences are often more consistent with expected value theory and/or expected utility theory when multiple plays are considered.

MEDICAL TREATMENTS FOR INDIVIDUALS AND GROUPS

One limitation of the research cited above is that the studies have focused almost exclusively on monetary gambles or other financial decisions (e.g., Joag et al., 1990, studied industrial purchasing decisions). A few researchers have attempted to assess whether the results generalize to decisions about medical treatments. For example, Redelmeier and Tversky (1992) reported that physicians and students were more likely to recommend a risky positive-EV treatment to an individual patient with chronic knee pain when they considered repeated treatments rather than a single treatment. This finding is consistent with those for monetary gambles.

More frequently, studies have involved the treatment of multiple patients rather than multiple treatments of the same patient. Redelmeier and Tversky (1990) reported that physicians and students who considered an individual patient (the individual perspective) often made different decisions from those who considered a group of comparable patients (the group perspective). In their adverse-outcomes scenario, for example, students who considered an individual woman with a blood condition were more likely to recommend a risky positive-EV treatment than were participants who considered many women. This result and others reported by Redelmeier and Tversky (1990) appear to contradict the literature on single-play and multiple-play gambles. If treating a group of similar patients is analogous to playing a gamble multiple times, one might
predict that a risky positive-EV treatment would be viewed more favorably in the group perspective than in the individual perspective.

However, other researchers have not found significant differences between medical treatments for individuals and groups. DeKay et al. (2000) were unable to replicate Redelmeier and Tversky’s (1990) results for the adverse-outcomes scenario, despite ample statistical power. Indeed, participants were slightly more likely to recommend treatment in the group perspective when the wording of the response options was improved. DeKay and Kim (2005) also reported no significant difference between the individual and group perspectives for a closely related scenario. Hux, Levinton, and Naylor (1994) found no evidence that physicians’ willingness to prescribe a medication to an individual patient differed from their willingness to recommend the medication in a practice guideline. Finally, Spranca, Minsk, and Baron (1991) and Ritov and Baron (1990) reported nonsignificant effects of perspective for students’ evaluations of a risky medical procedure and a risky flu vaccination, respectively.¹

Understanding these results is important because medical practice guidelines frequently reflect the group perspective adopted in randomized clinical trials, decision analyses, and cost-effectiveness analyses. If people think differently about medical treatments for individuals and groups, these differences may help to explain why physicians often deviate from practice guidelines when treating individual patients (Asch & Hershey, 1995; Kosecoff et al., 1987; Lomas et al., 1989; Sackett, 1989; Sorum et al., 2003; Timmermans, Sprij, & de Bel, 1996; Woo, Woo, Cook, Weisberg, & Goldman, 1985). If not, then explanations for the discrepancy between practice guidelines and actual practice must be sought elsewhere.

¹ Spranca et al.’s (1991) and Ritov and Baron’s (1990) studies of omission bias were not about risk per se, as the action and inaction alternatives both involved risk. The studies are cited here because preferences for the option with the higher expectation (action) did not depend on whether the decision affected one patient or many patients.
STUDY: RANKING SEVERAL TREATMENT OPTIONS 
AND A NO-TREATMENT OPTION

Overview

The study described here was designed to provide additional insight into the distinction between medical decisions for individuals and groups. Although the study did not involve monetary gambles, our use of risky positive-EV treatments allows a straightforward comparison to the literature on single and multiple plays of mixed monetary gambles.

This investigation extends previous research on medical decisions for individuals and groups in three ways. First, we utilized a new task that involved ranking several treatment options (different flu shots) and a no-treatment option, with the rank of the no-treatment option serving as the primary dependent measure. This task may have been more subtle than the dichotomous-choice and single-treatment rating tasks used in previous studies (DeKay et al., 2000; DeKay & Kim, 2005; Hux et al., 1994; Redelmeier & Tversky, 1990). It was somewhat similar to the separate ratings of the treatment and no-treatment options in studies of omission bias (Ritov & Baron, 1990; Spranca et al., 1991), although there were more treatment options in this study.

Second, we included the framing of uncertainty as an additional variable. Perspective (individual vs. group) and uncertainty frame (probability vs. frequency) have occasionally been confounded in past research (e.g., Redelmeier & Tversky, 1990; Spranca et al., 1991), presumably because it is natural to describe uncertainty in terms of probabilities when considering an individual and in terms of frequencies when considering a group. This confound is potentially important because reasoning is often improved when frequencies rather than probabilities are used, although the reasons and required conditions for this performance difference are still debated (Cosmides & Tooby, 1996; Gigerenzer, 1991, 1996a, 1996b;
Gigerenzer & Hoffrage, 1995; Kahneman & Tversky, 1996; Mellers, Hertwig, & Kahneman, 2001; Tversky & Kahneman, 1983). DeKay et al. (2000) crossed perspective and uncertainty frame and found that uncertainty frame was not a significant predictor of treatment recommendations. Our design was similar, but we used absolute frequencies (e.g., 600 out of 1000 people) rather than relative frequencies (e.g., 60% of people) in this study, because relative frequencies may be treated more like probabilities (Gigerenzer & Hoffrage, 1995).

Third, we asked medical experts (resident physicians and medical students) and undergraduates to complete the same task, because previous studies have varied in their use of physician and lay participants. Redelmeier and Tversky (1990) surveyed both physicians and students (for different questions), Hux et al. (1994) surveyed physicians, DeKay et al. (2000) surveyed the general public, and the remaining studies of medical decisions for individuals and groups used student participants.

**Method**

**Design.** Perspective (individual vs. group), uncertainty frame (probability vs. frequency), and participant population (resident physicians and medical students vs. undergraduates) were crossed in a $2 \times 2 \times 2$ between-participants factorial design. Participants from each population were randomly assigned to the four versions of survey materials.

**Participants.** Fifty-eight resident physicians in internal medicine and 13 advanced medical students from the Hospital of the University of Pennsylvania received cookies in return for their participation. The mean age was 28 (range = 22–48) and 49% were female. Two residents were dropped because they did not rank all of the treatment options.

Ninety-nine undergraduates were recruited by placing signs in the University of
Pennsylvania Department of Psychology. They received $6.00 per hour for participation in various experiments. Demographic data were not collected.

_Materials and procedures._ Participants read a cover story describing “a new strain of flu that is likely to sweep the region in the next few months.” In the frequency frame, participants were told: “If no vaccine is administered, 600 out of every 1000 people in this region are expected to catch the flu. 400 out of every 1000 people are expected not to catch the flu. Unfortunately, there is no way to predict ahead of time who will catch the flu and who will not.” The story also indicated that nine new vaccines had been developed to combat this strain of flu and that these vaccines had been tested on “a large sample of patients who are very similar to your patients.” In addition to reducing the number of patients who would catch the flu, the vaccines were also said to lead to occasional “adverse reactions” that were “TWICE AS BAD as catching the flu.”

We provided participants with a shuffled deck of 10 cards describing the vaccines and the “No Flu Shot” option and asked them to rank the options from best (1) to worst (10). In the frequency frame, the cards included bar graphs depicting the “Distribution of Patient Outcomes” (i.e., the “Number of Patients” expected to experience the three possible outcomes: “Reaction,” “Flu,” and “No Flu”), along with an “Average Quality of Life” score (defined as “the mean of the distribution of patient outcomes when the worst possible outcome is given a score of 0 and the best possible outcome is given a score of 100”) and an “Outcome Variability” score (the standard deviation of that distribution). In the probability frame, the text and graphs used “Percent Chance,” “Distribution of Possible Outcomes,” “Expected Quality of Life,” and “Outcome Uncertainty” instead. Figure 1 provides two examples of the stimuli. Table 1 describes all nine vaccines and the no-flu-shot option. Note that all of the vaccines had higher average-quality-of-life scores than the no-flu-shot option, so that they might appear realistic.
In the individual and group perspectives, participants were asked to think about which of the 10 options they would recommend to their “individual patient” or their “1000 patients,” respectively. Undergraduates also indicated whether the no-flu-shot option would appear better or worse if viewed from the other perspective. For example, participants who had ranked the options in the individual perspective were asked, “Do you think the No Flu Shot option would appear better or worse if you were treating 1000 similar patients?”

Hypotheses

Based on previous research, we expected that perspective would not significantly affect treatment preferences. We also expected that the effect of uncertainty frame would be nonsignificant, based on DeKay et al.’s (2000) result. We did not have specific expectations for the effect of participant population, for interactions between the three predictors, or for undergraduates’ intuitions regarding the adoption of the alternative perspective. The study was exploratory with respect to those issues.

Results

*Rank of the no-flu-shot option.* The mean rank of the no-flu-shot option was 8.29 (where 1 = best option and 10 = worst option; see Table 1); only flu shot D was ranked worse ($M = 9.02$). In fact, 56.5% of participants ranked the no-flu-shot option as worst (see Table 2), perhaps because all of the flu shots had higher EVs. The rank of the no-flu-shot option and the percentage of participants ranking it as worst may be considered measures of participants’ relative preference.
for treatment versus no treatment in the different conditions of the study. The higher the rank and the greater the percentage, the more treatment was preferred.

We conducted a $2 \times 2 \times 2$ ANOVA for predicting the rank of the no-flu-shot option, using standard regression techniques for contrast-coded predictors and their interactions (Judd & McClelland, 1989). Results indicated a nearly significant effect of uncertainty frame, $F(1, 160) = 3.20, p = 0.076$, such that the no-flu-shot option was ranked worse (treatment was ranked better) when frequencies were used ($M = 8.60$) than when probabilities were used ($M = 7.98$; see Table 2).

There was also a nearly significant interaction between perspective and uncertainty frame, $F(1, 160) = 2.89, p = 0.092$, such that the individual – group difference was greater in the probability frame (difference = 0.86) than in the frequency frame (difference = –0.40). However, the effect of perspective was not significant in either frame, both $F$s $\leq$ 2.22, both $p$s $\geq$ 0.141.

Looking at the interaction the other way, the frequency – probability difference was greater in the group perspective (difference = 1.25) than in the individual perspective (difference = –0.01).

The simple effect of uncertainty frame was significant in the group perspective, $F(1, 80) = 6.00, p = 0.016$, but not in the individual perspective, $F < 1$, suggesting that the distinction between probabilities and frequencies was more relevant when many patients were considered.

No other main effects or interactions approached significance, all $F$s $< 1$. In particular, the effect size for the main effect of perspective, $f^2 = 0.0026$ (Cohen, Cohen, West, & Aiken, 2003, p. 94), was much smaller than that for Redelmeier and Tvesky’s (1990) adverse-outcomes
scenario, $f^2 = 0.045$, and similar to that in DeKay et al.’s (2000) exact replication of the adverse-outcomes scenario, $f^2 = 0.0035$. In this study, the power for detecting an effect as large as that reported by Redelmeier and Tversky (1990) was 0.78. The power for detecting a “medium” effect was greater than 0.99 if medium is defined as $f^2 = 0.15$ (Cohen et al., p. 95). Our observed effect was noticeably smaller than Cohen et al.’s “small” effect of $f^2 = 0.02$. When only the individual/probability cell and the group/frequency cell were considered, as in Redelmeier and Tversky (1990), the difference was not significant, $F < 1, f^2 = 0.0097$.

In sum, these analyses indicate that the distinction between the individual and group perspectives was not particularly important for this task, although perspective may have moderated the effect of uncertainty frame. However, it is possible that the nonsignificant results were caused by a floor effect involving the rank of the no-flu-shot option. To address this concern, we dropped participants who ranked that option as worst. The remaining 73 participants ranked the no-flu-shot option very similarly in the individual and group perspectives ($M = 6.14$ and $M = 5.97$, respectively). There were no significant effects in the three-way ANOVA, all $F$s $\leq$ 2.19, all $p$s $\geq$ 0.144, indicating that the original results were not due to a floor effect.²

**Within-participant regressions.** Although the mean rank of the no-flu-shot option did not vary significantly as a function of perspective, it is possible that participants in the different perspectives responded in different ways to the uncertainty frame. We conducted a logistic regression to predict the percentage of participants ranking the no-flu-shot option as worst (i.e., to assess whether there were effects of perspective or other variables on the percentage of participants choosing the “floor” for that option). When participant population and its interactions were omitted, results were the same as those for the original analysis, except that the effect of uncertainty frame was not significant, $OR = 1.30$, Wald $\chi^2 = 2.77, p = 0.274$, and the simple effect of uncertainty frame was not quite significant in the group perspective, $OR = 1.54$, Wald $\chi^2 = 3.75, p = 0.053$. As before, the distinction between the individual and group perspectives appeared unimportant in these analyses.

² An additional concern about the original (full-sample) ANOVA involved the residuals. Because the distribution of no-flu-shot ranks was skewed, the distribution of residuals was also skewed. We addressed this issue in two ways. First, we re-ran the ANOVA after rank-transforming the original data. Skew in the residuals dropped from $-1.18$ to $-0.53$. Results were the same as before, except that the effect of uncertainty frame was not significant, $F(1, 160) = 1.85, p = 0.175$. The simple effect of uncertainty frame remained significant in the group perspective, $F(1, 80) = 4.71, p = 0.033$. Second, we conducted logistic regressions to predict the percentage of participants ranking the no-flu-shot option as worst (i.e., to assess whether there were effects of perspective or other variables on the percentage of participants choosing the “floor” for that option). When participant population and its interactions were omitted, results were the same as those for the original analysis, except that the effect of uncertainty frame was not significant, $OR = 1.30$, Wald $\chi^2 = 2.77, p = 0.274$, and the simple effect of uncertainty frame was not quite significant in the group perspective, $OR = 1.54$, Wald $\chi^2 = 3.75, p = 0.053$. As before, the distinction between the individual and group perspectives appeared unimportant in these analyses.
conditions weighted other information (e.g., the chance of an adverse reaction) differently when ranking the options, and that the no-flu-shot option was rated as systematically better or worse than might be expected on the basis of that information. To assess this possibility, we conducted a series of within-participant regressions. Two models (2 and 4) included a dummy code for the no-flu-shot option, whereas the other two (1 and 3) did not.

In model 1, we regressed the ranks of the 10 options onto the percentage of patients expected to experience adverse reactions and the percentage expected to experience neither adverse reactions nor the flu (in the individual perspective, we used the percent chance of these outcomes). The mean unstandardized regression coefficients appear in Table 3. As expected, participants gave higher (worse) ranks to options with more adverse reactions, and lower (better) ranks to options with more no-flu outcomes, both $|t|\geq 25.07$, both $p < 0.0001$. In model 3, we used average quality of life (or expected quality of life) and outcome variability (or outcome uncertainty) as predictors. As expected, participants gave lower (better) ranks to options with higher average quality of life and higher (worse) ranks to options with higher outcome variability, both $|t|\geq 8.40$, both $p < 0.0001$.

In models 2 and 4, we added a dummy code for the no-flu-shot option to models 1 and 3, respectively. The coefficient for the dummy code was not significantly different from zero in either case, $t(167) = -1.06$, $p = 0.293$ in model 2 and $t(167) = -0.08$, $p = 0.933$ in model 4. The fact that the no-flu-shot option was not given special standing suggests that status-quo bias (Samuelson & Zeckhauser, 1988) and omission bias (Baron, 1992; Ritov & Baron, 1990, 1992; Spranca et al., 1991) were relatively unimportant in the ranking task.
To assess whether information use varied across conditions, we used the coefficients for the predictor variables in models 2 and 4 as the dependent variables in a series of 2 (perspective) × 2 (uncertainty frame) ANOVAs. None of the main effects or interactions was significant, all $F$s $\leq$ 1.51, all $p$s $\geq$ 0.222, suggesting that reasoning was similar across conditions. Results were very similar when we dropped those participants who ranked the no-flu-shot alternative as worst.

*Undergraduates’ intuitions about different numbers of patients.* When asked whether the no-flu-shot option would appear better or worse if viewed from the other perspective (e.g., if they were treating 1000 patients instead of just one), 73% of undergraduates who responded indicated that it would. We used ordinal logistic regression to predict whether the no-flu-shot option would appear worse than, the same as, or better than it had in the original perspective, using perspective change, uncertainty frame, and their interaction as predictors. Responses were significantly related to perspective change, $OR = 1.82$, Wald $\chi^2 = 8.78$, $p = 0.003$, but not to uncertainty frame, $OR = 1.09$, Wald $\chi^2 = 0.18$, $p = 0.674$ (see Table 4). Participants who originally considered one patient said that the no-flu-shot option would appear worse (treatment would appear better) if they considered 1000 patients, $S = -144.5$ for the Wilcoxon signed rank test, $p = 0.002$. Participants who originally considered 1000 patients said that the no-flu-shot option would appear better (treatment would appear worse) if they considered one patient, but this trend was not significant, $S = 70$, $p = 0.174$.

The main effect of perspective change was qualified by a nearly significant interaction with uncertainty frame, $OR = 1.39$, Wald $\chi^2 = 2.73$, $p = 0.098$, with the above effects being stronger in
the frequency frame than in the probability frame. The simple effect of perspective change was significant for frequencies, $OR = 2.56$, Wald $\chi^2 = 9.83$, $p = 0.002$, but not for probabilities, $OR = 1.31$, Wald $\chi^2 = 0.93$, $p = 0.335$ (see Table 4). Viewing the interaction the other way, the effect of uncertainty frame was not significant for either shift of perspective, both $ps \geq 0.183$.

Undergraduates’ intuitions that treatment would be more attractive in the group perspective were consistent with the greater appeal of risky monetary gambles in multiple-play situations, but these intuitions were not borne out by the actual ranks of the no-flu-shot option (see Table 2). Interestingly, however, the interaction reported above was in the same direction as that in the analysis of actual ranks. Although the simple effects were somewhat different in the two analyses, the no-flu-shot option fared worst (treatment fared best) when frequency information was combined with the group perspective (or a shift to the group perspective).

**DISCUSSION**

In this study, preferences for treatment options were very similar in the individual and group perspectives. This result and those of other studies (DeKay et al., 2000; DeKay & Kim, 2005; Hux et al., 1994; Ritov & Baron, 1990; Spranca et al., 1991) conflict with Redelmeier and Tversky’s (1990) finding that treatment is more likely to be preferred for individuals than for groups. Our data help to eliminate differences in uncertainty frames and participant populations as explanations for this discrepancy in the literature. More important, all of these studies (including Redelmeier and Tversky, 1990) suggest that the relatively robust distinction between single-play and repeated-play monetary gambles does not extend to medical treatments for individuals and groups.

One promising explanation for this result is that people are willing to aggregate monetary outcomes over multiple plays (e.g., to think of five gains of $200 and five losses of $100 as a net...
gain of $500), but unwilling to aggregate outcomes of medical treatments over multiple patients (e.g., to think of five patients each gaining 10 years of life and five other patients each losing 2 years of life as a net gain of 40 years). For multiple-play gambles in which one person may win money on some plays and lose money on others, it is reasonable (even normative) to consider the distribution of aggregate outcomes (Benartzi & Thaler, 1999; Kahneman & Lovallo, 1993; Read, Loewenstein, & Rabin, 1990; Redelmeier & Tversky, 1992). Risky positive-EV gambles are often more appealing when possible outcomes are aggregated over multiple plays, perhaps because repetition reduces the probability of losing money (sometimes to near zero). However, when medical treatments for multiple patients are considered, aggregation may be inappropriate because the gains and losses experienced by different patients do not offset each other in any real sense (Asch, 1990; Asch & Hershey, 1995). If people are reluctant to aggregate outcomes over patients, then the analogy between monetary gambles and medical treatments breaks down, and treatments are likely to be evaluated similarly in the individual and group perspectives.

Evidence for this explanation comes from two sources. First, Redelmeier and Tversky’s (1992) result for multiple treatments of one patient paralleled results for repeated-play monetary gambles rather than those for the treatment of multiple patients, suggesting that participants were willing to aggregate medical outcomes experienced by an individual. Second, DeKay and Kim (2005; DeKay, Kim, & Tuma, 2003) reported that the perceived fungibility of outcomes over multiple plays (i.e., the appropriateness of aggregating outcomes over plays) was lower for risky medical treatments involving multiple patients—including versions of Redelmeier and Tversky’s (1990) adverse-outcomes scenario—than for multiple plays of risky monetary gambles involving a single person or firm. Moreover, the increased attractiveness of repeated plays relative to a single play (the standard result for monetary gambles) was lower in situations with less fungible
outcomes, even though probabilities and relative gains and losses were equated across situations. Thus, if most participants in this study were unwilling to aggregate gains and losses over patients, one would expect little or no difference between the individual and group perspectives.

In addition to our primary result, we observed a nearly significant interaction between perspective and uncertainty frame, with a significant simple effect of uncertainty frame in the group perspective only. Apparently, expressing uncertainty in terms of frequencies rather than probabilities led participants in the group condition to view treatment more favorably. One possible explanation is that the use of frequencies facilitated participants’ recognition that for each flu-shot option, more patients would be spared the flu than would experience adverse reactions (i.e., there would be a net increase in aggregate health, relative to the no-flu-shot option). This realization may have seemed more relevant to participants considering a decision for many patients than to those considering a decision for only one patient, assuming that at least some participants were willing to aggregate outcomes over patients.

Although the distinction between the individual and group perspectives was not significant for actual rankings, undergraduate participants expressed the belief that treatment would be evaluated more favorably for many patients than for one patient. It is not clear whether these intuitions were simply off the mark (i.e., “folk theory” did not match reality), or whether they represented underlying tendencies that were too weak to compete with other considerations in the ranking task. One possibility is that the two tasks (comparing many options in one perspective vs. comparing one option in two perspectives) focused participants’ attention on different aspects of the situation, just as different evaluation modes lead to preference reversals in other contexts (e.g., Hsee, Loewenstein, Blount, & Bazerman, 1999; Tversky, Sattath, &

---

3 In the frequency format, a net improvement was essentially guaranteed because no uncertainty was reported for the numbers of patients experiencing each outcome (see Figure 1). In the probability frame, a net improvement may have appeared less certain because participants themselves would have had to aggregate outcomes over patients.

Manuscript under review. Please do not quote without permission.
In contrast to this difference for the main effect of perspective, the nearly significant interactions between perspective and uncertainty frame were somewhat similar in the two tasks: treatment fared best when frequency information was coupled with decisions about many patients. Perhaps there was something to the undergraduates’ intuitions after all.

In summary, accumulating evidence indicates that the distinction between single and multiple plays of risky monetary gambles does not extend to risky medical treatments for individuals and groups, perhaps because people are reluctant to aggregate the results of medical treatments over patients. The intriguing intuitions of our undergraduate participants and the nearly significant interactions between perspective and uncertainty frame qualify this conclusion only slightly. As a practical matter, researchers interested in understanding discrepancies between clinical guidelines and the treatment of individual patients may wish to consider alternative explanations.

ACKNOWLEDGEMENTS

We are grateful to Christine Weeks for assistance with data collection and entry, and to Jon Baron, Tehila Kogut, and two anonymous reviewers for helpful comments on earlier versions of this article. Preparation of this article was supported by National Science Foundation grant SES-0218318 to the first author.

REFERENCES


Gigerenzer, G. (1996b). The psychology of good judgment: Frequency formats and simple


Manuscript under review. Please do not quote without permission.


Manuscript under review. Please do not quote without permission.


Manuscript under review. Please do not quote without permission.


TABLE 1
Attributes (Frequency Versions) and Mean Ranks of Treatment Options, Including the No-Flu-Shot Option

<table>
<thead>
<tr>
<th>Optiona</th>
<th>Adverse Reaction</th>
<th>Flu</th>
<th>No Flu</th>
<th>Average Quality of Life</th>
<th>Outcome Uncertainty</th>
<th>Mean Rankb</th>
</tr>
</thead>
<tbody>
<tr>
<td>No flu shot</td>
<td>0</td>
<td>600</td>
<td>400</td>
<td>70.0</td>
<td>24.5</td>
<td>8.29 (0.19)</td>
</tr>
<tr>
<td>Flu shot F</td>
<td>50</td>
<td>150</td>
<td>800</td>
<td>87.5</td>
<td>26.8</td>
<td>1.25 (0.07)</td>
</tr>
<tr>
<td>Flu shot A</td>
<td>50</td>
<td>250</td>
<td>700</td>
<td>82.5</td>
<td>28.6</td>
<td>3.38 (0.11)</td>
</tr>
<tr>
<td>Flu shot B</td>
<td>50</td>
<td>350</td>
<td>600</td>
<td>77.5</td>
<td>29.5</td>
<td>5.90 (0.14)</td>
</tr>
<tr>
<td>Flu shot G</td>
<td>100</td>
<td>100</td>
<td>800</td>
<td>85.0</td>
<td>32.0</td>
<td>2.72 (0.11)</td>
</tr>
<tr>
<td>Flu shot E</td>
<td>100</td>
<td>200</td>
<td>700</td>
<td>80.0</td>
<td>33.2</td>
<td>5.02 (0.08)</td>
</tr>
<tr>
<td>Flu shot I</td>
<td>100</td>
<td>300</td>
<td>600</td>
<td>75.0</td>
<td>33.5</td>
<td>7.61 (0.09)</td>
</tr>
<tr>
<td>Flu shot C</td>
<td>150</td>
<td>50</td>
<td>800</td>
<td>82.5</td>
<td>36.3</td>
<td>4.90 (0.17)</td>
</tr>
<tr>
<td>Flu shot H</td>
<td>150</td>
<td>150</td>
<td>700</td>
<td>77.5</td>
<td>37.0</td>
<td>6.90 (0.12)</td>
</tr>
<tr>
<td>Flu shot D</td>
<td>150</td>
<td>250</td>
<td>600</td>
<td>72.5</td>
<td>37.0</td>
<td>9.02 (0.09)</td>
</tr>
</tbody>
</table>

a Flu-shot options were assigned random letters for identification purposes.

b Options that were viewed more favorably have lower ranks (1 = best option, 10 = worst option). Standard errors are in parentheses.
### TABLE 2
Mean Ranks of the No-Flu-Shot Option and Percentages of Participants Ranking the No-Flu-Shot Option as Worst

<table>
<thead>
<tr>
<th>Perspective</th>
<th>Participants</th>
<th>Uncertainty Frame</th>
<th>Individual</th>
<th>Group</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>M</td>
<td>%</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>Physicians</td>
<td>Probability</td>
<td>8.33</td>
<td>66.7</td>
<td>7.29</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.66)</td>
<td>(11.1)</td>
<td>(0.78)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Frequency</td>
<td>8.56</td>
<td>68.8</td>
<td>9.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.58)</td>
<td>(11.6)</td>
<td>(0.40)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>8.44</td>
<td>67.6</td>
<td>8.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.44)</td>
<td>(8.0)</td>
<td>(0.45)</td>
</tr>
<tr>
<td></td>
<td>Undergraduates</td>
<td>Probability</td>
<td>8.44</td>
<td>56.0</td>
<td>7.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.45)</td>
<td>(9.9)</td>
<td>(0.53)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Frequency</td>
<td>8.28</td>
<td>48.0</td>
<td>8.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.45)</td>
<td>(10.0)</td>
<td>(0.41)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>8.36</td>
<td>52.0</td>
<td>8.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.32)</td>
<td>(7.1)</td>
<td>(0.34)</td>
</tr>
<tr>
<td></td>
<td>All participants</td>
<td>Probability</td>
<td>8.40</td>
<td>60.5</td>
<td>7.54</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.38)</td>
<td>(7.5)</td>
<td>(0.44)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Frequency</td>
<td>8.39</td>
<td>56.1</td>
<td>8.79</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.35)</td>
<td>(7.8)</td>
<td>(0.29)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>8.39</td>
<td>58.3</td>
<td>8.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.26)</td>
<td>(5.4)</td>
<td>(0.27)</td>
</tr>
</tbody>
</table>

*Note.* Options that were viewed more favorably have lower ranks (1 = best option, 10 = worst option). Higher means and percentages imply that the no-flu-shot option was viewed less favorably, and that treatment was viewed more favorably. Standard errors are in parentheses.
TABLE 3  
Mean Unstandardized Regression Coefficients from Within-Participant Regressions for Predicting Ranks of Treatment Options, Including the No-Flu-Shot Option

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept(^a)</td>
<td>17.305****</td>
<td>17.713****</td>
<td>34.861****</td>
<td>34.983****</td>
</tr>
<tr>
<td></td>
<td>(0.392)</td>
<td>(0.634)</td>
<td>(1.320)</td>
<td>(2.343)</td>
</tr>
<tr>
<td>Percentage with adverse reaction (worst outcome)</td>
<td>0.349****</td>
<td>0.343****</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.017)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage with no flu (best outcome)</td>
<td>-0.223****</td>
<td>-0.228****</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.008)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average quality of life</td>
<td></td>
<td>-0.432****</td>
<td>-0.433****</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.012)</td>
<td>(0.019)</td>
<td></td>
</tr>
<tr>
<td>Outcome variability</td>
<td></td>
<td>0.149****</td>
<td>0.147****</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.018)</td>
<td>(0.028)</td>
<td></td>
</tr>
<tr>
<td>No-flu-shot option (dummy)</td>
<td></td>
<td>-0.312</td>
<td>-0.028</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.296)</td>
<td>(0.340)</td>
<td></td>
</tr>
<tr>
<td>Mean adjusted (R^2)</td>
<td>0.847</td>
<td>0.903</td>
<td>0.836</td>
<td>0.898</td>
</tr>
</tbody>
</table>

*Note.* Options that were viewed more favorably have lower ranks (1 = best option, 10 = worst option). For predictor variables, a positive coefficient means that higher values of the predictor variable are associated with options that were ranked higher (worse). A negative coefficient means that higher values of the predictor variable are associated with options that were ranked lower (better). Standard errors of mean regression coefficients (not mean standard errors) are in parentheses.

\(^a\)Intercepts are not meaningful because no options had values of zero on all predictor variables.

**** \(p < 0.0001\) for a \(t\) test of the hypothesis that the mean regression coefficient equals zero.
TABLE 4
Numbers of Undergraduates Reporting that the No-Flu-Shot Option Would Appear Worse, the Same, or Better if Viewed from the Other Perspective

<table>
<thead>
<tr>
<th>Uncertainty Frame</th>
<th>Change in Perspective (Original ➞ New)</th>
<th>Evaluation of the No-Flu-Shot Option in the New Perspective</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Worse than in Original Perspective</td>
</tr>
<tr>
<td>Probability</td>
<td>Individual ➞ group</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Group ➞ individual</td>
<td>8</td>
</tr>
<tr>
<td>Frequency</td>
<td>Individual ➞ group</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Group ➞ individual</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>Individual ➞ group</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Group ➞ individual</td>
<td>13</td>
</tr>
</tbody>
</table>

Manuscript under review. Please do not quote without permission.
Figure 1. Examples of stimuli that appeared on cards. The no-flu-shot option is shown in the probability frame and flu shot F (the best of the flu shots) is shown in the frequency frame.