The role of numeracy and approximate number system acuity in predicting value and probability distortion

Andrea L. Patalano, Jason R. Saltiel, Laura Machlin & Hilary Barth

Psychonomic Bulletin & Review

ISSN 1069-9384
Volume 22
Number 6

Psychon Bull Rev (2015) 22:1820-1829
DOI 10.3758/s13423-015-0849-9

Springer

Psychon Bull Rev (2015) 22:1820–1829
DOI 10.3758/s13423-015-0849-9

BRIEF REPORT

The role of numeracy and approximate number system acuity in predicting value and probability distortion

Andrea L. Patalano[1,2] · Jason R. Saltiel[1] · Laura Machlin[1] · Hilary Barth[1]

Published online: 24 September 2015

Abstract It is well documented that individuals distort outcome values and probabilities when making choices from descriptions, and there is evidence of systematic individual differences in distortion. In the present study, we investigated the relationship between individual differences in such distortions and two measures of numerical competence, numeracy and approximate number system (ANS) acuity. Participants indicated certainty equivalents for a series of simple monetary gambles, and data were used to estimate individual-level value and probability distortion, using a cumulative prospect theory framework. We found moderately strong negative correlations between numeracy and value and probability distortion, but only weak and non-statistically reliable correlations between ANS acuity and distortions. We conclude that low numeracy contributes to number distortion in decision making, but that approximate number system acuity might not underlie this relationship.

Keywords Numeracy · ANS acuity · Cumulative prospect theory · Probability weighting

According to cumulative prospect theory (CPT), a dominant descriptive model of decision making under risk (Tversky & Kahneman, 1992), decision makers transform stated outcomes and their likelihoods during their valuation of choices. A basic tenet of CPT is that people generally exhibit diminishing sensitivity for gains; in other words, each additional unit is treated as less valuable than the previous one (such that, e.g., the difference between \$200 and \$300 is treated as greater than \$1000 and \$1100; Bernoulli, 1738/1954; Von Neumann & Morgenstern, 1947).[1] According to CPT, not only are outcomes transformed into subjective values with a concave curve, but probabilities are transformed into decision weights following an inverse S-shaped curve, where small probabilities are overweighted (e.g., people behave as if a 10 % chance is more likely than it actually is) and medium to large ones are underweighted. These functions explain a complex pattern of risk attitudes: risk aversion for high probability gains and risk seeking for high probability losses, but the reverse for low probability gains and losses (Tversky & Kahneman, 1992). We refer to these transformations as distortions in the mathematical sense that they reflect deviations from identity relationships (the diagonal $y = x$ on a graph) between stated outcomes and subjective values and between probabilities and decision weights, not in the decision-making sense of a deviation from normativity. In CPT, value functions are generally considered to have no normatively optimal shape, while the normatively optimal probability function is the identity relationship.

Aggregate patterns of value and probability functions notwithstanding (for psychological accounts, see, e.g., Hogarth & Einhorn, 1990; Rottenstreich & Hsee, 2001; Tversky & Kahneman, 1992), individual-level variation in degree of distortion is quite large (Gonzalez & Wu, 1999) and relatively stable over time (Birnbaum, 1999; Glöckner & Pachur, 2012; Zeisberger, Vrecko, & Langer, 2012), raising intriguing

✉ Andrea L. Patalano
apatalano@wesleyan.edu

1 Wesleyan University, Middletown, CT, USA

2 Department of Psychology, Wesleyan University, 207 High St, Middletown, CT 06459, USA

[1] We use the term "sensitivity" to refer to the fact that a flat curve (e.g., a value curve that gets flatter as the outcome grow larger, or a probability curve that is flatter around 50 %) indicates that individuals behave as if they are less sensitive to changes. Use of the term is not intended to imply a particular cognitive account how why this happens.

questions about the sources of variation. Here, building on limited past work, we investigate whether individual differences in cognitive capacities associated with numerical competence—in this case, *numeracy* and *approximate number system (ANS) acuity*—contribute to degree of distortion of both value and probability functions. We first describe existing literature on the relationship between numeracy and decision making, and also consider the possible role of ANS acuity in the representation and manipulation of numerical quantities during decision making. We then empirically test the claim that more numerate individuals might show less distortion during decision making and that this relationship might be explained by approximate number system acuity.

Numeracy as used here refers to one's learned skill at understanding, manipulating, and translating between numerical part-whole relationships presented in symbolic forms including fractions, percentages, and proportions. This type of numeracy has particular relevance to decision making from descriptions involving risk. It is measured with questions like those in the standard 11-item numeracy scale of Lipkus, Samsa, and Rimer (2001), such as "The chance of getting a viral infection is .0005. Out of 10,000 people, about how many of them are expected to get infected?" In choice tasks, more numerate individuals are less influenced by positive versus negative problem frames, frequency versus probability risk format (Pachur & Galesic, 2013; Peters, Västfjäll, Slovic, Mertz, Mazzocco, & Dickert, 2006), and narrative information (Dieckmann, Slovic, & Peters, 2009). More numerate individuals also tend to have less difficulty with utility elicitation (see Reyna, Nelson, Han, & Dieckmann, 2009, for a review), and are more likely to adopt number-based decision strategies than to rely on other decision heuristics, such as intuition (Cokely & Kelley, 2009; Pachur & Galesic, 2013). They also show greater probability sensitivity (Reyna et al., 2009) in that they modulate their choice behavior in response to even small changes in given probabilities, and they have better-calibrated subjective probability judgments (Winman, Juslin, Lindskog, Nilsson, & Kerimi, 2014); that is, judgments that more closely match actual relative frequencies.

People are also known to have a more basic innate capacity to represent and differentiate numbers of different magnitudes with an *approximate number system (ANS)*. This evolutionarily ancient system is more primitive, nonverbal, and imprecise, and is intuitive rather than relying on formal learning (see Libertus & Brannon, 2009, for a review). A standard measure of ANS acuity is a visual number discrimination task in which two sets of dots differentiated by color flash rapidly on a computer screen, and one judges which set is more numerous (see Fig. 1 for the version used here, developed by Halberda, Mazzocco, & Feigenson, 2008). Discrimination difficulty varies with the numerical ratio of the dots in the sets, and individuals differ in the difficulty at which they perform at chance levels. Although visual tasks are most commonly used to assess ANS acuity, a wealth of research has shown that the ANS is not specific to one sensory modality (see, e.g., Brannon, 2006). Rather, the ANS appears to underlie the ability to represent and manipulate approximate abstract numerical quantities.

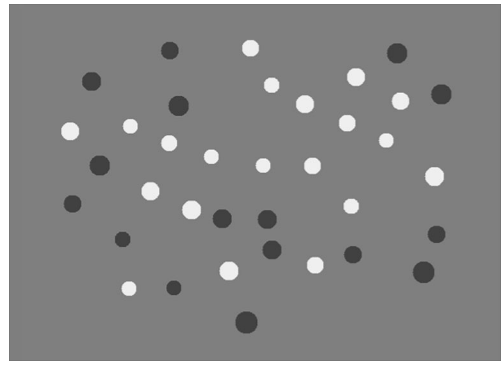

Fig. 1 Example stimulus from the approximate number system (ANS) acuity task. The task was to indicate as quickly as possible which color circle was more numerous. Circles were blue and yellow in the actual task

There is some evidence of a relationship between the approximate number system and formal math skills, including skills associated with numeracy. Scores on ANS tasks have been correlated with adult numeracy scores (Paulsen, Woldorff, & Brannon, 2010; but see Inglis, Attridge, Batchelor, & Gilmore, 2011), as well as other assessments of formal math ability (Halberda et al., 2008; Libertus, Odic, & Halberda, 2012). Even response times for magnitude discriminations between highly familiar numerals are influenced by numerical ratios, suggesting ANS involvement. For example, adults are slower in deciding whether 5 or 6 is the larger number than in deciding whether 5 or 8 is larger (Moyer & Landauer, 1967). Further, Peters, Slovic, Västfjäll, and Mertz (2008) found that numeracy scores are related to performance on magnitude comparison tasks using dots and numerals. Thus, there is reason to believe that the ANS could be linked to a wide range of activities that require use of numbers and, more specifically, that ANS acuity could be related to the distortion of numerical quantities in decision making.

Schley and Peters (2014) recently conducted an initial investigation of the relationship between numerical competence and value and probability functions. In addition to numeracy as measured by a verbal numeracy scale, they considered another form of numerical competence: performance on a number line task which required marking the appropriate locations of 11 numerals (0.1, 0.8. 1.5, 9.5, 23.2, 89.3, 268, 442, 682, 834, and 925) on a line labeled only with endpoints of 0 and 1000. Numeracy was negatively correlated with value (but not probability) distortion as inferred from gambling task behavior, and number line estimation accuracy mediated the relationship between numeracy and value distortion. However, because this number line task draws upon multiple distinct skills and capacities (e.g., mapping of symbols to mental magnitudes, discrimination among magnitudes, proportional mappings between numerical and spatial quantities), including facility with decimals which might be better understood as a component of numeracy, these findings leave open many questions about the source of the relationship between number line performance and CPT parameter distortions. Furthermore, the question of whether there is a relationship between ANS acuity (distinct from other skills such as symbol-magnitude mapping) and CPT distortions remains unexplored.

In the present study, we investigated the relationship of numeracy, ANS acuity, and value and probability distortions. Participants completed a standard numeracy scale, a visual numerosity discrimination task, and a gambling task from which indices of value and probability function distortion could be inferred from *certainty equivalent (CE)* data. A certainty equivalent is a monetary outcome identified by an individual as being as attractive as a stated risky gamble, and thus reflecting the overall value of a gamble to an individual. For example, the CE of the gamble "a 25 % chance of $100, otherwise nothing" would be $50 if receiving an outright $50 were deemed just as desirable as taking the gamble. CEs can be collected over many gambles and used to infer individual-level CPT coefficients that index value and probability distortion.

The goals of the present study were to further establish a relationship between numeracy and value and probability distortion, and to test the relationship between ANS acuity and these distortions, with one possibility being that ANS acuity might underlie the relationship between numeracy and CPT distortions. We also considered whether CPT models—which assume that decision makers interpret and integrate values and probabilities in a consistent manner—generally better fit the data for individuals higher in numerical competence.

Method

Participants

Participants were 57 undergraduate students (24 male, 33 female; 18–22 years old), who received either $10 per hour or introductory psychology course credit for completing two sessions. The gambling task was administered in the first session and numerical competence measures were given in the second session, with sessions at least 3 days apart ($M = 6$, range = 3–14). The first session lasted approximately 90 min and the second was 30 min.

Gambling task

Materials The task was similar to that of Gonzalez and Wu (1999). A set of 165 unique two-outcome gambles was created by crossing 15 pairs of dollar values with 11 probabilities. The pairs of dollar values were 25–0, 50–0, 75–0, 100–0, 150–0, 200–0, 400–0, 800–0, 50–25, 75–50, 100–50, 150–50, 150–100, 200–100, and 200–150. The probabilities were .01, .05, .10, .25, .40, .50, .60, .75, 90, .95, and .99. Eleven of these gambles, one at each probability level, were repeated as a reliability check, resulting in a total of 176 gambles. Ten random orders of these gambles were created; no gamble could appear twice in a row.

Procedure The gambling task was administered by computer in the laboratory to participants who completed the task individually at their own pace. Each trial began with a display as shown in Fig. 2. The task was to compare the stated gamble (e.g., "95 % chance of $100, otherwise $0") to each of six "sure-thing" dollar amounts (e.g., $100, $80, $60, etc.) and to indicate a preference between the gamble and each sure thing. The sure-thing amounts ranged from the largest possible gamble outcome to the smallest (i.e., from $100 to $0 in Fig. 2), with intermediate values at equally spaced intervals. Participants were told that it was expected that they would "cross over" from preferring the sure thing to preferring the gamble at some point for each display. When done with the first display, participants went on to a second display (for the same trial), which presented a narrower range of sure-thing values (the new endpoints were the sure-thing values bracketing the crossover point from the previous display, e.g., the endpoints might be from $80 to $60). The dollar value at the middle of the participant's crossover point for the second display was recorded as the certainty equivalent of the gamble, the amount of money that the gamble was worth to the individual.

Numerical competence measures

The numeracy scale described earlier (Lipkus et al., 2001; Peters et al., 2006), an 11-item questionnaire, was administered on paper. This was followed by the visual number discrimination task (Halberda et al., 2008; Panamath.org) in which trials of intermixed yellow and blue dots, as shown in

Gamble: 95% chance of $100, otherwise $0		
Sure Thing	Prefer Sure Thing	Prefer Gamble
$100	x	
$80	x	
$60	x	
$40		x
$20		x
$0		x

NEXT

Fig. 2 In this gambling task, participants were instructed to choose a preference between the gamble and each sure thing option; hypothetical responses are shown here

Fig. 1, were presented on a computer screen for 600 ms each, followed by a 200-ms multicolored-pixel mask. Participants indicated which color was more numerous (using "F" and "J" keys for yellow and blue, respectively), as quickly and accurately as possible, then pressed the space key to continue. A total of 328 trials with different ratios of yellow to blue dots were presented in random order without feedback in a ~10-min period. Yellow and blue dots were equated for cumulative surface area on half the trials and for average size on the other half. Number of dots per color ranged from 5 to 21, and ratio of sets ranged from 1.10 to 2.47 (the Panamath program's difficulty setting for a 21-year-old). Finally, a personality inventory was administered that is unrelated to the present report.

Results

For 6 participants, estimated value and probability coefficients for the gambling task were implausible (coefficients were outside a 0–3 range and > 4 SD above the mean). These participants, whose data were also associated with low response reliability across repeated trials ($M = .52$) and many excluded trials ($M = 8$), were not included in any analyses. For the remaining 51 participants, numerical competence scores were not related to number of missing trials or to response reliability across repeated trials (ps > .100).

Given some skewed variables (with |skewness| > 1.0; Bulmer, 1979) in our data set, we report skewness as a descriptive measure for all variables. We ran all analyses using both linear and ordinal correlation and regression methods, and we re-ran linear analyses using variables that were transformed to reduce skew (following Howell, 2007). Because these analyses did not lead to different conclusions, we report only linear analyses with untransformed variables here.

Gambling task

Data processing For an average of 8.5 trials ($SD = 7.6$, range = 0–30) per participant, a certainty equivalent could not be directly inferred (i.e., the participant either did not give any response or gave the same response to every comparison). These trials were excluded except for ones in which a participant indicated a crossover point on the first screen of a trial, but indicated all "sure thing" or all "gamble" responses on the second screen. In these cases, because the certainty equivalent range was narrowed by the first response, it was reasonable to conclude that a participant who chose all "sure thing" responses on the second screen was intending the lowest crossover point and a participant who chose all "gamble" responses intended the highest. Ultimately, an average of 2.7 trials ($SD = 3.6$, range = 0–24) per participant were excluded; only one participant had more than 10 excluded trials. Analyses run with the latter participant omitted yielded no differences in findings and are not reported here. As a measure of reliability, correlations were computed for certainty equivalents for the 11 repeated trials for each participant. Reliability was high ($r = .96$, $SD = .08$, range = .57 – 1.00).[2]

Parameter estimation procedure According to CPT, the value to an individual of a simple gamble with two possible non-negative outcomes represented as ($x_1, p; x_2, 1 - p$), where $x_1 > x_2 \geq 0$, is $v(CE) = v(x_1)w(p) + v(x_2)(1 - w(p))$. In this equation, $v(x)$ represents the value function that transforms the CE and each dollar value (x_1 and x_2) into subjective values, and $w(p)$ represents the weighting function that transforms the probability p associated with x_1 (the larger dollar value) into a decision weight. The decision weight for the second probability is simply $1 - w(p)$; that is, it is a subtraction of the weight of the first probability from 1. In Tversky and Kahneman's (1992) two-parameter CPT model, a general one-parameter power function $v(x) = x^{\alpha}$, is used to transform CE, x_1, and x_2 values to subjective values, and a one-parameter weighting function $w(p) = p^{\beta}/(p^{\beta} + (1\text{-}\ p)^{\beta})^{1/\beta}$ is used to transform p into a decision weight. Substituting these functions into the original equation and isolating the CE term, we obtain $CE = (x_1^{\alpha} \bullet p^{\beta}/(\ p^{\beta} + (1\text{-}\ p)^{\beta})^{1/\beta} + x_2^{\alpha} \bullet (1 - (p^{\beta}/(\ p^{\beta} + (1\text{-}\ p)^{\beta})^{1/\beta})))^{1/\alpha}$. To obtain parameter estimates for each participant, we simply fit this equation to each individual's CE data. SPSS statistical software, nonlinear least squares regression, and a Levenberg-Marquardt algorithm for parameter estimation were used. All parameter estimates were unconstrained with starting values of 0.5. Coefficients of 1 indicate no deviation from identity (e.g., a probability of 75 % is treated as 75 %), and thus no distortion of value or probability.

We also re-ran the parameter estimation procedure using a three-parameter model (Gonzalez & Wu, 1999) with the same value function but a two-parameter weighting function $\delta p^{\gamma}/(\delta p^{\gamma} + (1\text{-}\ p)^{\gamma})$, resulting in the equation: $CE = (x_1^{\alpha} \bullet \delta p^{\gamma}/(\delta p^{\gamma} + (1\text{-}\ p)^{\gamma}) + x_2^{\alpha} \bullet (1\text{-}\ \delta p^{\gamma}/(\delta p^{\gamma} + (1\text{-}\ p)^{\gamma})))^{1/\alpha}$. In this case, the weighting function has two components: γ indexes degree of curvature and δ separately indexes overall elevation of the curve. Coefficients of 1 indicate no deviation from identity. For γ, the greater the distance from 1, the more exaggerated the S-shaped over-under ($\gamma < 1$) or under-over ($\gamma > 1$) curve. For δ, the greater the distance from 1, the farther the curve is shifted above ($\delta > 1$) or below ($\delta < 1$) the identity line. While the three-parameter model has not been shown to improve fit over the two-parameter model, it is sometimes preferred for its component interpretation. In particular, γ has been proposed to reflect sensitivity to changes in probabilities (because

[2] We additionally re-ran all analyses after excluding all trials in which a certainty equivalent could not be directly inferred; there were no differences from reported findings. We also compared two groups, based on a median split on number of missing trials, and found no reliable differences between the groups on any analyses.

highly curved functions are flatter for mid-range probabilities, suggesting that these probabilities might be less well differentiated by the individual), while δ might reflect the overall attractiveness of gambling to an individual (because all probabilities are treated as if they are smaller or larger than they actually are; Gonzalez & Wu, 1999). We include the three-parameter model here because numerical competence might be related to only one of the two probability components. In particular, we considered that numerical competence, especially ANS acuity, might be related to curvature (because low acuity could explain low sensitivity to changes in mid-range probabilities) while elevation might arise from other sources.

Five fold cross-validation procedure Before addressing the central question of the relationship between numerical competence and deviation scores, we assessed the predictive accuracy of our CPT models with individual-level coefficients. We used a fivefold cross-validation procedure in which we randomly partitioned the trials into five sets with 20 % of trials per set. We fit the CPT model to 80 % of the trials, and used the obtained parameter estimates to predict the remaining 20 % of trials for each individual for each fold. Predictive accuracy using individual parameter estimates was compared with use of median estimates (parameters set to the median across all participants for the fold) and identity values (parameters set to 1) for both two- and three-parameter models. Prediction error (calculated as root mean squared error or *RMSE*) was calculated for each individual for each model. Table 1 shows *RMSE* averaged across individuals and across the five folds. As shown in the table, for both two- and three-parameter models, *RMSE* was lowest when individual coefficients were used. The three-parameter model did not make better predictions than the two-parameter model, consistent with past findings (Gonzalez & Wu, 1999). Overall, these results provide evidence that, rather than overfitting the data, CPT models that use individual coefficients predict individual responses better than those that use median or identity coefficients, and support the use of individual coefficients.

Parameter estimates We then fit the two- and three-parameter models to the full set of 165 trials to obtain the parameter estimates that we use for our central analyses. We found that the well-known aggregate-level patterns of diminishing sensitivity to outcome values, and of the overweighting of small probabilities and underweighting of large probabilities, emerged in median data using the two-parameter model (α_1 = .84, β = .59) and the three-parameter model (α_2 = .77, δ = .77, γ = .53), as indicated by parameter estimates below 1 for value and β and γ (i.e., curvature-related) probability coefficients, respectively. Individuals with value or curvature-related probability coefficients above 1 were also identified. That is, there were individuals showing increasing sensitivity to outcome values, and individuals who showed underweighting of large probabilities and overweighting of small ones. The majority of participants had coefficients < 1, for value α_1 (73 %) and probability weighting β (98 %); and for α_2 (71 %), and curvature γ (92 %). The majority also had coefficients < 1 for probability elevation δ (76 %) in the three-parameter model, reflecting a general treatment of probabilities as smaller than their actual magnitudes. The coefficient of δ largely does not reflect distortion (alteration of the shape of) the curve, just a change in its overall elevation. However, both for ease of exposition and because changes in elevation typically occur in the context of changes in curvature too, we refer to deviations from 1 in probability coefficients δ and γ in the three-parameter model as collectively characterizing distortion of the probability function.

Deviation measures Our main interest was in the degree of value and probability distortion as reflected in the deviations of α_1 and β (for the two-parameter model) and α_2, δ, and γ (for the three-parameter model) from 1. Deviation scores were created, computed as the absolute value of the parameter estimate minus 1 (e.g., $\beta_d = |1 - \beta|$), with 0 indicating no distortion, and larger numbers reflecting greater distortion. One might expect deviation measures to be highly skewed as a result of this transformation, but they were not (skewness = −.29 – .85), except for α_{1d} where skewness was 2.49 (though even here it was .47 with two extreme data points excluded). As noted earlier, no conclusions differed as a result of using ordinal rather than linear methods (or transformed variables), and so only linear analyses are reported here. Though we do not hold this view, one could also argue that deviations based on coefficients < 1 versus > 1 should not necessarily be treated as equivalent because they might arise for different reasons. Given this possibility, we also report correlation analyses run with raw coefficients, and we graph deviation findings separately for individuals with raw coefficients above versus below 1.

Numerical competence measures

Numeracy scale The mean numeracy score was 9.7 (*SD* = 1.5, range = 6 – 11, skewness = .80). There were no gender differences (*ps* > .100). While Cronbach's alpha for reliability was low for numeracy (α = .54), it was only slightly lower than in past reports (e.g., .57 – .63; Lipkus et al., 2001).

ANS acuity task ANS acuity score was computed as a Weber fraction (Halberda et al., 2008) ranging from 0 to 1 where 0 indicates perfect discrimination. The mean Weber fraction was .15 (*SD* = .04, range = .09–.30, skewness = 1.38), consistent with studies with young adults (e.g., Libertus et al., 2012). There were no gender differences (*ps* > .100); women have

Table 1 Cumulative prospect theory (CPT) model error in certainty equivalent (CE) predictions using the fivefold cross validation procedure

	Coefficient values		
	Identity $\alpha_1 = 1, \beta = 1$ $\alpha_2 = 1, \delta = 1, \gamma = 1$	Median Varied by fold only	Individual Varied by fold and individual
Two-parameter model	52.44	39.56	25.72
Three-parameter model	52.44	39.48	26.46

Note: Shown is root mean squared error (*RMSE*), averaged over all individuals and five folds. As shown, prediction error was lowest for the model using individual coefficients

performed marginally better in some past work but not reliably (e.g., Halberda & Feigenson, 2008). Split-half reliability for the Weber fraction was .68, consistent with past work (e.g., $r = .68$ in Inglis & Gilmore, 2014). The Weber fraction was correlated with percentage of items answered correctly (M = 87 %, skewness = −.69) at $r = .96$ and analyses run with percent correct did not yield different results. The directional sign of correlations involving the Weber fraction is reversed here and throughout to reflect the fact that a negative correlation with the Weber fraction indicates a positive conceptual relationship with ANS acuity.

Correlation between measures No reliable correlation between numeracy and ANS score was found after exclusion of one data point (a score of 6 on numeracy and .30 on ANS acuity) that had a large influence on the correlation ($r(48) = .15$, $p = .286$; otherwise $r(49) = .33$, $p = .018$). This was surprising given that ANS and numeracy scores here were in the same range as those in past work in which a correlation of $r = .47$ was found (Paulsen et al., 2010), but the finding is consistent with at least one report of no correlation between ANS and math achievement test scores in adults (Inglis et al., 2011). We include this data point in further analyses, but we note where its removal would alter conclusions.

Coefficient deviations and numerical competence measures

As shown in Table 2, moderate negative correlations were found between numeracy and all coefficient deviation measures except δ_d. Weak negative correlations between ANS score and deviation measures were observed but were not statistically reliable. Correlations between numerical competence measures and raw coefficients were uniformly weaker than those with deviation measures. Among the correlations between numeracy and deviation scores, no reliable differences in the magnitudes of pairs of correlations were found (Z_Hs < 1.0, ps > .100; except δ_d was lower than γ_d, $Z_H = 1.74$, $p = .08$). We also ran all analyses with the earlier described data point excluded. Results were the same except that for ANS acuity the range of correlations between ANS acuity and deviation scores narrowed (rs = .11 – .17). Additionally, there were no differences in numeracy scores of individuals with coefficients > 1 versus < 1, for any coefficients (ts < 1.00, ps > .100), suggesting that numeracy was not related to direction of deviation.

For each deviation measure, we ran a linear regression for predicting deviation score from numeracy and ANS scores. We first introduced numeracy into the regression equation and then added ANS acuity, to assess whether ANS acuity might

Table 2 Pearson correlations between numerical competence and parameter estimate deviation scores

	Parameter estimate deviation scores						
	Two-parameter model			Three-parameter model			
	α_{1d}	β_d	*PRE*	α_{2d}	δ_d	γ_d	*PRE*
Numeracy	−.33* (−.20)	−.46*** (.42**)	.45*** –	−.28* (.11)	−.17 (−.03)	−.48*** (.38**)	.43** –
ANS acuity	−.19 (−.17)	−.22 (.20)	.11 –	−.05 (−.09)	−.11 (.01)	−.22 (.20)	.27 –

***$p < .001$; **$p < .01$; *$p < .05$. $N = 51$

Notes: A larger approximate number system (ANS) acuity score indicates higher acuity. In parentheses are correlations between numerical competence measures and raw coefficients. Also reported are correlations between numerical competence measures and proportion reduction in error *(PRE)* when using the individual-level cumulative prospect theory (CPT) model to predict certainty equivalents (CEs) relative to use of a grand mean model

Table 3 Linear regression standardized coefficients for predicting parameter estimate deviations

	Parameter estimate deviation scores						
	Two-parameter model			Three-parameter model			
	α_{1d}	β_d	*PRE*	α_{2d}	δ_d	γ_d	*PRE*
Numeracy	−.32*	−.44**	.47***	−.30*	−.15	−.46***	.38**
ANS acuity	.08	.07	.05	−.05	.06	.07	−.14

****p* < .001; ***p* < .01; **p* < .05. *N* = 51

Note. Also reported are coefficients for predicting proportion reduction in error *(PRE)* from numerical competence measures

explain additional variance in this context. Consistent with reported correlations, the improvement in fit of the regression model containing numeracy was statistically reliable (*F*s > 4.00, *p*s < .050; except for δ_d where $F = 1.84$, $p = .181$), and there was no further improvement with the addition of ANS score (*F*-changes < 0.50, *ps* > .500). See Table 3 for regression coefficients for numeracy and ANS acuity when both variables are included in the model.

To visually illustrate the relationship between numeracy and value and probability deviation measures, we median-split participants based on numeracy (*mdn* = 10.5) and graphed value and probability functions for each numeracy group using the three-parameter model. Figure 3 shows value functions (using median parameter estimates) for individuals high and low on numeracy, further divided based on whether the α_2 value coefficient was < 1 or > 1. Figure 4a and b show probability functions for individuals high and low on numeracy further divided based on whether the γ probability coefficient was < 1 or > 1. These figures illustrate that more numerate individuals showed less distortion, whether coefficients were above or below 1.

We also considered the relationship between numerical competence and CPT model fit. Our measure of model fit here was *PRE*, the proportion of reduction in sum of squared error due to use of the CPT model over use of a grand mean model (the latter referring to use of one's average CE to predict all of their CEs, following Gonzalez & Wu, 1999). For the two-parameter model, *PRE* = .92 (*SD* = .06, range = .73 – 1.00; skewness = −1.40) and for the three-parameter model, *PRE* = .92 (*SD* = .06, range = .70 – 1.00, skewness = −1.33), meaning that 92 % of variance in response was explained by the individual-level CPT model. As shown in Table 2, there were moderate positive correlations between numeracy and *PRE;* the model explained more variation in response for more numerate individuals. There were no reliable correlations between ANS acuity and *PRE*, though the correlation for the three-parameter model approached significance (*p* < .100). The improvement in fit of the regression equation containing numeracy for predicting *PRE* was statistically reliable (see Table 3 for regression coefficients; *F*s > 10.00, *p*s < .001), and there was no further improvement with the addition of ANS acuity (*Fs* < .500, *ps* > .500). These findings suggest that not only do more numerate individuals distort values and probabilities to a lesser extent than less numerate individuals,

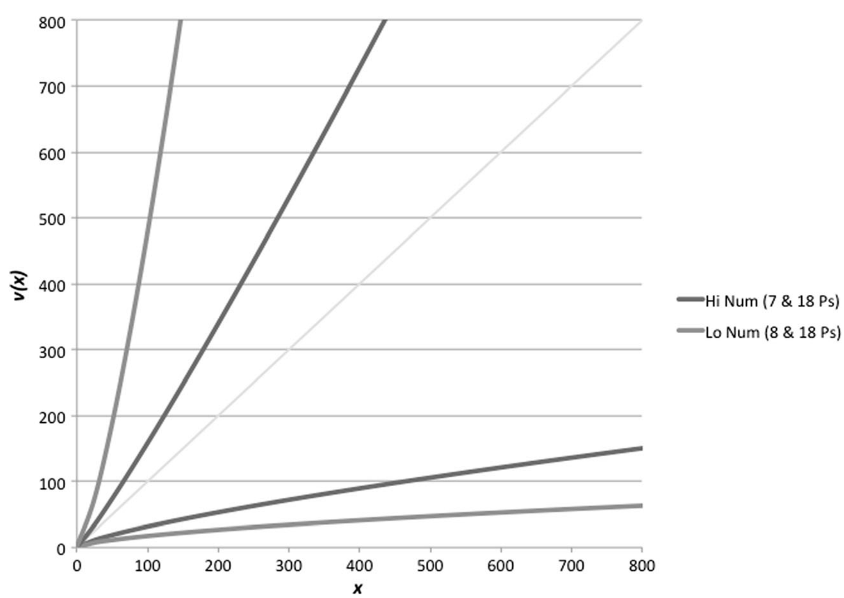

Fig. 3 Value functions for individuals high versus low on numeracy, further split based on whether an individual's value coefficient α was < 1 or > 1. To create each curve, we used the value function $v(x) = x^{\alpha}$, with α set to the median of the parameter estimates for the subgroup of participants. The estimates were from the three-parameter cumulative prospect theory (CPT) model. The figure illustrates that curves for high numeracy individuals were closer to the identity line than those for low numeracy individuals, whether curves were below ($\alpha < 1$) or above ($\alpha > 1$) the line

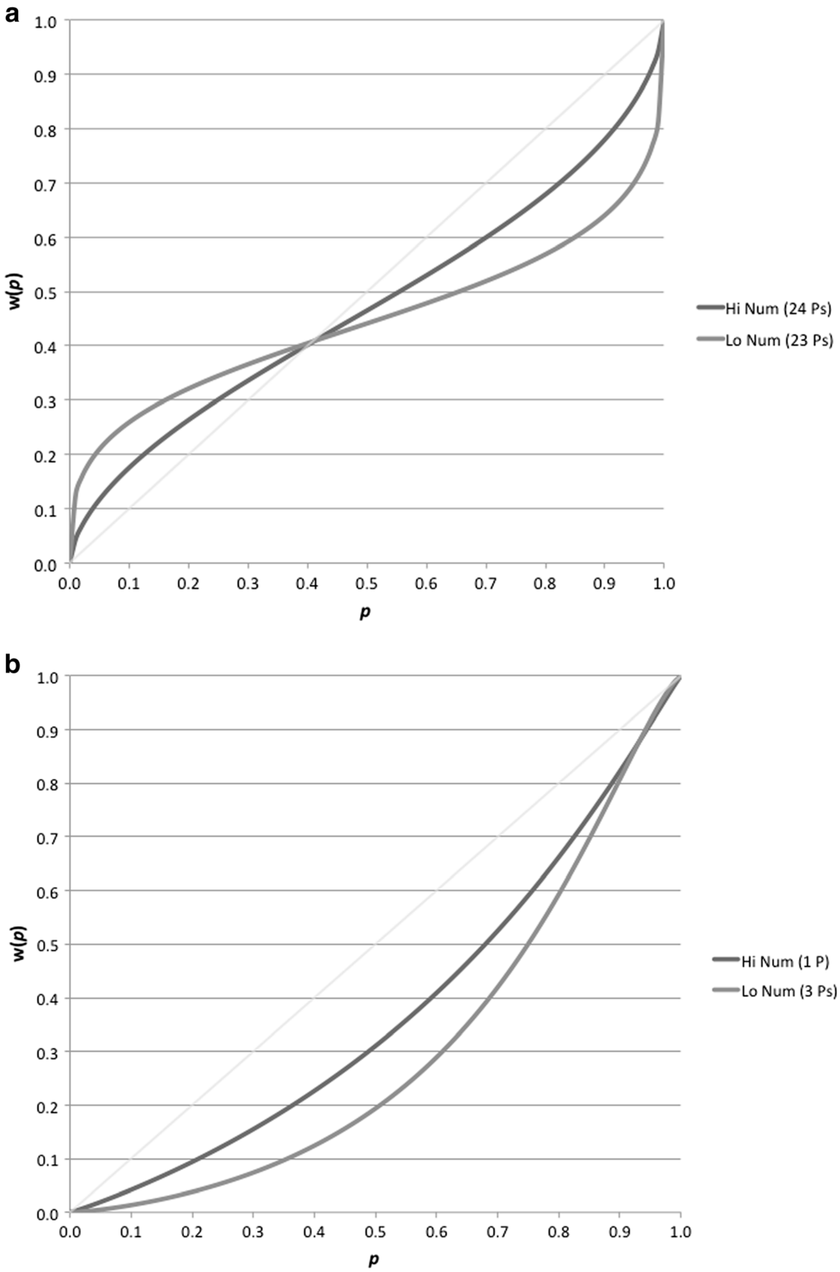

Fig. 4 Probability functions for individuals high versus low on numeracy, further split based on whether an individual's probability curvature coefficient (γ) was < 1 or > 1. To create each curve, we used the probability function $w(p) = \delta p^{\gamma}/(\delta p^{\gamma} + (1- p)^{\gamma}$, with δ and γ set to the medians of the parameter estimates for the subgroup of participants. The estimates were from the three-parameter cumulative prospect theory (CPT) model. The figures illustrate that curves for high numeracy individuals were closer to the identity line than those for low numeracy individuals, regardless of the shape of the curve. (**a**) Curves for individuals with probability curvature $\gamma < 1$, and (**b**) curves for those with a probability curvature $\gamma > 1$

but they also appear to more consistently integrate outcomes and probabilities in accordance with CPT.

Discussion

The findings replicate past work showing that those high in numeracy have less distorted value functions (Schley & Peters, 2014), and provide evidence that they also have less distorted probability functions, consistent with past findings of greater sensitivity to changes in probabilities (see Reyna et al., 2009). While curvature of the probability function was moderately correlated with numeracy in the present study, elevation was only weakly correlated, suggesting different sources of variation. In contrast to the present findings, Schley and Peters (2014) did not find a relationship between numeracy and probability distortion, but they argued that because distortions can be captured through either the value or the probability coefficient, their null finding might reflect issues in the fitting of CPT parameters. There were also some procedural differences across studies (e.g., internet sample, paired choices for gambling task, possibility of losses in their study), though no obvious sources of differences in the probability findings. The present findings are also broadly consistent with those of Winman et al. (2014), who found numeracy to be associated with better-calibrated subjective probability judgments. In future work, it will be important to replicate existing findings using diverse methods for eliciting subjective values and probability weights. At present, the finding from this study that there is a relationship between numeracy and *both* value and probability distortion makes a coherent case that some distortion may be explained by differences in the processing of symbolic numbers rather than by differences unique to the assessment of either subjective value or probability weight.

Correlations between numeracy and deviation measures here were in the range of $|rs| = .30 - .50$, which indicate moderate effect sizes. In contrast, we found no more than weak correlations between ANS acuity and deviation measures, with $|rs| = .05–.22$, which would require a larger sample size ($n = 150$) to reach statistical significance at $\alpha = .05$. From the present findings, it appears that the relationship between ANS acuity and deviation scores is weaker than that between numeracy and these scores, and that ANS acuity does not appear to underlie the relationship between numeracy and number distortions. Perhaps the weak correlation exists because ANS acuity is more generally related to the development of formal math skills (e.g., Halberda et al., 2008), but is not centrally involved in our gambling task. Our ANS findings should perhaps be interpreted with caution, given that we also found only a weak relationship between ANS score and numeracy. Some recent reports have suggested that adult performance across different tasks assumed to measure the ANS may be uncorrelated (Gilmore, Attridge, & Inglis, 2011), that correlations between adult ANS acuity and measures involving symbolic number processing are inconsistent (see Libertus et al., 2012 for a review), and that inconsistencies might be due in part to low reliability of the ANS task (Lindskog, Winman, Juslin, & Poom, 2013). Our findings are compatible with the suggestion of Schley and Peters (2014) that what may differentiate individuals in value and probability distortion is not mental magnitude discrimination acuity *per se,* but rather the ability to map Arabic numerals to mental magnitudes in a consistent and precise manner, which was not tested in the present study. Additionally, our findings are compatible with those of Winman et al. (2014), who found that while numeracy predicted calibration of subjective probability judgments, ANS acuity did not.

Much remains to be explained about potential interactions between individual differences in various aspects of number processing and other influences on the qualitative shapes taken by value and probability functions. Whatever the other factors that contribute to the shapes of value and probability functions, it might be that individuals high in numeracy are less influenced by these factors. Tversky and Kahneman (1992) proposed a psychophysical account of the qualitative shape of these functions whereby decreasing sensitivity to change occurs with increased distance from reference points such as 0 % and 100 % (or $0). It might be that more numerate individuals are better able to judge change at points more distant from the reference point (though this admittedly does suggest involvement of the ANS). Rottenstreich and Hsee (2001), in contrast, proposed that overestimation of small probabilities for positive outcomes arises from the hope that the outcome will occur, while underestimation of large probabilities arises from the fear that it will not occur. In this case, more numerate individuals might be less influenced by affective responses to the meanings of various numbers. Of course, these accounts (and others) must ultimately explain not just the most typical distortion patterns but also why some individuals show opposite patterns of varying magnitudes. Future work related to numerical competence will no doubt contribute to the development of a better understanding of how component systems of numerical cognition support the extraction and integration of magnitudes in judgment and decision tasks, and how number skills interact with other processes (e.g., affective systems) in the service of decision making. Ultimately, translating this understanding into the development of effective practices for teaching decision-related skills is likely to have a meaningful impact on human decision quality.

Acknowledgments This work was supported in part by NSF DRL-0950252 to Hilary C. Barth. We thank Elizabeth Chase, Lily Kaplan, Rachel Santiago, and Elizabeth Tammaro for their help with participant recruitment and data collection, Jessica Taggart for her assistance in manuscript preparation, and Pavel Oleinikov for his statistical expertise.

References

Bernoulli, D. (1738/1954). Exposition of a new theory on the measurement of risk. *Econometrica, 22,* 23–36.

Birnbaum, M. H. (1999). Testing critical properties of decision making on the Internet. *Psychological Science, 10,* 399–407.

Brannon, E. M. (2006). The representation of numerical magnitude. *Current Opinion in Neurobiology, 16,* 222–229.

Bulmer, M. G. (1979). *Principles of statistics*. New York, NY: Dover.

Cokely, E. T., & Kelley, C. M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgment and Decision Making, 4,* 20–33.

Dieckmann, N. F., Slovic, P., & Peters, E. M. (2009). The use of narrative evidence and explicit likelihood by decision makers varying in numeracy. *Risk Analysis, 29,* 1473–1488.

Gilmore, C., Attridge, N., & Inglis, M. (2011). Measuring the approximate number system. *Quarterly Journal of Psychology, 64,* 2099–2109.

Glöckner, A., & Pachur, T. (2012). Cognitive models of risky choice: Parameter stability and predictive accuracy of prospect theory. *Cognition, 123,* 21–32.

Gonzalez, R., & Wu, G. (1999). On the shape of the probability weighting function. *Cognitive Psychology, 38,* 129–166.

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the "number sense": The approximate number system in 3-, 4-, 5-, 6-year-olds and adults. *Developmental Psychology, 44,* 1457–1465.

Halberda, J., Mazzocco, M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with mathematics achievement. *Nature, 455,* 665–668.

Hogarth, R. M., & Einhorn, H. J. (1990). Venture theory: A model of decision weights. *Management Science, 36,* 780–803.

Howell, D. C. (2007). *Statistical methods for psychology.* Belmont, CA: Wadsworth.

Inglis, M., Attridge, N., Batchelor, S., & Gilmore, C. (2011). Non-verbal number acuity correlates with symbolic mathematics achievement: But only in children. *Psychonomic Bulletin & Review, 18,* 1222–1229.

Inglis, M., & Gilmore, C. (2014). Indexing the approximate number system. *Acta Psychologica, 145,* 147–155.

Libertus, M. E., & Brannon, E. (2009). Behavioral and neural basis of number sense in infancy. *Current Directions in Psychological Science, 18,* 346–351.

Libertus, M. E., Odic, D., & Halberda, J. (2012). Intuitive sense of number correlates with scores on college-entrance examination. *Acta Psychologica, 141,* 373–379.

Lindskog, M., Winman, A., Juslin, P., & Poom, L. (2013). Measuring acuity of the approximate number system reliably and validly: The evaluation of an adaptive test procedure. *Frontiers in Cognition, 4,* 510.

Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making, 21,* 37–44.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgments of numerical inequality. *Nature, 215,* 1519–1520.

Pachur, T., & Galesic, M. (2013). Strategy selection in risky choice: The impact of numeracy, affect, and cross-cultural differences. *Journal of Behavioral Decision Making, 26,* 260–271.

Paulsen, D. J., Woldorff, M. G., & Brannon, E. M. (2010). Individual differences in non-verbal number discrimination correlate with event-relate potentials and measures of probabilistic reasoning. *Neuropsychologia, 48,* 3687–3695.

Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological Science, 17,* 408–414.

Peters, E., Slovic, P., Västfjäll, D., & Mertz, C. K. (2008). Intuitive numbers guide decisions. *Judgment and Decision Making, 3,* 619–635.

Reyna, V. F., Nelson, W. L., Han, P. K., & Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making. *Psychological Bulletin, 135,* 943–973.

Rottenstreich, Y., & Hsee, C. S. (2001). Money, kisses, and electric shocks: On the affective psychology of risk. *Psychological Science, 12,* 185–190.

Schley, D. R., & Peters, E. (2014). Assessing economic value: Symbolic-number mappings predict risky and riskless valuations. *Psychological Science, 25,* 753–761.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5,* 297–323.

Von Neumann, J., & Morgenstern, O. (1947). *The theory of games and economic behavior.* Princeton, NJ: Princeton University Press.

Winman, A., Juslin, P., Lindskog, M., Nilsson, H., & Kerimi, N. (2014). The role of ANS-acuity and numeracy for the calibration and the coherence of subjective probability judgments. *Frontiers in Psychology, 5,* 851.

Zeisberger, S., Vrecko, D., & Langer, T. (2012). Measuring the time stability of prospect theory preferences. *Theory and Decision, 72,* 359–386.