Research Report

VIOLATIONS OF MONOTONICITY AND CONTEXTUAL EFFECTS IN CHOICE-BASED CERTAINTY EQUIVALENTS

Michael H Birnbaum

California State University, Fullerton, and Irvine Research Unit in Mathematical Behavioral Science

Abstract—*This article investigates choices between gambles and amounts of money to explore two issues in decision making First, in recent studies, judgments of the values of gambles violated monotonicity (dominance), yet choices between the same gambles satisfied monotonicity, producing reversals of preference This experiment tested whether certainty equivalents based on choices between gambles and money would violate monotonicity Results indicated that these choices violated monotonicity in the same way as had judgments Second, this experiment investigated whether the certainty equivalent of a gamble would depend on the distribution of amounts offered for comparison It was found that certainty equivalents based on choices depended on the context in the same fashion as psychophysical comparisons Apparently, paradoxes of behavioral decision making are not eliminated by using choices instead of judgments to investigate human preferences*

Monotonicity is one of the most compelling principles of normative decision making The principle, sometimes called dominance, dates back at least to the time of Arnobius of Sicca (Grier, 1981) The principle can be stated briefly as follows If two alternatives are otherwise identical but one gamble has one outcome that is better, then the gamble with the higher outcome is better Despite the reasonableness of monotonicity, Birnbaum, Coffey, Mellers, and Weiss (1992) discovered a situation in which people assigned systematically lower judgments to gambles when the value of an outcome was increased

Let (x, p, y) represent the binary gamble to receive x with probability p and otherwise receive y Monotonicity requires that (x, p, y) is preferred to (z, p, y) if and only if x is preferred to z However, Birnbaum et al (1992) found that when $p \leq 2$, (\$0, p, \$96) receives a higher judgment than (\$24, p, \$96), even though \$24 is higher than \$0 The judged values of the highest price that a buyer should pay to play the gamble, the lowest price that a seller should accept to sell the gamble, and the "fair" price (from a neutral point of view) all showed similar results Higher judgments were assigned to the dominated gamble when $p = 05$, 1, and 2, but not when $p \geq 4$ Similar results were also found when \$72 replaced \$96 as the higher outcome Because it is so reasonable to assume that people prefer more money to less, these results seem a striking violation of monotonicity

Mellers, Weiss, and Birnbaum (1992) conducted a series of experiments to explore when the violations occur in judgment They used a different format for presentation of the probabilities (pie charts), which might reduce tendencies to attempt numerical calculations, but the violations persisted They found that violations of monotonicity occurred consistently when judgments of (x, p, y) were compared with judgments of $(0, p, y)$, when $p < 2$ and $0 < x < y/3$ The violations persisted when y was increased from \$81 50 to \$960, they were also found for negative values (when $p < 2$ and $0 > x > y/3$), but not when x and y were of opposite sign The pattern of results matched the predictions of the configural weight model of Birnbaum et al (1992), described in the appendix

Mellers et al (1992) found that violations persisted even when real money was used as an incentive, but they found one condition that significantly reduced violations of monotonicity When the key gambles were printed on the same page, there were fewer violations, as if conditions that facilitate comparisons among the gambles reduce violations

Birnbaum and Sutton (1992) replicated the results of Birnbaum et al (1992) for judgments of buyer's prices and seller's prices Birnbaum and Sutton also presented the gambles in pairs and found that although judgments of the gambles violated monotonicity when gambles were judged one at a time, it was extremely rare for subjects to choose the dominated gamble when both were presented simultaneously in a direct comparison Because judged values showed a different ordering from that obtained in direct choice, Birnbaum and Sutton identified their finding as a new type of preference reversal

In the "classic" preference reversal (Lichtenstein & Slovic, 1971, Lindman, 1971), gambles with equal or nearly equal expected values are compared Subjects assign higher prices to gambles with a small probability to win a high outcome [e g , (\$0, 95, \$96)] than they do to gambles with a high probability to win a small outcome [e g , (\$0, 2, \$6)] However, when subjects are offered a direct comparison, they choose the gamble with the higher probability to win

Researchers in preference theory were both disturbed and excited by these reversals, because they seemed to show that the most fundamental relationship in the theory, the preference relation itself, is difficult to operationalize in an internally consistent fashion (Krantz, Luce, Suppes, & Tversky, 1971) A number of theories were proposed to explain why different methods of elicitation yield different preferences (Birnbaum, in press, Busemeyer & Goldstein, in press, Mellers, Ordóñez, & Birnbaum, in press, Schoemaker & Hershey, in press, Slovic, Lichtenstein, & Fischhoff, 1988, Tversky, Sattath, & Slovic, 1988, von Winterfeldt & Edwards, 1986)

Bostic, Herrnstein, and Luce (1990) found that these classic preference reversals were reduced when choice-based certainty equivalents were used instead of judged certainty equivalents Choice-based certainty equivalents appeared

Address correspondence to Michael H Birnbaum, Department of Psychology, California State University, Fullerton, CA 92634, e-mail mbirnbaum@fullerton edu

better than judged certainty equivalents in predicting which of two gambles a judge would choose when offered a direct comparison

One might conclude, then, from Bostic et al (1990), Birnbaum and Sutton (1992), and Mellers et al (1992), that perhaps reversals of preference, including both the classic reversals and the reversals due to monotonicity violations, might be reduced if certainty equivalents were determined by the method of choice, rather than by judgment

However, presenting the choice between a gamble and a cash value does not offer the "transparent" situation afforded in a direct choice between gambles, since each gamble is compared to an amount and the experimenter tests monotonicity by examining how each gamble stacks up against money

In decision making, there has been a standard assumption that when faced with a choice between a gamble and a sure amount, the subject compares the utility of the gamble with the utility of the comparison amount and chooses the gamble if and only if its utility exceeds that of the money However, there is evidence that choice is not that simple

In psychophysics, judgment and choice are known to be subject to contextual effects (Birnbaum, 1982, Birnbaum, Parducci, & Gifford, 1971, Mellers & Birnbaum, 1982, Parducci, 1990, Parducci & Haugen, 1967, Poulton, 1989) Garner (1954) attempted to find a tone that would seem "half as loud" as a standard tone by asking subjects in a choice-based procedure to judge whether each comparison tone was "more" or "less" than "half as loud" as the standard He manipulated the context (distribution) of comparison stimuli and found that the tone inferred to be "half as loud" as the standard was the median of the comparison stimuli

Garner concluded that subjects may have no idea what tone is half as loud as another and that the experimenter determines the result by the selection of comparisons, two experimenters would find two different "half-loudness" values by using different comparisons with the same standard Since Garner's paradigm is analogous to the procedures used to find choice-based certainty equivalents, it is reasonable to ask if choices in decision making might also be susceptible to the same psychological processes of context as psychophysical judgments

The present experiment investigated this question by manipulating the distribution of comparison amounts It also investigated whether monotonicity violations persist in the comparison procedure

METHOD

Instructions read (in part) as follows

> On each trial you will be offered a comparison between an amount of money and a gamble, or lottery Your task is to decide whether you would prefer the money (for sure) or the chance to play the lottery (the gamble) Compare the gamble to each amount If you prefer the amount of money circle the amount Circle all of the sure amounts of money that you would prefer to the gamble

Stimuli and Design

Gambles were displayed as in the following example

2	8
$24	$96

This display represents a probability of 2 to win $24 and a probability of 8 to win $96 Subjects were instructed to imagine a can with 20 slips specifying the smaller amount and 80 specifying the larger amount, 1 slip would be chosen at random to determine the amount won Probabilities displayed always summed to 1

The 30 binary gambles were generated from a factorial design of six pairs of amounts [(x, y) = ($0, $24), ($0, $48), ($0, $96), ($24, $48), ($24, $96), ($48, $96)] combined with five levels of the probability of receiving the smaller amount (p = 05, 2, 5, 8, or 95)

Comparison Contexts

Each of the 30 gambles was presented for comparison with two sets of comparison amounts Context 1 (positively skewed distribution) included the following dollar amounts 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50, 60, 70, 80, and 90 Context 2 (negatively skewed distribution) included the following dollar amounts 10, 20, 30, 40, 50, 60, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, and 90 These comparison amounts were printed in ascending order, in a vertical column below each gamble

Procedure and Subjects

The 30 gambles, with their comparison amounts, were printed in random order in two booklets, each booklet began with instructions and six warm-up trials Each booklet contained either Context 1 or Context 2 comparison amounts, and the warm-ups for each context used the appropriate series of comparisons for the condition Half of the subjects completed Context 1 followed by Context 2, and half of the subjects received the booklets in the opposite order In addition, half of the subjects received a different order of trials (these order effects were negligible)

Instructions stated that subjects should prefer a gamble to any amount less than the least amount the gamble would offer and that they should prefer amounts of money that exceeded the most the gamble could offer Subjects who violated these properties during the warm-up were directed to reread instructions before proceeding [This instruction would have the effect of reducing violations of monotonicity because it rules out any response less than $24 for the ($24, p, $96) gamble, but does not rule out small values for the ($0, p, $96) gamble]

The subjects were 46 undergraduates at California State University, Fullerton They received extra credit in an introductory psychology course

RESULTS AND DISCUSSION

The minimum value of money preferred to each gamble was taken as a dependent variable Table 1 displays the means of this variable for each gamble in each context Rows indicate amounts to win, and columns depict probabilities to receive the larger amount in each pair

Violations of Monotonicity

Figure 1 plots the mean (of the minimum amount preferred to each gamble)

Table 1 *Mean value of smallest comparison preferred to gamble*

Amounts to win	1 − p				
	05	2	5	8	95
		Context 1			
($0, $24)	16 4	16 4	21 3	24 5	26 9
($0, $48)	23 0	26 6	33 8	41 6	48 9
($0, $96)	25 8	32 7	53 8	74 0	79 2
($24, $48)	33 4	33 0	39 1	45 4	47 4
($24, $96)	37 5	37 3	56 4	69 4	77 3
($48, $96)	55 2	54 1	68 1	70 2	81 1
		Context 2			
($0, $24)	20 7	22 0	23 3	26 7	29 1
($0, $48)	33 7	33 9	38 7	45 0	47 8
($0, $96)	38 8	42 4	58 0	78 2	82 8
($24, $48)	36 1	37 3	43 2	49 3	51 2
($24, $96)	42 2	47 2	60 1	74 3	79 4
($48, $96)	58 5	62 1	70 7	77 5	85 8

Note Each entry is the mean of the smallest comparison amount that is just preferred to each gamble, $1 - p$ is the probability to receive the larger amount Underlined values show violations of monotonicity discussed in the text

as a function of $1 - p$, with separate curves for ($0, p, $96) and for ($24, p, $96), averaged over contexts Monotonicity implies that judgments of ($24, p, $96) should exceed judgments of ($0, p, $96) for all values of p, that is, the curves should not cross Instead, the mean value for ($0, p, $96) is higher than the mean for ($24, p, $96) for $p =$ 2 and $p =$ 05 For these two values of p, mean responses are significantly higher for the

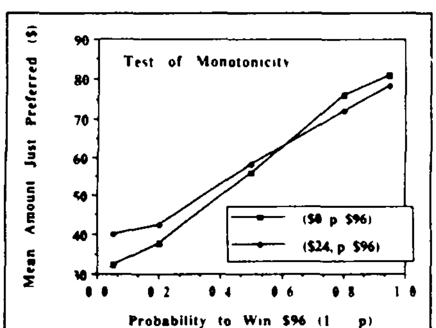

Fig 1 Mean value of the smallest comparison amount just preferred to each gamble, plotted against probability to receive the larger value ($1 - p$), averaged over contexts Monotonicity implies that the curve for ($24, p, $96) should not cross below the curve for ($0, p, $96)

dominated than for the dominating gambles, $F(1, 44) = 5\ 81$ For these four tests (underlined in Table 1), 70% of the subjects showed at least one violation of monotonicity, 50% of the subjects violated monotonicity more often than they satisfied it, whereas only 25% satisfied it more often than not (the others came out even) Violations of monotonicity were similar in both contexts (see Table 1), and they fit the pattern previously obtained with judgments of value (Birnbaum et al , 1992)

The violations of monotonicity in Figure 1 combined with previous results seem to imply an intransitivity of choice because there should exist an amount of money, c^*, such that ($0, 05, $96) is preferred to c^*, which is preferred to ($24, 05, $96), yet the latter gamble is preferred in direct choice to the former (Birnbaum & Sutton, 1992) Perhaps one could even find a gamble, G^*, such that G^* could replace the monetary amount c^* in the above relations A finding by Tversky and Kahneman (1986) suggests that it may be possible to observe such an intransitivity between subjects However, it seems unlikely that such an intransitivity could be maintained within subjects

Nevertheless, the present findings offer practical advice to people who are sales oriented If a savings and loan company, trying to sell low-risk investments, were to say, " and, in the unlikely event that we go bankrupt, you will receive 25% of your investment back," this statement would make the investment seem worse Apparently, an unlikely zero outcome is easier to ignore than a small payoff Configural weight theory can predict the violation of monotonicity by assuming that when the lowest outcome of a gamble is zero, the lowest outcome receives lower weight than when it is positive (see the appendix)

Contextual Effects

Table 1 shows that the means were higher for 29 of the 30 gambles in Context 2 than in Context 1 The average difference was $4 86, which is statistically significant, $F(1, 44) = 26\ 68$ At low levels of probability to win the larger amount ($1 - p$), the context effect was greater than at higher levels of probability, especially for gambles in which $x = 0$, the Context × Probability interaction was significant, $F(4, 176) = 4\ 13$, but other interactions involving context were not

Figure 2 plots the percentage of times that the comparison amount (sure thing)

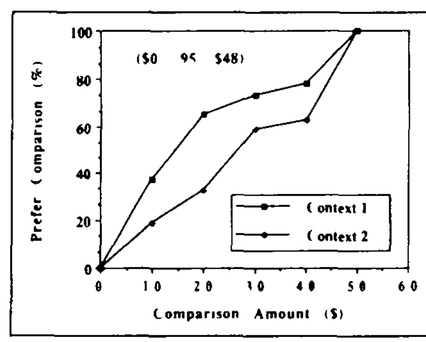

Fig 2 Percentage of choices favoring the certain money over the gamble (with a probability of 05 to win $48, and otherwise to win $0), plotted as a function of the amount of money offered A separate curve is shown for each context (in Context 1, most of the comparisons presented to the subjects were less than $20, in Context 2, most of the comparisons were greater than $70) Inferred certainty equivalents are projections on the abscissa corresponding to the ordinate value of 50%

was preferred to the ($0, 95, $48) gamble, with a separate curve for each context (Data were also graphed as in Fig 2 for each gamble separately, the typical shape of the curves and the direction of the contextual effects were similar to those illustrated in Fig 2, although Fig 2 depicts one of the larger contextual effects)

The certainty equivalents (projections of the curves at 50% on the ordinate onto the abscissa in Fig 2) are larger in Context 2 than in Context 1 The values are $14 61 and $26 67 for Contexts 1 and 2, respectively Contextual effects also explain why the certainty equivalents for this gamble, which has an expected value of only $2 40, are so high in both contexts—most of the comparison values exceed the expected value Recall that the medians of comparisons were $14 and $77 in the two contexts

One viewpoint might hold that the "true" value of this gamble should be closer to its expected value, but these results were "biased" by the distribution of comparisons However, such a position quickly becomes circular If we already knew the "right" values, and if we understood contextual effects, then we could select comparisons for each gamble to produce the "right" answer But if we do not already know the answer, how do we choose the "right" context? These issues are discussed from contrasting viewpoints by Poulton (1989) and Birnbaum (1992)

Procedures have been advocated that are intended to find "unbiased" values (e g , counterbalancing, staircase methods, between-subjects designs), however, disagreements between opposite proposals have not yet been resolved For example, some investigators would argue that the context should be tailored for each gamble separately However, others would prefer to unconfound the distribution of comparisons from the particular gamble, they might prefer the method of the present experiment, which uses a factorial combination of gambles and comparisons

The contextual effects in this experiment are probably only the tip of the iceberg In the present study, the median of the comparison stimuli was manipulated, but the distributions of the features of the gambles (*x*, *p*, and *y*) were fixed Recent studies have found surprising effects caused by manipulations of other aspects of the context (Mellers et al , in press, Stevenson, in press)

CONCLUSIONS

The present results indicate that the procedure of comparison per se does not eliminate monotonicity violations Comparison judgments are also susceptible to contextual effects that complicate the interpretation of choice-based certainty equivalents These findings rule out a simple world in which contextual effects, monotonicity violations, and preference reversals could all be attributed to the peculiarities of direct judgment They also rule out the assumption that subjects compare gambles with amounts by contrasting their invariant utilities Instead, the values of—or the comparisons between—alternatives depend on the distributions that form the context of choice

Acknowledgments—Thanks are due Suzanne Mistretta and Steve McCormick for assistance Support was received from National Science Foundation Grant SES 8921880 and a Senior Faculty Fellowship from California State University, Fullerton

REFERENCES

Birnbaum M H (1982) Controversies in psychological measurement In B Wegener (Ed) *Social attitudes and psychophysical measurement* (pp 401–485) Hillsdale NJ Erlbaum

Birnbaum M H (1992) Should contextual effects in human judgment be avoided? [Review of *Bias in quantifying judgments*] *Contemporary Psychology 37* 21–23

Birnbaum M H (in press) Issues in utility measurement *Organizational Behavior and Human Decision Processes*

Birnbaum M H Coffey G Mellers B A & Weiss, R (1992) Utility measurement Configural-weight theory and the judge s point of view *Journal of Experimental Psychology Human Perception and Performance 18* 331–346

Birnbaum M H Parducci A & Gifford, R K (1971) Contextual effects in information integration *Journal of Experimental Psychology 88* 158–170

Birnbaum M H , & Sotoodeh, Y (1991) Measurement of stress Scaling the magnitudes of life changes *Psychological Science 2* 236–243

Birnbaum M H & Stegner S E (1979) Source credibility in social judgment Bias expertise and the judge's point of view *Journal of Personality and Social Psychology 37* 48–74

Birnbaum M H , & Sutton, S E (1992) Scale convergence and utility measurement *Organizational Behavior and Human Decision Processes 52* 183–215

Bostic R Herrnstein, R J & Luce, R D (1990) The effect on the preference-reversal phenomenon of using choice indifference *Journal of Economic Behavior and Organization 13* 193–212

Busemeyer J R (1985) Decision making under uncertainty Simple scalability fixed sample, and sequential sampling models *Journal of Experimental Psychology Learning Memory and Cognition 11* 538–564

Busemeyer, J R & Goldstein W M (in press) Linking together different measures of preference A dynamic model of matching derived from decision field theory *Organizational Behavior and Human Decision Processes*

Garner W R (1954) Context effects and the validity of loudness scales *Journal of Experimental Psychology 48* 218–224

Grier B (1981, August) *The origin of risky choice and the dominance principle* Paper presented at the Fourteenth Annual Mathematical Psychology meetings Santa Barbara CA

Krantz D H Luce R D , Suppes P & Tversky A (1971) *Foundations of measurement Vol I Additive and polynomial representations* New York Academic Press

Lichtenstein S & Slovic P (1971) Reversals of preference between bids and choices in gambling decisions *Journal of Experimental Psychology 89* 46–55

Lindman H R (1971) Inconsistent preferences among gambles *Journal of Experimental Psychology 89* 390–397

Mellers B A & Birnbaum M H (1982) Loci of contextual effects in judgment *Journal of Experimental Psychology Human Perception and Performance 8* 582–601

Mellers B A Ordóñez L & Birnbaum M H (in press) A change of process theory for contextual effects and preference reversals in risky decision making *Organizational Behavior and Human Decision Processes*

Mellers B A Weiss R & Birnbaum M H (1992) Violations of dominance in pricing judgments *Journal of Risk and Uncertainty 5* 73–90

Parducci A (1990) Response bias and contextual effects When biased? In J P Caverni J -M Fabre & M Gonzales (Eds) *Cognitive biases* (pp 207–219) Amsterdam North Holland

Parducci A & Haugen R (1967) The frequency principle for comparative judgments *Perception & Psychophysics 2* 81–82

Poulton E C (1989) *Bias in quantifying judgments* Hove United Kingdom Erlbaum

Schoemaker, P J & Hershey J C (in press) Utility measurement Signal noise and bias *Organizational Behavior and Human Decision Processes*

Slovic P Lichtenstein S & Fischhoff B (1988) Decision making In R C Atkinson R J Herrnstein G Lindzey & R D Luce (Eds) *Stevens handbook of experimental psychology* (Vol 2 pp 673–738) New York Wiley

Stevenson M K (in press) The impact of temporal context and risk in the judged value of future outcomes *Organizational Behavior and Human Decision Processes*

Tversky A & Kahneman D (1986) Rational choice and the framing of decisions *Journal of Business 59* S251–S278

Tversky A Sattath S & Slovic, P (1988) Contingent weighting in judgment and choice *Psychological Review 95* 371–384

von Winterfeldt D , & Edwards, W (1986) *Decision analysis and behavioral research* Cambridge England Cambridge University Press

Wakker, P (in press) Additive representations on

rank-ordered sets Part II The topological approach *Journal of Mathematical Economics*
Weber, E U Anderson C J & Birnbaum M H (in press) A theory of perceived risk and attractiveness *Organizational Behavior and Human Decision Processes*

(RECEIVED 8/26/91, REVISION ACCEPTED 2/28/92)

APPENDIX

Birnbaum et al (1992) represented judgments of binary gambles, (x, p, y), by the following configural weight model

$$U_V(x, p\ y) = \frac{a_V S_x(p)u(x) + (1 - a_V)(1 - S_x(p))u(y)}{a_V S_x(p) + (1 - a_V)(1 - S_x(p))}, \quad (A1)$$

where $U_V(x\ p, y)$ is the utility of the gamble in point of view V, a_V is the configural weighting parameter for point of view V, $u(x)$ and $u(y)$ are the utilities of the lower- and higher-valued outcomes and $S_x(p)$ is a function of the probability of the lower-valued outcome that depends on value There are different S functions for $x > 0$ and for $x = 0$ This model fit the data of Birnbaum et al (1992) and predicted the patterns obtained by Birnbaum and Sutton (1992) and by Mellers et al (1992)

Figure A1 illustrates predictions of this model for gambles of the form $(x\ p\ y)$, for different values of x and p with y fixed Predictions were calculated with $a_V = 5$, $u(x) = x$ $S(p) = 30 + 58p$ for $x > 0$, and $S_0(p) = 14 + 73p$ for $x = 0$ (For $05 \leq p \leq 95$, these expressions give values of $S_x(p)$ that are close to the estimates of Birnbaum et al , 1992, who set $a_V = 5$ for the seller s point of view) The open and solid arrows in Figure A1 show that the predicted value of ($0, p y) can exceed ($24, p y) so the model predicts the violations of monotonicity

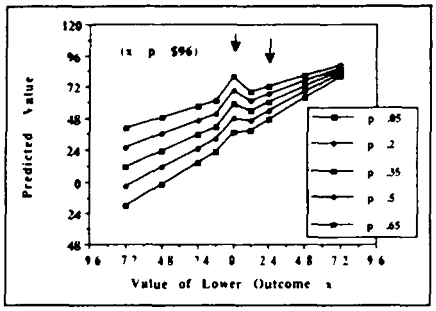

Fig A1 Predicted judgments, according to the configural weight model of Birnbaum et al (1992), for gambles of the form $(x, p, \$96)$ plotted as a function of x, with separate curves for different values of p The open and solid arrows highlight values of $x = \$0$ and $x = \$24$, respectively (see discussion in the text)

The term configural is used to indicate that the parameter representing a stimulus component may depend on the relationships between that component and others that make up the stimulus The configural weight model (Equation A1) allows the weight of an outcome to depend on its rank among the other outcomes in the gamble Therefore, the weight of the same outcome with the same probability can be different in different gambles (Birnbaum 1982 Birnbaum et al 1992 Birnbaum & Sotoodeh, 1991 Birnbaum & Stegner 1979 Weber Anderson & Birnbaum, in press)

The model is closely related to rank-dependent utility theories, which were developed independently (see review by Wakker, in press), except that configural weighting allows weights to depend on point of view and to differ for the zero-valued outcomes, which allows configural weight theory to explain violations of monotonicity

Changes in the configural weight parameters, a_V explain why the rank order of gambles changes in different points of view Configural weight theory led to estimated $u(x)$ functions that were invariant with respect to point of view (Birnbaum et al , 1992), estimated $u(x)$ functions also agree with estimates based on subtractive theory applied to judgments of ratios and differences" of riskless utility (Birnbaum & Sutton, 1992)

Contextual effects can be discussed with respect to the following model of choice

$$P(G, c) = F[U(G) - u(c)], \quad (A2)$$

where $P(G\ c)$ is the probability of choosing the gamble $G = (x\ p\ y)$ over the sure amount c U is a function (e g , Equation A1) that assigns an overall utility to each gamble, u is a utility function for money, and F is a monotonic function that maps a given utility difference into a choice probability The results in Table 1 and Figure 2 are inconsistent with this theory of comparative judgment, if the functions are assumed to be invariant If Equation A2 is to be saved, then F U, or u must be subscripted with the context Busemeyer (1985) observed violations of scalability that would also constitute evidence against Equation A2 which could be saved if the function F were permitted to depend on the variance of the outcomes in each gamble (see also Schoemaker & Hershey, in press)