Estimation as a Catalyst for Numeracy: Micro-Interventions that Increase the Use of Numerical Information in Decision-Making

Luke F. Rinne, Michael A. Ranney
Graduate School of Education, University of California, Berkeley, 4533 Tolman Hall, Berkeley, CA, 94720-1670
Nicholas H. Lurie
College of Management, Georgia Institute of Technology, 800 West Peachtree St. NW, Atlanta, GA, 30308-0520
Email: lrinne@berkeley.edu, ranney@berkeley.edu, lurie@gatech.edu

Abstract: The utilization of available numerical information represents an important aspect of numeracy. Accordingly, we seek to develop decision-making scaffolds that increase the incorporation of numerical information into preferences. Two experiments contrast five techniques for presenting students with numerical information and prompting them to make decisions related to that information. The techniques are based on a method called EPIC (Estimate, Prefer, Incorporate feedback, and Change preference) from the Numerically-Driven Inferencing (NDI) paradigm. Results regarding post-feedback preferences indicate that having students estimate a quantity's value beforehand increases the impact on preferences when the quantity's true value is presented. However, having students state initial preferences does not yield an analogous effect. An attractive explanation for the estimation effect is that estimating diminishes hindsight bias, making information that diverges from expectations more surprising. Variants of the EPIC method may serve as "micro-interventions" that increase the impact of numerical data on preferences.

Introduction

People often fail to fully utilize available numerical information when making decisions (e.g., Bartol, Koehl, & Martin, 1987). The utilization of numerical information may be seen as an aspect of "numeracy," an idea that has been advanced most famously by Paulos (1990) as a parallel in the quantitative realm for what is termed literacy in the verbal realm. Paulos shows how innumeracy often leads to poor decisions by consumers, employees, and citizens. He believes the main cause of innumeracy to be a general lack of mathematical ability, and indeed, in the cases he describes, the best remedy may be to improve people's understanding of mathematical concepts. However, Viswanathan (1993) argues that the notion of numeracy should include not only one's *ability* to use numerical information, but also one's *proclivity* toward using numerical information. That is, he holds that two individuals with comparable mathematical abilities may often differ in the extent to which they use the numbers they have learned to help them make decisions. Viswanathan supported this claim with several studies demonstrating the validity of a measure he developed called "Preference for Numerical Information" (PNI), which is meant to capture individual differences regarding that aspect of numeracy that is not subsumed by mathematical ability.

We embrace Viswanathan's (1993) broader notion of numeracy—according to which the use of numerical information is thought to depend on more than mathematical competence. Consequently, we believe that it may be possible to increase students' use of numerical information in decision-making by means other than improving their mathematics skills. In this paper, we argue that this can be achieved through "micro-interventions" based on methods from the Numerically-Driven Inferencing (NDI) paradigm (e.g., Munnich, Ranney, & Appel, 2004; Munnich, Ranney, Nelson, Garcia de Osuna, & Brazil, 2003; Ranney, Cheng, Garcia de Osuna, & Nelson, 2001). By "micro-interventions," we mean brief activities conducted by teachers in which the presentation of numerical information is preceded and/or followed by particular exercises that scaffold students' decision-making procedures.

Work on preference formation in behavioral decision research supports the promise of a "micro-intervention" approach. Several decision theorists have argued that preferences—which are taken to be what cause individuals to choose in particular ways—should not necessarily be seen as previously-existing, stable features. Rather, in novel situations, preferences are better conceived of as "constructed" in the process of elicitation based on the accessibility of various pieces of information and the ease with which information may be evaluated (Hoeffler & Ariely, 1999; Hsee, 1996; Payne, Bettman, & Johnson, 1992; Slovic, 1995). This contradicts neoclassical economic theory, which assumes that decisions do not depend on the processes through which they are made and that individuals' preferences are simply "revealed" by their decisions (Rabin, 1998). On the constructed preference view, however, prompts for novel decisions do more than simply tap existing preferences; rather, the need to make a novel decision often leads to preferences being formed in the first place. Further, because the prompt itself serves as an additional source of information for the decision-maker, its format may significantly influence the preference that is constructed. This has important consequences, because once preferences in particular domains have been constructed, they may be retained in memory and influence subsequent decisions in the same—or related—domains.

Following the constructed preferences view, it seems likely that when one receives numerical information in a domain with which one has little experience, that information's influence on preferences will be heightened if one is prompted to make a relevant decision right after receiving the information. One will construct a preference in order to reach a decision, and at this stage the numerical information provided will still be highly salient, increasing its incorporation into the preference formation process relative to what would have occurred had much time (e.g., months or years) elapsed between the receipt of information and the first instance in which one was prompted to make a decision for which the information would be relevant. Thus, given that our goal is to increase the impact of numerical information on preferences, it seems clear that our micro-interventions should include prompts for decisions given soon after numerical information is presented. But how else might it be possible to increase the impact of presenting numbers to students? In this study, we investigate whether certain NDI methods—specifically, those that include prompts for estimates and/or initial preferences *before* numerical data are presented—may *further* increase the influence of numerical information. For this purpose, post-feedback decisions, in addition to being part of the intervention, provide a measure of such potential effects.

It will be helpful to review prior research in the NDI paradigm at this point. An initial goal of NDI research was to assess how views on social issues (e.g., abortion or immigration policy) are influenced by relevant pieces of numerical information (e.g., rates of abortion or immigration). To this end, a method called EPIC was developed, which involves prompting people to: 1) **E**stimate the value of a quantity, 2) state a **P**reference related to that quantity, 3) **I**ncorporate feedback in the form of the estimated quantity's true value, and 4) **C**hange their preference (if they wish) in response to this feedback. The EPIC method allows for assessments both of prior beliefs and of the influence of numerical feedback on preferences. Early NDI experiments showed that offering a single, surprising numerical datum can indeed greatly influence preferences, and that changes between initial and final preferences increase with estimate inaccuracy (Garcia de Osuna et al., 2004; Munnich et al., 2003; Ranney et al., 2001).

Recently, Lurie and Ranney (submitted) conducted NDI experiments to study the impact of estimation on preferences and preference change. Preferences elicited using the EPIC method were compared to those obtained using two variants—called PEIC and IC—in which, respectively, the first two steps, **E**stimate and **P**refer, were either reversed in order or left out altogether. (Although no initial preferences were elicited in the IC condition, "C" for "Change preference" is retained for simplicity.) Each group used one of the three procedures—EPIC, PEIC, or IC—throughout, so that the preferences produced by each could be contrasted. In one experiment, estimates and feedback were for the number of mortalities per 100,000 deaths for each of a pair of diseases: either (a) lung cancer and skin cancer, or (b) heart disease and breast cancer (among women only for the latter pair). Preferences were elicited by asking students how they would divide $100 in research funding between prevention efforts for the two diseases in each pair (funds were dedicated to women only for both heart disease and breast cancer). By estimating mortality rates for each disease in a pair, students implicitly estimated the *ratio* of mortalities from one disease to mortalities from the other disease. It was thought that prior beliefs about these ratios would affect preferences for $100 splits, and that in some cases, but not others, these beliefs would deviate considerably from the true values later provided as feedback. The true values—based on Centers for Disease Control and Prevention ("CDC") statistics from 2001—are 8.2:1 and 20.9:1, respectively, for heart disease/breast cancer and lung cancer/skin cancer. While the true ratio for the heart disease/breast cancer pair is 3.2 times higher than the average estimate of 2.6:1, the true ratio for the lung cancer/skin cancer pair is only 1.3 times higher than the average estimate of 15.6:1.

When participants decided on initial allocations before performing any other activities (as in the PEIC condition), they allocated an average of $44.92 for heart disease (vs. breast cancer) and $48.04 for lung cancer (vs. skin cancer). However, when participants estimated mortality rates before stating initial preferences (as in the EPIC condition), average allocations for heart disease were lower—$38.40—while average lung cancer allocations were higher—$57.34. The conclusion was that estimating mortality rates increased the extent to which participants based their decisions on their prior beliefs about mortality ratios, instead of on other reasons (e.g., family members having a disease, concerns about personal responsibility for getting a disease) that would have motivated higher allocations for heart disease and skin cancer. Because estimates of mortality ratios for the lung cancer/skin cancer pair were relatively accurate, EPIC participants' initial preferences (M = $57.34) were closer than PEIC participants' initial preferences (M = \$48.04) to IC participants' preferences (M = \$59.93), which were elicited after simply giving participants the true mortality rate data. Because mortality ratio estimates for the heart disease/breast cancer pair were relatively inaccurate, in this case the effect was reversed; EPIC participants' initial preferences (M = \$38.40) were further away from IC participants' preferences (M = \$55.14) than were PEIC initial preferences (M = \$44.92).

As for final allocations, Lurie and Ranney (submitted) found that in both the EPIC and PEIC conditions, change between initial and final preferences was greater when estimates were inaccurate; the conclusion here was that preference change was driven by how "shocking" the information received was. Interestingly, although differences between groups regarding *post*-feedback allocations were not predicted, such differences emerged nonetheless. Mean post-feedback allocations for the bigger killer in each pair were higher in both the EPIC and PEIC conditions than in the IC condition (lung cancer: $M_{EPIC/PEIC}$ = \$67.54 vs. M_{IC} = \$57.08; heart disease: $M_{EPIC/PEIC}$ = \$64.62 vs. M_{IC} = \$54.68). Because their study was rather tangential to issues of information utilization, Lurie and Ranney did not scrutinize this result. However, such results are critical in the context of our current investigation; it seems that estimating and/or stating initial preferences *before* receiving feedback may have increased the influence of mortality rates *after* students learned the true values. We hypothesized that one of these activities (or perhaps the combination of the two) increased the influence of the actual values on final allocations. This spawned the idea that we might develop instructional techniques involving these activities with the aim of increasing students' incorporation of newly received numerical information into their preferences—beyond what can be achieved by prompting for decisions after information is presented. The present study represents the fruition of this idea.

The remainder of this paper will proceed as follows: First, we present our hypotheses. Second, we present new analyses of data from the aforementioned experiment by Lurie and Ranney (submitted), including some data they collected but did not use—students' written reasons explaining their preferences. Third, we present the results of a new experiment we conducted in order to 1) replicate Lurie and Ranney's results, and 2) tease apart estimation and initial preference effects. The fourth section offers a general discussion, including a proposed explanation of our findings. Finally, we suggest educational implications and directions for future research.

Hypotheses

Hypothesis 1 is that post-feedback allocations to the disease in each pair that kills more people will be significantly higher for EPIC and PEIC students than for IC students. Hypothesis 2 is that such results are due to the increased utilization of numerical information. This can be expressed in terms of two sub-hypotheses, such that support for both would demonstrate support for Hypothesis 2: The first (2a) is that written rationales provided by students in the EPIC and PEIC conditions will cite the mortality rate information provided as explaining their final allocations more frequently than will students in the IC condition. The second sub-hypothesis (2b) is that students who cite mortality rate information as explaining their final allocations will allocate more money to the bigger killer in each pair than will students who do not cite mortality rate information as explaining their allocations. In short, if both of these sub-hypotheses were supported, this would indicate that effects of estimation and/or initial allocation on final allocations are linked to the presence of written rationales that cite the numerical information provided as a reason for the final allocation decision. Hypotheses 1 and 2 are addressed in Experiment 1.

The next three hypotheses correspond to alternative explanations for why post-feedback preferences for the disease in each pair that kills more people might be higher in the EPIC and PEIC conditions than in the IC condition. The first possibility is that (a) there is an independent effect of estimation that increases final allocations, but no similar effect for stating initial preferences. The second possibility is the reverse: (b) there is an independent effect of stating initial preferences that increases final allocations, but no similar effect for estimation. Finally, the third possibility is that (c) increases in final allocations arise due to a joint effect of estimating and stating initial preferences, which means that one or both of the following are true: 1) there are independent, additive, estimation and initial preference effects, both of which increase final allocations, or 2) there is an interaction between estimating and stating initial preferences that increases final allocations. Although these hypotheses do not exhaust all possibilities (e.g., there could be effects of the length of the procedure, the order of estimation and initial preference, etc.), each of these options is mutually exclusive of the others, and each, if manifested, seems unlikely to be confounded with other possible effects. These hypotheses, which will be referred to as the "estimate effect," "initial preference effect," and "joint effect" hypotheses, are addressed in Experiment 2.

Experiment 1

Method

Undergraduates (N = 208) were randomly assigned to one of three groups: EPIC, PEIC, or IC. EPIC participants first estimated the annual number of mortalities, per 100,000 deaths, for each of a pair of diseases, either a) lung cancer and skin cancer, or b) heart disease and breast cancer. After estimating, participants were asked to state how they would prefer that $100 be divided between respective efforts for prevention of the two diseases, and to provide a written reason explaining their preferred allocation. Participants were instructed that estimates for the heart disease/breast cancer pair were to be restricted to the female population, and that all funds allocated for heart disease or breast cancer would be devoted exclusively to prevention among women. Lung cancer and skin cancer prompts were not gender specific. Next, participants received true mortality rates from the CDC for each disease in the pair, also per 100,000 deaths (nb. female deaths for the heart disease/breast cancer pair). After learning the true rates, participants were again asked to divide $100 between respective prevention efforts and to provide a written reason explaining their allocation. This procedure was then repeated for the second disease pair. Both the order in which the pairs were presented and the order of the diseases within each pair were counterbalanced. For PEIC participants, the first two steps of the EPIC procedure were reversed in order, and for IC participants, they were left out altogether. Apart from these changes, the procedures for the PEIC and IC groups were the same as for the EPIC group. The experiment also involved data collection and another manipulation that are not relevant here (see Lurie & Ranney, submitted, for details). There is no indication that these affected the data analyzed in this study.

Written reasons for final preferences were coded to identify instances in which students indicated that mortality rates for either disease in a given pair helped to explain their final allocations. These were coded as "numerical" (NUM), although they may have been attended by reasons of other types as well. Responses of other types were coded "non-numerical" (NON). Cases in which no rationale was given (< 10%) were excluded.

Results

Preferences for $100 splits are given in terms of the disease in each pair that kills more people. EPIC and PEIC students were pooled together to increase statistical power. Planned comparisons assessed Hypothesis 1; final allocations from the pooled group of EPIC and PEIC students were significantly higher than those from the IC group for lung cancer ($M_{\mathrm{EPI\underline{C}}}$ = $69.72 & $M_{\mathrm{PEI\underline{C}}}$ = $63.98, vs. $M_{\mathrm{I\underline{C}}}$ = $57.08), $t(183) = 2.617$, $p < .01$ and heart disease ($M_{\mathrm{EPI\underline{C}}}$ = $67.93 & $M_{\mathrm{PEI\underline{C}}}$ = $60.66 vs. $M_{\mathrm{I\underline{C}}}$ = $54.68), $t(195) = 2.979$, $p < .01$.

To analyze the effect of both estimating and stating initial preferences on the use of numerical reasons for final allocations (Hypothesis 2a), EPIC and PEIC responses were again pooled together and compared to IC responses (see Table 1). Note that because of this pooling, the null hypothesis predicts that within each column, the EPIC/PEIC cell count will be roughly twice the IC cell count. NUM/NON percentages are given only to aid the reader. Chi-Square tests indicate significant positive associations between membership in the EPIC/PEIC group (vs. the IC group) and the use of numerical reasons (vs. non-numerical reasons) for both disease pairs: lung cancer/skin cancer, $\chi^2(1) = 3.96$, $p < .05$, and heart disease/breast cancer, $\chi^2(1) = 5.67$, $p < .05$.

<u>Table 1. Counts of Reason-Types for Final Allocations (C)—by Group, Disease Pair, and Rationale Type</u>

	Lung Cancer/Skin Cancer		Heart Disease/Breast Cancer	
	NUM	NON	NUM	NON
EPI<u>C</u>/PEI<u>C</u>	79 (64%)	44 (36%)	85 (71%)	35 (29%)
I<u>C</u>	32 (49%)	33 (51%)	37 (54%)	32 (46%)

But did participants who gave numerical reasons ultimately allocate more money during the final preference phase for research on diseases that kill more people (Hypothesis 2b)? Since different groups gave numerical reasons at different rates, within-group analyses were conducted to control for effects of experimental group that may not have been captured by the mere presence or absence of a particular type of written reason (see Table 2). Within each disease pair/condition, stating a numerical reason was indeed significantly related to allocating more money to the bigger killer (lung cancer: EPIC—M_{NUM} = $76.86, M_{NON} = $59.14, $t(70) = 3.979$, $p < .01$, PEIC—M_{NUM} = $77.76, M_{NON} = $43.68, $t(46) = 6.684$, $p < .01$, and IC—M_{NUM} = $69.70, M_{NON} = $44.06, $t(63) = 4.146$, $p < .01$; heart disease: EPIC—M_{NUM} =$75.04, M_{NON} = $49.21, $t(67) = 5.904$, $p < .01$, PEIC—M_{NUM} = $68.14, M_{NON} = $50.00, $t(57) = 3.682$, $p < .01$, IC—M_{NUM} = $68.69, M_{NON} = $38.91, $t(66) = 6.406$, $p < .01$).

<u>Table 2. Mean Final Allocations (C), by Group, Disease Pair, and Rationale Type</u>

	Lung Cancer (vs. Skin Cancer)		Heart Disease (vs. Breast Cancer)	
	NUM	NON	NUM	NON
EPIC	$76.86	$59.14	$75.04	$49.21
PEIC	$77.76	$43.68	$68.14	$50.00
IC	$69.70	$44.06	$68.69	$38.91

Note that average lung cancer and heart disease final allocations for students giving non-numerical reasons were sometimes strikingly low, particularly considering that these diseases kill many more people than do the diseases they were paired with. Written rationales indicate that lower final preferences for the bigger killer often stem from concerns about tragedy (e.g., breast cancer) or culpability for contracting the disease (e.g., lung cancer).

Discussion

Results from Experiment 1 indicate: 1) the robustness of the effect that estimating and stating initial preferences has on final allocations for the bigger killer in each pair (Hypothesis 1), and 2) that this effect appears to occur due to an increased emphasis on the numerical information provided (Hypothesis 2). However, the results of this experiment offer little help in explaining *why* estimating and stating initial preferences increased students' utilization of the numerical information they received. Because both estimates and initial preferences were elicited in both the EPIC and PEIC conditions, it is unclear whether estimating, stating initial preferences, or some joint effect of doing both produced the differences that were observed. Our next experiment was designed to clarify matters by separating the effects of estimating and stating initial preferences from one another, as well as from potential joint effects. Experiment 2 also served to replicate some of the effects that had been observed to date.

Experiment 2

Method

In order to discriminate among the "estimate effect," "initial preference effect," and "joint effect" hypotheses, we created two new NDI methods, EIC and PIC, which exclude the initial preference and estimation phases, respectively. Thus, we may separate estimation and initial preference effects from one another and from effects of combining the two. Undergraduates (N = 153) were randomly assigned to one of five groups: EPIC, PEIC, EIC, PIC, or IC. Apart from the addition of the EIC and PIC groups, the design, materials, and prompts at each stage (Estimate, Prefer, Incorporate feedback, and Change preference) were the same as in Experiment 1.

Results

If the "estimation effect" hypothesis were true, then one would expect EPIC, PEIC, and EIC students' final allocations for the bigger killer in each pair to be higher than those of PIC and IC students. Alternatively, if the "initial preference effect" hypothesis were true, one would expect EPIC, PEIC, and PIC students' final allocations for the bigger killers to be higher than those of EIC and IC students. Finally, if the "joint effect" hypothesis were true, one would expect EPIC and PEIC students' final allocations for the bigger killer in each pair to be higher than those of EIC, PIC, and IC students. Note that these three alternatives are mutually exclusive.

Generally, this experiment produced significant results that adjudicate among the alternative hypotheses for the heart disease/breast cancer pair, but not for the lung cancer/skin cancer pair. Likely reasons for this are given in the discussion below. Consequently, due to space limitations, only results for the heart disease/breast cancer pair are discussed here. For this pair, however, we see strong evidence for the presence of an estimation effect, and no evidence for other effects (see Table 3 below). Planned comparisons were conducted to test each of the alternative hypotheses. One-tailed tests of the contrasts corresponding to the joint effect and initial preference effect hypotheses were not significant (joint effect: $t(148) = -1.264$, *n.s.*; initial preference effect: $t(148) = -.014$, *n.s.*), while a one-tailed test of the contrast corresponding to the estimation effect hypothesis was significant, $t(148) = -2.378$, $p < .01$.

Table 3. Mean Final Allocations (C) to Heart Disease, by Group

	EPIC	PEIC	EIC	PIC	IC
Mean	$65.48	$63.28	$66.45	$56.17	$56.94

Discussion

The results for the heart disease/breast cancer pair provide strong evidence of an estimation effect, while no evidence was found for either an initial preference effect or a joint effect. Although no significant results were observed for the lung cancer/skin cancer pair, we believe that this may have been due to insufficient statistical power; effect size calculations showed that in Experiment 1, the effect of estimation on final allocations for the lung cancer/skin cancer pair, while still significant, was somewhat smaller than that for the heart disease/breast cancer pair—exacerbating the fact that the number of participants per condition in Experiment 2 was less than half of that in Experiment 1. The effect for the lung cancer/skin cancer pair might be smaller for multiple reasons. First, final allocations for the lung cancer/skin cancer pair appear to have been more variable, perhaps because concerns about responsibility were more influential. Second, as discussed in the next section when we present our explanation for the estimation effect, the size of the effect may well depend on the "surprisingness" of the numerical information presented. Recall that we mentioned earlier that students' estimates for the lung cancer/skin cancer pair were considerably more accurate than their estimates for the heart disease/breast cancer pair. It could be that estimates of the lung cancer/skin cancer mortality ratio were too accurate for an effect to be observed with the sample size used.

General Discussion

This paper's main experimental finding is that estimating the values of quantities before learning their true values increases the influence of the true values on preferences. Why might this be the case? An attractive explanation involves a form of overconfidence called "hindsight bias," which is "the tendency for people with outcome knowledge to believe falsely that they would have predicted the outcome of the event" (Hawkins and Hastie, 1990, p.311). It seems likely that hindsight bias would manifest itself when people are simply given a quantity's value; they will likely tend to believe that, had they estimated the value of the quantity beforehand, their estimate would have been more accurate than would really have been the case. But if one estimates before learning the true value, the potential for hindsight bias becomes greatly diminished or fully eliminated, since one's "cards are on the table," so to speak. This means that, to the extent that estimates of a quantity's value *would, if elicited,* diverge from the quantity's true value, the act of estimating will make the true value of the quantity more surprising.

It seems likely that more surprising information (assuming it is seen as reliable) will have greater influence during preference formation than more mundane information—perhaps for no other reason than that people attend more to surprising information, making it more salient. If inaccurately estimating a quantity's value makes learning the true value more surprising, this may explain why the final allocations of students in groups who estimated mortality rates (EPIC, PEIC, and EIC) were higher than those of students in groups who did not (IC and PIC); students who estimated were more surprised when they learned the true mortality rates, likely causing them to place more weight on mortality ratios. This explanation coheres with and helps account for past NDI findings showing larger differences between estimates and true values to be linked to larger differences between initial and final preferences (Garcia de Osuna et al., 2004; Lurie & Ranney, submitted; Munnich et al., 2003; Ranney et al., 2001).

Educational Implications and Future Directions

If one's goal is to maximize students' utilization of numerical information (although we realize that this may not always be appropriate), then "micro-interventions" based on NDI methods as simple as EIC may offer the best means currently known. One domain in which micro-interventions might be particularly useful is the study of "current events." Students might estimate the values of socially relevant quantities, receive true statistics, and then decide which public policies they think ought to be pursued. Similar applications abound elsewhere, as well.

Although it does not appear that stating initial preferences increases the utilization of numerical information, more research involving NDI methods like EPIC and PEIC is required. First of all, there may be pedagogical, motivation-related, reasons to include initial preferences, such as to increase student engagement. Second, results from pilot experiments show that when people rate NDI procedures in terms of the quality of the decisions they would expect them to yield, procedures that include both estimating and stating initial preferences are generally rated highest. Finally, results from a recent study involving the EPIC method—a study that showed post-feedback preferences to persist several months later—also showed that students' memories of numerical feedback seem to be linked to their memories for how their preferences *changed* in response to feedback (Munnich, Ranney, & Bachman, 2005). In order for such an effect to occur, it may be that students would need to explicitly state initial preferences. Thus, micro-interventions that include prompts for initial preferences may yet prove to be quite useful.

The research presented herein shows that merely providing statistics to students without engaging their expectations—or offering potential applications—is metaphorically like dropping seeds onto unplowed earth. This is reminiscent of science education research on Predict-Observe-Explain activities (White & Gunstone, 1992). In practice, our micro-interventions for increasing the impact of numerical information may be seen as analogous to Ausubel's (1960) "advance organizers" for learning verbal material. Ultimately, we hope that the techniques we suggest will help people incorporate numerical information into their preferences, leading to better decisions.

References

Ausubel, D. P. (1960). The use of advance organizers in the learning and retention of meaningful verbal material. *Journal of Educational Psychology, 51,* 267-272.

Bartol, K., Koehl, D., & Martin, D. (1987). Quantitative versus qualitative information utilization among college business students. *Educational and Psychological Research, 7*(1), 61-74.

Garcia de Osuna, J., Ranney, M., & Nelson, J. (2004). Qualitative and quantitative effects of surprise: (Mis)estimates, rationales, and feedback-induced preference changes while considering abortion. In K. Forbus, D. Gentner, & T. Regier (Eds.). *Proceedings of the Twenty-sixth Annual Conference of the Cognitive Science Society* (pp. 422-427). Mahwah, NJ: Erlbaum.

Hawkins, S. & Hastie, R. (1990). Hindsight: biased judgments of past events after the outcomes are known. *Psychological Bulletin, 107*(3), 311-327.

Hoeffler, S. & Ariely, D. (1999). Constructing stable preferences: A look into dimensions of experience and their impact on preference stability. *Journal of Consumer Psychology, 8*(2), 113-139.

Hsee, C. K. (1996). The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational Behavior and Human Decision Processes, 67*(3), 247-257.

Lurie, N., & Ranney, M. (submitted). *The Effects of Estimates and Base Rates on Preferences and Preference Change: How Thinking About What Is Can Change What One Wants.* University of North Carolina, Chapel Hill, and University of California, Berkeley.

Munnich, E. L., Ranney, M. A., & Appel, D. M. (2004). Numerically-Driven Inferencing in instruction: The relatively broad transfer of estimation skills. In K. Forbus, D. Gentner, & T. Regier (Eds*.). Proceedings of the Twenty-sixth Annual Conference of the Cognitive Science Society* (pp. 987-992). Mahwah, NJ: Erlbaum.

Munnich, E., Ranney, M., & Bachman, M. L. N. (2005). The longevities of policy-shifts and memories due to single feedback numbers. In B.G. Bara, L. Barsalou, & M. Bucciarelli (Eds.) *Proceedings of the Twenty-seventh Annual Conference of the Cognitive Science Society* (pp. 1553-1558). Mahwah, NJ: Erlbaum.

Munnich, E., Ranney, M., Nelson, J., Garcia de Osuna, J., & Brazil, N. (2003). Policy shift through Numerically-Driven Inferencing: An EPIC experiment about when base rates matter. In R. Alterman & D. Kirsh (Eds.), *Proceedings of the Twenty-Fifth Annual Conference of the Cognitive Science Society* (pp. 834-839). Mahwah, NJ: Erlbaum.

Paulos, J. A. (1988). *Innumeracy: Mathematical illiteracy and its consequences.* New York: Hill & Wang.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1992). Behavioral decision research: A constructive processing perspective. *Annual Review of Psychology, 43,* 87-131.

Rabin, M. (1998). Psychology and economics. *Journal of Economic Literature 36*(1), 11-46.

Ranney, M., Cheng, F., Garcia de Osuna, J., & Nelson, J. (2001, November). *Numerically driven inferencing: A new paradigm for examining judgments, decisions, and policies involving base rates.* Paper presented at the Annual Meeting of the Society for Judgment & Decision-making, Orlando.

Slovic, P. (1995). The construction of preference. *American Psychologist*, *50*, 364–371.

Viswanathan, M. (1993). Measurement of individual differences in preference for numerical information. *Journal of Applied Psychology, 78*(5), 741-752.

White, R. T., & Gunstone, R. F. (1992). *Probing Understanding.* Great Britain: Falmer Press.

Acknowledgments

We thank Christine Diehl, Patti Schank, and Geoff Saxe for their help and advice. We also thank Ed Munnich, Sujata Ganpule, Janek Nelson, Michelle Million, Kelvin Chan, Jed Stamas, Sarah Cremer, Sasha Raskin, Myles Crain, Nelson Bradley, Lauren Barth-Cohen, Andrew Galpern, Tammie Chen, Roxanne Garbez, the rest of the UCB Reasoning Group, and the ICBS faculty for their comments and suggestions. Comments and suggestions from three anonymous reviewers were also appreciated. This work was funded by faculty research grants from UCB and the University of North Carolina, Chapel Hill, and various fellowships from UCB and its GSE.