Empirical Research

Check for updates

Do Errors on Classic Decision Biases Happen Fast or Slow? Numeracy and Decision Time Predict Probability Matching, Sample Size Neglect, and Ratio Bias

Ryan Corser [1], Raymond P. Voss Jr. [2] (iD) , John D. Jasper [3]

[1] *Owen Graduate School of Management, Vanderbilt University, Nashville, TN, USA.* [2] *Department of Psychology, Purdue University – Fort Wayne, Fort Wayne, IN, USA.* [3] *Department of Psychology, University of Toledo, Toledo, OH, USA.*

Abstract

Higher numeracy is associated with better comprehension and use of numeric information as well as reduced susceptibility to some decision biases. We extended this line of work by showing that increased numeracy predicted probability maximizing (versus matching) as well as a better appreciation of large sample sizes. At the same time, we replicated the findings that the more numerate were less susceptible to the ratio bias and base rate neglect phenomena. Decision time predicted accuracy for the ratio bias, probability matching, and sample size scenarios, but not the base rate scenarios. Interestingly, this relationship between decision time and accuracy was positive for the ratio bias problems, but negative for the probability matching and sample size scenarios. Implications for research on cognitive ability and decision biases are discussed.

Keywords

numeracy, decision bias, probability matching, ratio bias, base rates, sample size neglect

Non-Technical Summary

Background

People's ability to understand and apply mathematical concepts (i.e., numeracy) are important skills for everyday reasoning and decision-making. Previous research shows that higher numeracy is generally associated with better comprehension and use of numeric information. It also predicts greater accuracy on decision scenarios that tend to elicit an easy, but incorrect response rather than a correct one requiring the application of a numeric or statistical concept. When recording people's decision time for these scenarios, researchers sometimes find that quick decisions yield correct answers, while other times slower decision times lead to more correct answers.

Why was this study done?

Our goal was to examine whether numeracy predicted reasoning about probability and small sample sizes in two classic decision scenarios. We also measured decision time to understand how numeracy and decision time predict accuracy in these scenarios.

What did the researchers do and find?

We presented participants with several common decision-making tasks while simultaneously measuring the amount of time it took them to make a decision. We also measured participants' numeracy by having them complete commonly used math tests. Finally, we related their numeracy ability and decision time to their performance on the tasks. Results indicated that higher numeracy was related to increased accuracy on the decision-making tasks. Decision times, however, showed that longer deliberation times were better for some tasks (i.e., reasoning about proportions), while quicker decisions were better for others (i.e., reasoning about probability and small sample sizes).

What do these findings mean?

This work provides evidence that numeracy is related to better decision making amongst previously untested decision tasks and work should continue to test additional tasks. The decision time data provides additional evidence that long deliberate processing is not required to make rational optimal choices. This understanding has potential impacts on several theories of decision-making and rationality.

Highlights

- Increased numeric ability was related to more optimal decision making across several different tasks.
- Additional investigations indicated that decision time was also related to performance, but the direction of results differed across tasks.

The ability to understand and apply mathematical concepts (i.e., numeracy) has been shown to predict the quality of people's decisions as well as how they process decision information (for review see Peters, 2012). Higher numeracy is generally associated with better comprehension and use of numeric information. It is also accompanied by a decreased susceptibility to certain decision biases, such as ratio bias, attribute framing, and base rate neglect (Obrecht & Chesney, 2013, 2016; Peters et al., 2006).

There is less work on how numerical ability relates to probability matching and sample size neglect. The former is a persistent bias in which people's predictions of a probabilistic outcome tend to match the actual probabilities of each outcome rather than choosing the most probable outcome every time (i.e., probability maximizing strategy). For example, when given a 6-sided die with 4 red and 2 green sides, and then asked to guess the outcome of each individual roll in a series, people tend to predict that (over the course of six rolls) 4 will result in red and 2 will result in green, rather than predicting red for each individual roll in the series, which is the optimal solution. Sample size neglect is the tendency to underestimate how small samples have greater sampling error (Kahneman & Tversky, 1982; Yoon et al., 2021). Scenarios assessing this bias have people consider, for example, whether a more skilled squash player will have a better chance of winning if the game is played to 9 or 15 points or if it makes no difference. Typically, people believe that the length of the game should make no difference and thus neglect how a small sample (in this case, a shorter game) is more likely to produce an atypical outcome (i.e., a less skilled player winning).

Additionally, we recorded decision times to examine whether high and low numerates differ in their deliberation time and strategies. Some previous research indicates that highly numerate individuals employ more effortful and elaborative search procedures, spending significantly more time attending to all aspects of the problem, including the numeric components (Jasper et al., 2017). Rooted in the assumptions of dual process theories, it would appear from this work that high numerates are more likely to engage in System 2 based processing. In fact, independent of numeracy, decision time has been linked to accuracy in some tasks. For example, Rubinstein (2013) compared fast and slow responders on several decision problems. For some, such as the Wason card task, the conjunction fallacy, and the framing of roulette gambles, fast responders were less likely to provide optimal solutions than slow responders. Other problems, however, did not show such a difference between fast and slow responders (e.g., the Ellsberg paradox and risky choice framing). Rubinstein concluded that for problems with a clear mistake (e.g., failing to choose a disconfirmation strategy in the Wason card task), there is a positive relationship between decision time and accuracy. In contrast, other problems

aimed at demonstrating how behavior violated normative axioms did not adhere to Kahneman's (2011) thinking fast and slow framework positing that errors are the product of quick, intuitive responses.

Other more recent research, however, argues that there is not strong evidence for the assumptions of the traditional fast/slow dual process theories, with several studies providing evidence that human decision making often violates these traditional theories (De Neys, 2023; De Neys & Pennycook, 2019). For example, several studies have employed a two-response paradigm where participants provide an initial intuitive response to a decision problem, followed by a deliberate response, allowing researchers to directly compare and contrast these responses (Thompson et al., 2011; Thompson & Johnson, 2014). Bago and De Neys (2017) then extended this paradigm to include a direction of change analysis wherein the 4 possible directions of a choice (Incorrect intuition/incorrect deliberation, incorrect intuition/correct deliberation, correct intuition/incorrect deliberation, and correct intuition/correct deliberation) could be analyzed. Using this paradigm, it became clear that participants who generated the correct response during the deliberation stage were actually likely to have already generated the correct response intuitively. Bago and De Neys (2019) then used the Bat and Ball problem from the Cognitive Reflection Test (CRT; Frederick, 2005) to investigate the direction of change. As this test was specifically designed to prime an incorrect intuitive response that needs to be overridden by deliberation, it should provide a more stringent test of participant's correct intuitions. Consistent with the results from the previous study, Bago and De Neys (2019) found that, amongst cases where participants correctly answered the problem, the majority had actually already generated the correct answer intuitively.

Extending this line of work, Raoelison et al. (2020) gave participants several decision tasks while using this two-response paradigm and related it to measures of cognitive capacity. As predicted by the traditional System 1 / System 2 theories, deliberative decision making was related to cognitive capacity. However, opposite of the traditional views, the relationship between cognitive capacity and intuitive responses was actually stronger than the relationship with deliberation. Other research has also pointed to mind-ware instantiation (Rapan & Valerjev, 2021) and mind-ware automatization (Burič, 2023) as potential mechanisms for correct intuitions and more optimal performance on decision tasks. Whatever the mechanisms, it seems clear that decision makers are capable of producing optimal choices intuitively and that cognitive capacity seems to be strongly related to this intuitive processing.

To extend this work, we sought to investigate numeracy in a similar manner as cognitive capacity to determine if high numerates were spending more time on the decision problems than low numerates and making more optimal choices across multiple tasks. This would provide evidence for more effortful processing of numeric information similar to the recent work of Jasper et al. (2017) and a general consensus with the traditional system 1 / system 2 models. On the other hand, if high numerates were making more optimal choices regardless of deliberation time, or in fact quicker than low numerates in some tasks, it would provide evidence for numeracy's relationship to quick optimal intuitions consistent with the work of Raoelison et al. (2020). Though we conducted the study in a similar manner to Rubinstein (2013), we thought it was important to select different biases than those used in that study to investigate the relationship between decision time and accuracy in other contexts.

Lastly, we found it important to explore the interrelationships among several decision biases at once. Specifically, we selected four biases that on the surface involve reasoning with probabilities, which previous research has indicated is often challenging for individuals to understand and can lead to suboptimal decision making (Reyna & Brainerd, 2007). Although these biases are speculated to rely on similar underlying processes, there is a dearth of research testing the relationships among different decision biases. Some existing research suggests that not all decision biases are correlated with one another (Del Missier et al., 2012). Combined with our analysis of decision time, this approach helps illuminate potential mechanisms operating between the tasks and could help explain high numerates' increased accuracy on these tasks.

Study 1

Numeracy and Decision Biases

Numeracy predicts people's ability to transform numerical information into different formats, and therefore, is helpful in reducing certain decision biases. For example, compared to low numerates, high numerates are less likely to treat

equivalent probabilities (e.g., 10% versus 10 out of 100) presented in the same problem differently (Peters et al., 2006). One potential explanation is that high numerates are more likely to transform one numerical format (frequency) into another (percentage) (Peters, 2012).

Other evidence suggests that high and low numerates not only differ with respect to their transformation skills, but also their ability to obtain accurate information from numbers. For example, high numerates have more precise mental number lines than low numerates (Peters et al., 2008). High numerates also tend to use numeric information to guide their decision making and their choices more often conform to normative models, such as expected utility theory (Cokely & Kelley, 2009; Dieckmann et al., 2009; Jasper et al., 2013). In contrast, low numerates tend to be influenced by non-numeric information, such as mood and verbal descriptions (Dieckmann et al., 2009; Peters et al., 2009). Study 1 investigated numeracy's relationship with probability matching and two other previously examined decision biases—ratio bias and base rate neglect. Each of these decision problems will be discussed in turn.

Probability Matching

Probability matching is a probabilistic contingency task in which participants must predict a randomly determined outcome over several trials. Sometimes participants are given the proportion of times each outcome will occur as with the die roll; other times they must learn the outcome proportions through experience. In both formats, people tend to switch back and forth matching the proportions. Mathematically, this matching strategy is inferior to the maximizing strategy of choosing the most probable outcome for all trials. We expected numeracy would be related to performance on this task because previous research has shown that higher cognitive ability (measured by SAT scores or the CRT) is also associated with the use of a maximizing strategy (Gaissmaier et al., 2016; Koehler & James, 2010; Stanovich & West, 2008). While numeracy and other cognitive ability measures are components of the positive manifold of 'g' or general intelligence, research suggests that numeracy explains additional variance in performance beyond these other measures (Cokely & Kelley, 2009; Liberali et al., 2012; Sinayev & Peters, 2015).

Non-Causal Base Rate

De Neys and Glumicic (2008) presented participants with scenarios such as the following:

> In a study 1000 people were tested. Among the participants there were 995 lawyers and 5 engineers.
> Jack is a randomly chosen participant of this study.
> Jack is 36 years old. He is not married and is somewhat introverted. He likes to spend his free time
> reading science fiction and writing computer programs.
> What is most likely?
>> a. Jack is an engineer
>> b. Jack is a lawyer

Here participants receive base rate information (i.e., 995 lawyers and 5 engineers) plus a biographical sketch of one person in the sample that is incongruous with the typical characteristics of the larger group (i.e., lawyers) and more stereotypical of the smaller group. In this scenario, the base rate information and biographical sketch conflict. Participants tend to use the representativeness of the biographical sketch to guide their judgments and choose that Jack is an engineer rather than use the base-rate information, which points to Jack being a lawyer. In contrast, when the base rate information and biographical sketch do not conflict, participants are much more likely to choose the more probable group. Numeracy, indeed, predicts base rate usage independent of other manipulations aimed at de-biasing participants (Obrecht & Chesney, 2016).

Ratio Bias

We sought to conceptually replicate Peters et al. (2006) who showed that low numerates were more likely to exhibit ratio bias or denominator neglect, which is the tendency to prefer a prospect offering a greater absolute number of winning chances (e.g., 9 in 100) over a proportionally, superior prospect (e.g., 1 in 10). Additionally, we tested whether numeracy predicted optimal responses on problems in which the normatively correct answer was also the option with

the largest numerator. On these "no-conflict" trials, we expected that low numerates should perform better and more similarly to high numerates.

Correlations Among Different Decision Biases

Previous research has found that normative responses on one decision problem are moderately correlated with normative responses on other decision problems (Stanovich & West, 1998). For example, performance on ratio bias and probability matching tasks were positively correlated (Koehler & James, 2010). Other studies have used composite measures of typical heuristics and biases as dependent measures. These studies have often shown that cognitive ability measures, such as CRT (Toplak et al., 2011) and SAT (Stanovich & West, 1998) have predicted performance on these composite tasks. Few studies, however, report the relationships between these decision tasks.

In fact, in one of the few studies to report the relationships between decision tasks, Del Missier et al. (2012) found evidence suggesting that while some decisions tasks were in fact correlated with one another, others showed no relationship. This examination, however, did not include all of the decision tasks investigated in this report. As such, a tertiary aim of this study was to explore the interrelationships between the ratio bias, probability maximization, and base rate tasks.

Hypotheses

In sum, we aimed to document that highly numerate individuals were less likely to exhibit probability matching, base rate neglect, and ratio bias (Hypothesis 1). We also tested how numeracy and decision time predicted accuracy on these tasks. Specifically, we hypothesized that higher numeracy and longer deliberation time would be associated with more optimal choices (Hypothesis 2). Finally, we investigated the interrelationships among the three decision biases. Here, we predicted that the purported similarities in underlying mechanisms would result in significant positive correlations among the three decision biases (Hypothesis 3). To investigate these hypotheses, participants completed a probability maximization task, a ratio bias task, and a base-rate neglect task while the computer simultaneously recorded decision times. Participants also completed an objective measure of numeric competency.

Method

Participants

One hundred seventy-two undergraduates (101 females) at a large Midwestern university participated in this study to partially fulfill an introductory psychology course requirement and completed an informed consent at the beginning of the session.

Materials and Procedure

All questionnaire materials were presented using MediaLab[TM] software (Jarvis, 2010). Individually, participants completed the following decision making tasks in the order presented below along with a mood questionnaire (The Brief Mood Introspection Scale; Mayer & Gaschke, 1988), a short form of the Decision Making Inventory (White & Nygren, 2009), and a handedness survey. The results of these were not analyzed for this paper. Please see Corser et al., 2024S for exact wording for each task.

Probability Matching Task — We used a probability matching problem by Stanovich and West (2008), which read:

> Consider the following situation: a die with 4 red faces and 2 green faces will be rolled 6 times. Your task is to predict which color (red or green) will show up once the die is rolled. Imagine that you will win $1 for each color you correctly predict.

Participants then indicated which color would appear on each of six die rolls (e.g., which color is most likely to show up after roll #1? 1 = Red or 2 = Green; . . . which color is most likely to show up after roll #6? 1 = Red or 2 = Green). Participants were not told the outcome of their choices.

Base Rate Task — The base rate problems were borrowed from De Neys and Glumicic (2008). Participants read and listened to six passages that provided base rate and biographical information about a randomly selected person. For each passage, participants indicated which population group the target person was most likely to belong. In the congruent condition, the base rate and biographical information were consistent (e.g., 995 engineers and 5 lawyers; Jack is introverted, reads science fiction and writes computer programs). In the incongruent condition, the two pieces of information conflicted. Participants read either six congruent or six incongruent passages in a fixed order.

Ratio Bias Task — Modified from Stanovich and West (2008), the instructions to the ratio bias task read:

> You will be presented with two trays of red and grey marbles, a large tray that contains 100 marbles and a small tray that contains 10 marbles. The marbles are spread in a single layer in each tray. Imagine you must draw out one marble (without peeking, of course) from either tray. If you draw a red marble, you win a prize. For each problem, select which tray you would prefer to draw from. We will select one of the problems and you will actually get to draw a marble from the tray you selected to see if you win a prize.

There were "conflict trials" in which the smaller tray was the optimal choice and "no-conflict trials" in which the larger tray was the optimal choice, similar to Bonner and Newell (2010). During conflict trials, the base proportions of the small tray (i.e., 10%, 20%, or 30%) were pitted against three different, inferior proportions displayed in the large tray: 7, 8, and 9%; 17, 18, and 19%; 27, 28, and 29%, respectively. For no-conflict trials, the same base proportions of the small tray were presented with the following superior proportions displayed in the large tray: 11, 12, and 13%; 21, 22, and 23%; 31, 32, and 33%, respectively. In total, participants completed 18 trials. The correct answer appeared on each side of the screen 50% of the time and the problems were randomly presented. Finally, participants completed a live version of the task in which they chose between a 10% chance in the small tray and 9% chance in the large tray. If they selected a winning bead they would win one of the three candy bars presented before beginning the task.

Cognitive Ability and Other Measures — Lastly, participants completed Lipkus et al.'s (2001) numeracy scale[1] and the computerized adaptive version of the Berlin numeracy test (Cokely et al., 2012). Participants also answered some demographic questions and reported their standardized American College Testing (ACT) scores if they could recall their score (n = 140). 32 students did not provide their ACT scores. Analyses comparing cases with versus without ACT scores showed no performance differences in terms of numeracy or any of the three decision biases suggesting data were missing completely at random.

Results

Numeracy Measures

As shown in Table 1, the Lipkus scores ranged from 1-10 (M = 6.56; SD = 1.97; Z_{skew} = -2.86), whereas the Berlin scores ranged from 1-4 (M = 1.67; SD = 0.85; Z_{skew} = 6.59). Typical of the Lipkus and the Berlin numeracy tests, their distributions were negatively and positively skewed, respectively. As suggested by Cokely et al. (2012), we summed the scores to obtain a composite numeracy measure that was more normally distributed (M = 8.23; SD = 2.43; Z_{skew} = -1.21). The two numeracy tests were also significantly correlated ($r_{Spearman's}$ = .45, p < .001).

---

1) Question 4 of the Lipkus numeracy scale (which of the following numbers represents the biggest risk of getting a disease? 1 in 100, 1 in 1,000, or 1 in 10) had a typographical error in one of the response options. Instead of presenting a 1 in 10 option, participants saw a 1 in 10,000 option making 1 in 100 the correct answer.

Table 1

Distribution of Numeracy Scores

Score	Lipkus Frequency	%	ANT Frequency	%	Lipkus + ANT Frequency	%
1	1	.6	90	52.3	0	0.0
2	5	2.9	58	33.7	0	0.0
3	9	5.2	15	8.7	6	3.5
4	12	7.0	9	5.2	8	4.7
5	22	12.8			10	5.8
6	23	13.4			20	11.6
7	37	21.5			21	12.2
8	38	22.1			21	12.2
9	18	10.5			30	17.4
10	7	4.1			22	12.8
11					21	12.2
12					10	5.8
13					2	1.2
14					1	0.6
Total	172	100	172	100	172	100

Probability Matching Task

Participants who predicted "red," the more probable outcome, for all six die rolls were classified as using the maximizing strategy ($n = 34$, 20%). Participants who predicted some combination of four reds and two greens were coded as using a matching strategy ($n = 98$, 57%). The remaining participants, whose responses did not fit into the previous two groups, were labeled as other ($n = 40$, 23%).

For subsequent analyses, we combined the matching and other strategies into one group and created high and low numerate groups based on a median split (i.e., numeracy scores ≥ 8.5 classified as high numerate and < 8.5 classified as low numerate). Among the high numerate, 28 out of 86 (32.56%) selected the maximizing strategy, while only 6 out of the 86 (6.98%) low numerate participants selected the same option. Numeracy scores were, indeed, significantly correlated with strategy, $r = .35$, $p < .001$. Individuals with higher numeracy scores were more likely to choose the maximizing strategy.

Decision Time and Probability Matching

Interestingly, higher levels of numeracy were associated with faster response times for answering the probability matching question ($r = -0.28$, $p < .001$). At the same time, less time spent on the task was associated with better performance ($r = -0.33$, $p < .001$). In a logistic regression model, numeracy ($b = 0.46$, $SE = .14$, $p = .001$) and response time[2] ($b = -0.33$, $SE = .09$, $p < .001$) predicted probability maximizing, $\chi^2(2) = 41.27$, $p < .001$, Nagelkerke $R^2 = 0.40$ (see Table 2, Model 1). The effect of numeracy and decision time remained significant even after adding ACT score as a predictor. ACT score, in fact, did not predict die roll strategy beyond numeracy ($b = .13$, $SE = .08$, $p = .14$), $\chi^2(1) = 2.36$, $p = .12$ (see Table 2, Model 2). Finally, the Numeracy × Decision Time interaction was not significant ($b = -0.02$, $SE = 0.05$, $p = .65$; see Table 2, Model 3).

---

[2] A natural log transformation was applied to the response time recorded when deciding to choose red or green for each of the 6 rolls. These transformed times were summed to create the overall time spent answering the probability matching problem. Continuous variables were mean-centered before submitting them to the regression analysis.

Table 2

Unstandardized and Standardized Regression Coefficients for Predicting Probability Maximizing

Predictor	Model 1				Model 2				Model 3			
	b	*SE*	*p*	*OR*	*b*	*SE*	*p*	*OR*	*b*	*SE*	*p*	*OR*
Intercept	-2.15	0.36	< .001		-2.20	0.37	< .001		-2.16	0.36	< .001	
Numeracy (N)	0.46	0.14	.001	1.59	0.34	0.16	0.03	1.41	0.32	0.16	0.05	1.38
Decision time (D)	-0.33	0.09	< .001	0.72	-0.31	0.09	< .001	0.74	-0.28	0.11	0.01	0.76
ACT					0.13	0.08	0.14	1.13	0.12	0.08	0.15	1.13
N × D									-0.02	0.05	0.65	0.98

Base Rate Task

Responses to the six base rate questions were coded '1' if they were base rate-consistent and '0' otherwise. The main dependent variable was the proportion of correct responses (out of 6). The distributions for the percentage of base-rate consistent responses in the congruent and incongruent conditions significantly departed from normality (Congruent: Z_{skew} = -12.40, $Z_{kurtosis}$ = 30.02; Incongruent: Z_{skew} = 4.60, $Z_{kurtosis}$ = 0.83). Consequently, we analyzed the data using Spearman's rank order correlation coefficient to examine the relationship between numeracy and base-rate consistent responses for each condition. Numeracy was significantly related to the number of base-rate consistent responses in the incongruent condition (r_s = .27, p = .01), but not the congruent condition (r_s = -.02, p = .88).

To examine the effect of numeracy on base rate performance controlling for general aptitude, the proportion of correct responses were regressed onto the base rate condition (0 = congruent; 1 = incongruent), mean-centered numeracy, and mean-centered ACT scores using PROCESS (Model 3; Hayes, 2017). The overall model was significant, $F(7,134)$ = 47.87, p < .001, R^2 = .71. Replicating De Neys and Glumicic (2008), participants selected more base-rate consistent responses in the congruent (M = 92%, SD = 15%) versus incongruent condition (M = 28%, SD = 29%), b = -65.36, p < .001 (see Table 3).

Table 3

Unstandardized Regression Coefficients for Predicting Base Rate Performance

Predictor	*b*	*SE*	*p*
Intercept	91.09	2.96	< .001
Base rate condition (BR)	-65.36	4.20	< .001
ACT	-0.54	0.78	.49
Numeracy	0.24	1.29	.85
BR × Numeracy	3.86	1.92	.045
ACT × Numeracy	0.17	0.20	.39
BR × ACT	1.19	1.10	.28
BR × Numeracy × ACT	0.40	0.39	.31

Note. Model R^2 = .71, *MSE* = 456.80.

In addition, the predicted Numeracy × Base Rate Condition interaction was significant, and indicated that increased numeracy predicted greater accuracy for the incongruent trials, but not in the congruent trials. Based on a median-split of numeracy, high numerates (M = 35%, SD = 32%, n = 43) selected more base-rate consistent responses than the low numerate (M = 20%, SD = 23%, n = 45) in the incongruent condition, while there was no difference between high (M = 91%, SD = 18%, n = 43) and low numerates (M = 93%, SD = 11%, n = 41) in the congruent condition.

Decision Time and Base Rate Neglect

Similar to the probability matching data, we examined how numeracy and decision time predicted accuracy on the incongruent base rate scenarios. Numeracy was a significant predictor ($b = 4.39$, $SE = .1.42$, $p = .003$), but not response time ($b = 1.65$, $SE = 1.18$, $p = .17$) or the interaction term ($b = -0.27$, $SE = .49$, $p = .59$), $F(3,84) = 3.52$, $p = .02$, $R^2 = .11$.

Ratio Bias Task

The ratio bias data were analyzed in two separate regression analyses. Numeracy and ACT scores were entered as predictors and the outcome variables were the percentage correct on no-conflict trials (Model 1) and conflict (Model 2) trials. Both numeracy ($b = 3.65$, $SE = 1.37$, $p = .01$) and ACT ($b = 1.77$, $SE = 0.80$, $p = .03$) scores independently predicted accuracy on conflict trials, $F(2, 137) = 15.06$, $p < .001$, $R^2 = .18$. For no-conflict trials, neither numeracy ($b = 1.95$, $SE = 1.26$, $p = .13$) nor ACT scores ($b = 0.51$, $SE = 0.73$, $p = .49$) predicted accuracy, $F(2, 137) = 3.21$, $p = .04$, $R^2 = .05$. Similar results were obtained for the live choice in which participants decided between a 9% or 10% chance of winning a candy bar. Sixty-three out of the 86 high numerate individuals (73%) selected the optimal tray, whereas only 44 out of 84 low numerate individuals (52%) selected the optimal tray. Numeracy scores were significantly correlated with choice, $r = .25$, $p = .001$.[3]

Decision Time and Ratio Bias

Because numeracy only predicted performance on the conflict trials, we analyzed these trials when examining decision time (with natural log transformation applied). Numeracy ($b = 4.47$, $SE = 0.91$, $p < .001$) and decision time ($b = 27.69$, $SE = 5.27$, $p < .001$) predicted accuracy independently, $F(3, 168) = 22.48$, $p < .001$, $R^2 = .29$. Unlike the probability matching results, increased deliberation as signaled by longer decision times was associated with better performance. The Numeracy × Decision Time interaction was not significant, $b = -0.40$, $SE = 2.14$, $p = .85$).

Relationships Among Decision Biases and Cognitive Ability Measures

Performance on the ratio bias, probability matching, and base rate tasks were generally positively correlated. Choosing the maximizing strategy was associated with better performance on the ratio bias ($r_s = 0.36$) and incongruent base rate problems ($r_s = 0.38$). Meanwhile the relationship between ratio bias and incongruent base rate performance did not quite reach statistical significance ($r_s = 0.21$, $p = .06$). See Table 4 displaying the correlations and statistical significance among the cognitive ability measures, the decision tasks and response time.

Numeracy and ACT scores were highly correlated ($r = 0.60$) with each other, but numeracy predicted performance for all three decision biases, whereas ACT scores did not correlate with base-rate performance. This is consistent with studies that have found no relationship between standardized test scores (i.e., SAT) and non-causal base rate reasoning (Stanovich & West, 1998, 2008). As reported above, when numeracy and ACT scores simultaneously predicted performance in a regression analysis, numeracy was the only significant predictor of probability matching and base rate performance. For ratio bias problems, ACT and numeracy independently predicted performance.

Additional Analyses

Because the nature and design of the classic decision tasks are inherently different, direct comparisons are difficult. In an effort to increase the commonalities between the tasks, the ratio bias data was reanalyzed while isolating the first conflict trial, and the base rate data was reanalyzed by selecting only participants who completed incongruent trials and isolating the first trial.

Base Rate Neglect — A total of 88 participants completed the incongruent base rate trials. Their responses on the first base rate scenario were coded as base rate consistent (1) or base rate inconsistent (0). Because the outcome was a dichotomous response, the responses were entered as the outcome in a logistic regression. All predictor variables were again mean centered prior to analysis. First, participant numeracy was entered as a predictor in the analysis. Neither the

---

3) Two participants' choices were not recorded, reducing the sample to 170.

Table 4

Intercorrelations for Cognitive Ability Measures, Decision Tasks, and Response Time (RT)

Measure	1	2	3	4	5	6	7	8
1. Numeracy	–	0.60**	*0.27**	*0.40**	0.35**	*-0.17*	0.21**	-0.28**
2. ACT		–	*0.16*	*0.37**	0.38**	*-0.41**	0.07	-0.38**
3. Base-rate accuracy (% correct incongruent trials)			–	*0.21†*	*0.38**	*0.15*	*0.04*	*-0.10*
4. Ratio bias (% correct conflict trials)				–	0.36**	*-0.05*	0.40**	-0.19*
5. Probability matching					–	*-0.15*	0.03	-0.33**
6. Base-rate RT (incongruent trials)						–	*0.19†*	*0.41**
7. Ratio bias RT (conflict trials)							–	0.23**
8. Probability matching RT								–
N per measure	172	140	88	172	172	88	172	172
M	8.23	22.88	27.65	66.47	0.20	49.13	8.17	46.19
SD	2.43	4.20	4.01	33.29	0.40	2.54	0.43	3.69

Note. Distributions for Variables 3, 4, and 6 were non-normal therefore Spearman rho correlation coefficients reported in italics; the remaining are Pearson correlation coefficients. Matching is a dichotomous variable 0 = matching, 1 = maximizing.

$^{†}p < .10$. $^{*}p < .05$. $^{**}p < .01$.

overall model, $\chi^2(1, N = 88) = 0.962$, $p = .327$, *Nagelkerke* $R^2 = 0.017$; nor numeracy within the model, $B = 0.115$, *Wald*$(1) = 0.942$, $p = .332$, $OR = 1.122$, were significant predictors of choice. To remain consistent with the original models analyzed above, we also entered in overall ACT score as a predictor. Within the participants who received the inconsistent base rate scenarios, 19 participants did not disclose their ACT score, leaving a total of 69 participants for this analysis. Again, the overall logistic model, $\chi^2(2, N = 69) = 1.440$, $p = .487$, *Nagelkerke* $R^2 = 0.031$ was not a significant predictor of choice. Within the model, neither numeracy, $B = 0.176$, *Wald*$(1) = 1.241$, $p = .265$, $OR = 1.193$ nor ACT score, $B = -0.019$, *Wald*$(1) = 0.053$, $p = .817$, $OR = 0.981$, were significant. Finally, the influence of decision time and numeracy were entered as predictors of choice. This overall model did significantly predict choice compared to a constant alone model, $\chi^2(2, N = 88) = 17.836$, $p < .001$, *Nagelkerke* $R^2 = 0.279$. Within this analysis, decision time was a significant predictor of choice $B = 1.788$, *Wald*$(1) = 12.374$, $p < .001$, $OR = 5.977$ with longer decision times leading to more optimal choices. Numeracy, however, was not significant, $B = 0.237$, *Wald*$(1) = 3.109$, $p = .078$, $OR = 1.267$.

Ratio Bias — The original format of the ratio bias stimuli randomly presented participants with both conflict and no conflict trials. The first conflict trial that a participant experienced was isolated and entered into an analysis, predicting optimal jar selection (1) or not (0). As with previous analyses, all predictor variables were mean centered prior to analysis. The overall model significantly predicted choice compared to the constant alone model, $\chi^2(1, N = 172) = 10.354$, $p = .001$, *Nagelkerke* $R^2 = 0.078$. Within the model, numeracy significantly predicted jar choice, $B = 0.211$, *Wald*$(1) = 9.672$, $p = .002$, $OR = 1.235$. To maintain consistency, ACT score was also entered into the analysis. Within this analysis, 32 participants failed to provide their ACT scores, leaving a total of 140 participants for the analysis. Here, the overall model remained statistically predictive of jar choice compared to a constant alone model, $\chi^2(2, N = 140) = 7.742$, $p = .021$, *Nagelkerke* $R^2 = 0.073$, but numeracy dropped to marginal, $B = 0.177$, *Wald*$(1) = 3.637$, $p = .056$, $OR = 1.194$, while ACT was not a significant predictor, $B = 0.025$, *Wald*$(1) = 0.214$, $p = .643$, $OR = 1.025$. Finally, numeracy and decision time were entered as predictors of choice. The overall model statistically predicted jar choice compared to a constant alone model, $\chi^2(2, N = 172) = 12.247$, $p = .002$, *Nagelkerke* $R^2 = 0.092$. Within the model, numeracy statistically predicted choice, $B = 0.211$, *Wald*$(1) = 9.527$, $p = .002$, $OR = 1.235$, but decision time did not predict choice in the model, $B = 0.289$, *Wald*$(1) = 1.863$, $p = 0.172$, $OR = 1.335$.

Relationships Among Decision Biases and Cognitive Ability Measures — Similar to the above analyses, the performance on the first incongruent base rate trial, the performance on the first conflict ratio-bias trial, performance on the probability matching task, ACT score, numeracy score, and response times were correlated to investigate possible relationships between the tasks, see Table 5.

Table 5

Intercorrelations for Cognitive Ability Measures, Decision Tasks, and Response Time (RT) Limited to the First Inconsistent Base-Rate and Ratio-Bias Trials

Measure	1	2	3	4	5	6	7	8
1. Numeracy	–	.60**	.10	.24**	.35**	-.15	.03	-.28**
2. ACT		–	.05	.17*	.38**	-.39**	-.07	-.38**
3. Base-rate accuracy (first incongruent trial)			–	-.02	.20	.40**	-.15	.04
4. Ratio bias (first conflict trial)				–	.25**	.030	.11	-.22**
5. Probability matching					–	–	.00	-.33**
6. Base-rate RT (first incongruent trial)						–	.03	.17
7. Ratio bias RT (first conflict trial)							–	.12
8. Probability matching RT								–
N per measure	172	140 (69)	88	172	172	88	172	172
M	8.23	22.88	0.23	0.57	0.20	8.64	8.99	46.19
SD	2.43	4.20	0.42	0.50	0.40	0.69	0.76	3.69

Note. Distributions for variables 3, 4, and 6 were non-normal therefore Spearman rho correlation coefficients reported in italics; the remaining are Pearson correlation coefficients. Matching is a dichotomous variable 0 = matching, 1 = maximizing. (N) corresponds to the reduction in participants for the inconsistent base-rate condition.

[†]$p < .10$. [*]$p < .05$. [**]$p < .01$.

Discussion

With the analysis of the full data set, numeracy and decision time both predicted accuracy on the ratio bias and probability maximizing tasks, while only numeracy predicted base rate performance. Contrary to our second hypothesis, shorter decision times were associated with accuracy on the probability matching task. One reason why decision time and accuracy may be negatively correlated for probability maximizing is that for those who possess the knowledge or mindware (i.e., understanding the independence of successive trials; Gal & Baron, 1996) the task becomes quite easy and requires choosing the most probable outcome. For the ratio bias, having the mindware (i.e., knowing how to convert frequency information into percentages by attending to the numerator and denominator) is necessary, but not sufficient. Participants also require extra time to compute or compare the probabilities (generally small differences between jars) engaging in sustained inhibition or cognitive decoupling as theorized by Stanovich and West (2008). In the probability maximizing problem, it is sufficient to have the mindware without precise calculations.

When isolating the first incongruent trial of the base rate task and the first conflict trial of the ratio-bias task, a slightly different pattern emerged. Within the base-rate task, only decision time significantly predicted performance. While only numeracy predicted performance in the ratio-bias task. The differing patterns observed could be the result of several factors. First, the base-rate task was presented between subjects, so isolating the incongruent trials effectively reduced the sample size and power of this analysis. Alternatively, recent research suggests that participant behavior can change throughout repeated trials of a task (Li et al., 2022). If mindware instantiation reflects the automatization of the specific process as suggested by other recent research, the differences between the first task and overall averages could reflect the acquisition of this mindware (Raoelison et al., 2021). The results from this study seem to align with the more recent conceptualizations of dual process theories wherein intuition can lead to optimal decision, especially if the mindware is present. Future studies should investigate the differences in decision time and performance across several decisions as well as investigating the impact of numeracy within a two-response paradigm.

Study 2

We conducted a second study to test the robustness of our findings especially with respect to probability maximizing and decision time. Additionally, participants answered a variation of Kahneman and Tversky's (1982) squash decision scenario that tested statistical intuitions about sample size. The scenario substituted squash for table tennis and read:

> A game of table tennis can be played either to 9 or 15 points. Holding all other rules of the game constant, if player A is better than player B, which scoring system will give A a better chance of winning?

Participants then selected which scoring system would best benefit the better player: a) a game to 9 points, b) a game to 15 points, or c) neither scoring system matters. Like the probability maximizing scenario, the squash scenario involves reasoning about probability distributions and how large samples are less likely to yield atypical outcomes. Therefore, in the long run, it is better to choose the more probable outcome or have the skilled player play a longer match. We predicted that numeracy would be positively associated with accuracy in the squash scenario, and accuracy would be negatively related to decision time. For those possessing the mindware or correct statistical intuitions, we expected that they would arrive at the answer more quickly than those who do not or those who fail to apply their understanding of large samples.

Method

Participants

376 undergraduate students from both a midsized southern private university and a midsized midwestern regional public university completed the research for partial course credit. Of these participants, 203 identified as female, 165 identified as male, and the remaining identified as nonbinary or wished not to disclose. Participants completed the study online and provided their consent electronically prior to completing the material. An additional 5 individuals accessed the survey but did not complete enough measures for analysis.

Materials and Procedure

Participants completed the study online using the Qualtrics survey platform. The study consisted of three types of decision scenarios. The first type was the same probability matching task with all six decisions presented on one page (instead of sequentially like Study 1). The second type presented two base rate scenarios (one incongruent and one congruent in a counterbalanced order) from Study 1. The last type was the squash/table tennis scenario from Kahneman and Tversky (1982). Finally, participants completed an eight-item objective numeracy scale (Weller et al., 2013) and demographic questions.

Results

Numeracy

Numerical ability was negatively skewed in this sample with a mean score of 5.02 (SD = 2.18) and median of 5. Applying a median split, we labeled participants scoring 5 or less as lower numerate (n = 198) and the remaining as higher numerate (n = 178). See Table 6.

Probability Matching

A higher proportion of participants choose the maximizing strategy (n = 175, 46.5%) compared to Study 1. Again however, the majority of participants did not choose this optimal strategy, showing either a matching (n = 143, 38%) or alternate strategy (n = 58, 15.4%). As in Study 1, we combined the matching and alternate strategies into one group. There was a significant relationship between dichotomized numeracy and strategy choice, $\chi^2(1, N = 376)$ = 95.34, $p <$ .001. Specifically, among the high numerate, 130 out of 178 (73%) selected the maximizing strategy, while only 45 out of the 198 (22.7%) low numerate participants selected the same option.

Table 6

Distribution of Numeracy Scores in Study 2

Numeracy Score	n	%
0	6	1.6
1	21	5.6
2	29	7.7
3	49	13.0
4	43	11.4
5	50	13.3
6	55	14.6
7	74	19.7
8	49	13.0

As in Study 1, the data was also examined using continuous numeracy. Numeracy scores were positively correlated with the optimal strategy, $r = .525$, $p < .001$ (see Table 7). Higher levels of numeracy were also associated with faster response times for answering the probability matching question ($r = -0.154$, $p < .001$). At the same time, less time spent on the task was associated with better performance ($r = -0.232$, $p < .001$).

Table 7

Intercorrelations for Numeracy, Decision Tasks, and Response Time (RT)

Measure	1	2	3	4	5	6	7	8	9
1. Numeracy	–	0.525**	.194**	.191**	.208**	-.154**	-.123*	0.067	-0.069
2. Matching		–	.012	.182**	.280**	-.232**	-.123*	-.031	-0.101*
3. Base-rate Accuracy (Congruent)			–	.026	.116*	-.030	-.010	-.018	.033
4. Base-rate Accuracy (Incongruent)				–	.144**	-.070	-.037	.037	-.048
5. Sample Size Neglect					–	-.103*	-.094†	-.052	-.139**
6. Matching RT						–	.589**	.461**	.601**
7. BR RT (Incongruent)							–	.505**	.543**
8. BR RT (Congruent)								–	.456**
9. Sample Size Neglect RT									–
M	5.02	0.47	0.90	0.41	0.35	3.58	3.19	2.78	3.18
SD	2.18	0.50	0.30	0.49	0.48	0.63	0.59	0.62	0.74

Note. N = 376. A natural log transformation was applied to all reaction time measures and descriptives for these transformed variables are shown.
$^\dagger p < .10.$ $^* p < .05.$ $^{**} p < .01.$

Finally, numeracy, response time, and the interaction were entered into a logistic regression model predicting strategy. Overall, the model significantly predicted strategy use beyond a constant alone model, $\chi^2(3, N = 376) = 129.75$, $p < .001$, Nagelkerke $R^2 = 0.40$. Within the logistic regression model, the main effects of numeracy ($b = 0.61$, $SE = .07$, $Wald\ \chi^2(1) = 76.55$, $p < .001$, $OR = 1.84$) and response time ($b = -0.68$, $SE = .22$, $Wald\ \chi^2(1) = 9.95$, $p = .002$, $OR = 0.51$) predicted probability maximizing., The interaction term, however, was not significant, ($b = -0.04$, $SE = .12$, $Wald\ \chi^2(1) = 0.12$, $p = .73$, $OR = 0.96$).

Base Rate

Also replicating Study 1, high numerates were more likely to select the base rate consistent response (86 out of 178; 48.3%) than low numerates (67 out of 198; 33.8%) during incongruent trials, $\chi^2(1, N = 376) = 8.14$, $p = .004$. During

congruent trials, the majority of both high numerates (168 out of 178; 94.4%) and low numerates (170 out of 198; 85.9%) selected the base rate consistent response, though numeracy was still significantly related to selection, $\chi^2(1, N = 376)$ = 7.50, p = .006. As with previous tasks, the various bivariate correlations were also investigated. While specifically focusing on incongruent trials, numeracy score was significantly correlated with base rate consistent responses, $r(376)$ = 0.191, $p < .001$, and response time, $r(376)$ = -0.123, $p = .017$. The relationship between base rate consistent responses and response time, however, was not significant, $r(376)$ = -0.037, $p = .477$. When specifically investigating congruent trials, numeracy was also significantly correlated with base rate consistent responses, $r(376)$ = 0.194, $p < .001$, but was not related to response time, $r(376)$ = 0.067, $p = .198$. Base rate consistent response and response time were also unrelated, $r(376)$ = -0.018, $p = .733$.

Finally, a series of logistic regression models were calculated to investigate the impact of numeracy and response time on predicting choices within each type of base rate task. When investigating incongruent trials, the overall model was statistically better at predicting choice than the constant alone model, $\chi^2(3, N = 376)$ = 15.15, $p = .002$, Nagelkerke R^2 = 0.053. Within the model, only numeracy significantly contributed to the model, $Wald \chi^2(1)$ = 12.71, $p = < .01$, OR = 1.20. Neither response time, $Wald \chi^2(1)$ = 0.05, $p = .82$, OR = 0.96, nor the interaction, $Wald \chi^2(1, N = 376)$ = 1.00, p = .317, OR = 1.09 significantly contributed to the model. When investigating congruent trials, the overall model was also statistically better at predicting choice than a constant model alone, $\chi^2(3, N = 376)$ = 17.62, $p < .001$, Nagelkerke R^2 = 0.095. Within the model, centered numeracy again significantly predicted choice, $\chi^2(1, N = 376)$ = 13.62, $p = < .001$, OR = 1.36. Neither response time, $\chi^2(1, N = 376)$ = 2.17, $p = .14$, OR = 0.66, nor the interaction, $\chi^2(1, N = 376)$ = 3.71, $p = .054$, OR = 0.82 significantly predicted choice.

Sample Size Neglect

Participants were asked to select which scoring system would benefit the better player during a game of table tennis. 27.66% of participants (104 out of 376) selected the 9-point game, 34.84% chose the optimal solution of 15 points (131 out of 376), while the remaining 37.5% (141 out of 376) said neither point system mattered. As in the probability maximization problem, the non-optimal solutions were merged to form a single group and were then compared to dichotomous numeracy. Numeracy was significantly related to optimal choices in the task, $\chi^2(1, N = 376)$ = 15.20, $p < .001$. Specifically, 44.94% of high numerates (80 out of 178) selected the optimal choice while only 25.76% of low numerates did so (51 out of 198).

Investigating the bivariate correlations revealed that numeracy was related to optimal choice in the task, $r(374)$ = 0.208, $p < .001$. Response time was also related to optimal choice within the task, $r(374)$ = -0.139, $p = .007$. Numeracy and response time, however, were unrelated, $r(374)$ = -0.069, $p = .181$.

Finally, the impact of numeracy and response time on choice were analyzed using a binary logistic regression. Overall, the model significantly predicted choice above chance levels, $\chi^2(3, N = 376)$ = 23.57, $p < .001$, Nagelkerke R^2 = 0.084. Within the model, numeracy, $Wald \chi^2(1)$ = 14.16, $p < .001$, OR = 1.23, and response time, $Wald \chi^2(1)$ = 6.33, $p = .012$, OR = 0.66 both significantly contributed to the model. The interaction, however, was not significant, $Wald \chi^2(1)$ = 0.152, $p = .696$, OR = 0.972.

Discussion

Largely, the results of Study 2 mirrored those of Study 1. Specifically, higher numeracy was related to improved performance across both the probability maximization and base rate tasks. Extending this line of work, higher numeracy was also related to increased performance on the sample size neglect task. Also consistent with Study 1, the response time was related to performance within the probability maximization task, with quicker responses showing increased probability maximization, but not the base rate task. Interestingly, the sample size neglect problem showed a similar relationship, with quicker response times being related to more optimal choices.

When investigating the relationships between numeracy and response time, there was a significant relationship in the probability maximization and base-rate incongruent tasks, but not in the base-rate congruent task or sample size neglect task. The former was expected due to no conflict in the congruent base-rate problem, while the latter was slightly more surprising, but the task differences between the probability maximizing and sample size problems likely

explain these differences. The non-significant relationship between numeracy and decision time during the sample size problem could reflect that it takes a similar amount of time to apply the incorrect solution for low numerates as it takes for high numerates to apply the correct solution. With everyone having to make one decision, it becomes less surprising that numeracy was not related to decision time for this problem. In contrast, the negative correlation between numeracy and decision time on the probability maximization problem could be due to high numerates knowing they should always choose red (or the more probable outcome each time), which essentially simplifies the decision-making process making it one decision versus six sequential decisions. Low numerates who do not retrieve the correct solution try an alternative strategy requiring them to spend extra time on each of the six rolls as they expect the outcome to somehow match the overall proportions (66.67% red; 33.33% green). These differences between the probability maximization and sample size problem help explain why numeracy would be predictive of decision time in one context but not the other. As with study 1, future research should look to investigate these tasks within a two-response paradigm to determine if the quicker decision times are related to improved intuitions.

General Discussion

Higher levels of numeric competency were associated with more optimum solutions across all four decision tasks confirming Hypothesis 1. We found new relationships among numeracy, probability maximization, and sample size neglect. Replicating previous research, we found a positive relationship between numeracy and performance on the base rate (Obrecht & Chesney, 2016) and ratio bias tasks (Liberali et al., 2012; Peters et al., 2006). We also found support for Hypothesis 3 predicting performance across the different decision tasks was positively correlated.

There was mixed support for Hypothesis 2, which predicted a positive relationship between decision time and normative performance. The hypothesis held true for the ratio bias conflict trials with longer decision times associated with more optimal responding. No significant relationship was found between decision time and accuracy for the base rate scenarios. In contrast, decision time was inversely related to performance on the probability matching and sample neglect scenarios. Together these results show that individuals with higher levels of numeric competency are more likely to possess the mindware necessary to answer these questions correctly and require less time to deploy this knowledge. For ratio bias problems, accuracy depends on having the necessary mindware and performing the calculations. It makes sense that higher accuracy should be achieved by those who take more time to compare the two ratios and this is particularly likely among the high numerate. For base rate problems, the decision time to arrive at an answer is likely similar regardless of whether participants rely more on the base rate or stereotypic information. But it is the high numerates who are more likely to use the base rate versus stereotypic information compared to less numerate individuals.

As described above, the differences in decision times and correlations between tasks can be easily explained by the different decisions made between high and low numerates within each situation, however, potential alternate explanations must also be discussed. As each task did seem to show a slightly different pattern, it could indicate that numeracy relates to several different potential mechanisms. First, it's possible that each task requires different specialized mindware. The relationship between numeracy and optimal choices could then emerge because high numerates are more likely to possess the correct mindware for the tasks. The differences in decision times then could be the result of differences in the actual processing time required by the specific mindware used for each problem (Stanovich, 2016; Stanovich et al., 2011). It's also possible that the automatization of the mindware could lead to reduced decision times and more optimal intuitions. Future research should investigate these possibilities.

Though some researchers question the veracity of dual-process accounts of reasoning (Keren & Schul, 2009; Kruglanski & Gigerenzer, 2011; Osman, 2004), the overwhelming depth of previous evidence for this approach at minimum warrants discussion as a possible mechanism. The traditional System 1 vs System 2 account often emphasizes the quick nature of System 1 processing vs the more deliberate thinking of System 2. This theory would seem to imply that quicker response time would be related to System 1 processing while longer response times would be related to System 2 processing. However, in the conceptualization provided by Evans and Stanovich (2013), the only thing that is required for a System 2 process (referred to as a Type 2 process in this clarified theory) is the involvement of working

memory, the possibility for mental simulation, and the ability to cognitively decouple. This difference may seem quite minute, but it actually provides an important distinction between the theory and traditional understanding. Specifically, this theory allows for the distinction that some Type 1 processes may in fact be longer than a Type 2 response, even if it is less likely, and that both Type 1 and 2 systems are capable of "rational" thought.

Leading directly from these theories, one such dual-process account that offers an interesting explanation of the effects, and could easily be tested, argues that each process relies on a different strategy type, which in turn accounts for the differences in performance and decision times (Markovits et al., 2021). An interesting prediction that can be drawn from this theory is that repeated exposure to similar problems results in the automatization of the process and thus a quicker response time, which seems to be in line with other work in cognitive psychology arguing that experts often rely on intuition/heuristic based processing (Cokely et al., 2018; Kahneman & Klein, 2009).

To investigate potential mechanisms, future research could be conducted to differentiate between these possible explanations. For example, a study could be devised that involved participants completing several trials of a completely novel task, such as completing mathematical operations and making decisions in a system other than base 10. This would ensure that participants come into the task with no preestablished understanding. In this case, each theory would predict different patterns of results. The mindware account would predict that individuals do not possess the mindware to solve the problem, and thus we should see no differences between high and low numerates. If a difference did develop while completing the task, we would expect to see a sharp increase in performance after participants acquired the correct mindware. The traditional System 1 vs System 2 account would predict higher numerates would be more likely to rely on System 2 processing and would thus take longer to answer the novel decisions. Finally, the dual process-strategy account would predict that participants would show a gradual improvement in the decision task and a gradually decreased response time as exposure to the tasks would lead to a gradual automatization. Another future study that could help to illuminate the results would be to conduct a similar study to the one presented above, while providing participants with a time pressure while answering the questions in an attempt to experimentally force a quick answer to highlight the differences in mindware between high and low numerates.

Overall, the present work shows the robustness of the relationship between numeracy and some decision biases. It also shows that deliberation time can be positively or negatively related to normative decision making as well as completely unrelated. These findings should encourage others to examine numeracy and decision time simultaneously as they often independently predict performance and may offer their own insight into the mechanisms of numeracy and decision making.

Supplementary Materials

The Supplementary Materials contain the following items:

- The research data that support the findings of the study (Voss, 2024S)
- The decision tasks used in the study (Corser et al., 2024S)

Index of Supplementary Materials

Corser, R., Voss, R. P., Jr., & Jasper, J. D. (2024S). *Supplementary materials to "Do errors on classic decision biases happen fast or slow? Numeracy and decision time predict probability matching, sample size neglect, and ratio bias"* [Decision tasks used in the study]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.15501

Voss, R. P., Jr. (2024S). *Do errors on classic decision biases happen fast or slow? Numeracy and decision time predict probability matching, sample size neglect, and ratio bias – JNC Submission* [Research data]. Mendeley Data. https://doi.org/10.17632/5z68zdgzk5.1

References

Bago, B., & De Neys, W. (2017). Fast logic? Examining the time course assumption of dual process theory. *Cognition, 158*, 90–109. https://doi.org/10.1016/j.cognition.2016.10.014

Bago, B., & De Neys, W. (2019). The intuitive greater good: Testing the corrective dual process model of moral cognition. *Journal of Experimental Psychology: General, 148*(10), 1782–1801. https://doi.org/10.1037/xge0000533

Bonner, C., & Newell, B. R. (2010). In conflict with ourselves? An investigation of heuristic and analytic processes in decision making. *Memory & Cognition, 38*(2), 186–196. https://doi.org/10.3758/MC.38.2.186

Burič, R. (2023). Acquired knowledge and bias susceptibility: Mindware automatization measured with a two-response paradigm and its relationship with conflict detection. *Studia Psychologica, 65*(4), 320–336. https://doi.org/10.31577/sp.2023.04.883

Cokely, E. T., Feltz, A., Ghazal, S., Allan, J. N., Petrova, D., & Garcia-Retamero, R. (2018). Skilled decision theory: From intelligence to numeracy and expertise. In K. A. Ericsson, R. R. Hoffman, A. Kozbelt, & A. M. Williams (Eds.), *The Cambridge handbook of expertise and expert performance* (2nd ed., pp. 476-505). https://doi.org/10.1017/9781316480748.026

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin numeracy test. *Judgment and Decision Making, 7*(1), 25–47. https://doi.org/10.1017/S1930297500001819

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*(1), 20–33. https://doi.org/10.1017/S193029750000067X

Del Missier, F., Mäntylä, T., & De Bruin, W. B. (2012). Decision-making competence, executive functioning, and general cognitive abilities. *Journal of Behavioral Decision Making, 25*(4), 331–351. https://doi.org/10.1002/bdm.731

De Neys, W. (2023). Advancing theorizing about fast-and-slow thinking. *Behavioral and Brain Sciences, 46*, Article e111. https://doi.org/10.1017/S0140525X2200142X

De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition, 106*(3), 1248–1299. https://doi.org/10.1016/j.cognition.2007.06.002

De Neys, W., & Pennycook, G. (2019). Logic, fast and slow: Advances in dual-process theorizing. *Current Directions in Psychological Science, 28*(5), 503–509. https://doi.org/10.1177/0963721419855658

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*(10), 1473–1488. https://doi.org/10.1111/j.1539-6924.2009.01279.x

Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science, 8*(3), 223–241. https://doi.org/10.1177/1745691612460685

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*(4), 25–42. https://doi.org/10.1257/089533005775196732

Gaissmaier, W., Wilke, A., Scheibehenne, B., McCanney, P., & Barrett, H. C. (2016). Betting on illusory patterns: Probability matching in habitual gamblers. *Journal of Gambling Studies, 32*(1), 143–156. https://doi.org/10.1007/s10899-015-9539-9

Gal, I., & Baron, J. (1996). Understanding repeated simple choices. *Thinking & Reasoning, 2*(1), 81–98. https://doi.org/10.1080/135467896394573

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression based approach.* The Guilford Press.

Jarvis, B. G. (2010). *MediaLab* (Version 2010.2.19) [Computer software]. New York, NY, USA: Empirisoft Corporation.

Jasper, J. D., Bhattacharya, C., & Corser, R. (2017). Numeracy predicts more effortful and elaborative search strategies in a complex risky choice context: A process-tracing approach. *Journal of Behavioral Decision Making, 30*(2), 224–235. https://doi.org/10.1002/bdm.1934

Jasper, J. D., Bhattacharya, C., Levin, I. P., Jones, L., & Bossard, E. (2013). Numeracy as a predictor of adaptive risky decision making. *Journal of Behavioral Decision Making, 26*(2), 164–173. https://doi.org/10.1002/bdm.1748

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *The American Psychologist, 64*(6), 515–526. https://doi.org/10.1037/a0016755

Kahneman, D., & Tversky, A. (1982). On the study of statistical intuitions. *Cognition, 11*(2), 123–141. https://doi.org/10.1016/0010-0277(82)90022-1

Keren, G., & Schul, Y. (2009). Two is not always better than one: A critical evaluation of two-system theories. *Perspectives on Psychological Science, 4*(6), 533–550. https://doi.org/10.1111/j.1745-6924.2009.01164.x

Koehler, D. J., & James, G. (2010). Probability matching and strategy availability. *Memory & Cognition, 38*(6), 667–676. https://doi.org/10.3758/MC.38.6.667

Kruglanski, A. W., & Gigerenzer, G. (2011). Intuitive and deliberative judgments are based on common principles. *Psychological Review, 118*, 97–109. https://doi.org/10.1037/a0020762

Li, Y., Krefeld-Schwalb, A., Wall, D. G., Johnson, E. J., Toubia, O., & Bartels, D. M. (2022). The more you ask, the less you get: When additional questions hurt external validity. *Journal of Marketing Research, 59*(5), 963–982. https://doi.org/10.1177/00222437211073581

Liberali, J. M., Reyna, V. F., Furlan, S., Stein, L. M., & Pardo, S. T. (2012). Individual differences in numeracy and cognitive reflection, with implications for biases and fallacies in probability judgment. *Journal of Behavioral Decision Making, 25*(4), 361–381. https://doi.org/10.1002/bdm.752

Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making, 21*(1), 37–44. https://doi.org/10.1177/0272989X0102100105

Markovits, H., de Chantal, P.-L., Brisson, J., Dubé, É., Thompson, V., & Newman, I. (2021). Reasoning strategies predict use of very fast logical reasoning. *Memory & Cognition, 49*(3), 532–543. https://doi.org/10.3758/s13421-020-01108-3

Mayer, J. D., & Gaschke, Y. N. (1988). The experience and meta-experience mood. *Journal of Personality and Social Psychology, 55*(1), 102–111. https://doi.org/10.1037/0022-3514.55.1.102

Obrecht, N. A., & Chesney, D. L. (2013). Sample representativeness affects whether judgments are influenced by base rate or sample size. *Acta Psychologica, 142*(3), 370–382. https://doi.org/10.1016/j.actpsy.2013.01.012

Obrecht, N. A., & Chesney, D. L. (2016). Prompting deliberation increases base-rate use. *Judgment and Decision Making, 11*(1), 1–6. https://doi.org/10.1017/S1930297500007543

Osman, M. (2004). An evaluation of dual-process theories of reasoning. *Psychonomic Bulletin & Review, 11*(6), 988–1010. https://doi.org/10.3758/BF03196730

Peters, E. (2012). Beyond comprehension: The role of numeracy in judgments and decisions. *Current Directions in Psychological Science, 21*(1), 31–35. https://doi.org/10.1177/0963721411429960

Peters, E., Dieckmann, N. F., Vastfjall, D., Mertz, C. K., & Slovic, P. (2009). Bringing meaning to numbers: The impact of evaluative categories on decisions. *Journal of Experimental Psychology: Applied, 15*(3), 213–227. https://doi.org/10.1037/a0016978

Peters, E., Slovic, P., Västfjäll, D., & Mertz, C. K. (2008). Intuitive numbers guide decisions. *Judgment and Decision Making, 3*(8), 619–635. https://doi.org/10.1017/S1930297500001571

Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological Science, 17*(5), 407–413. https://doi.org/10.1111/j.1467-9280.2006.01720.x

Raoelison, M., Boissin, E., Borst, G., & De Neys, W. (2021). From slow to fast logic: The development of logical intuitions. *Thinking & Reasoning, 27*(4), 599–622. https://doi.org/10.1080/13546783.2021.1885488

Raoelison, M., Thompson, V. A., & De Neys, W. (2020). The smart intuitor: Cognitive capacity predicts intuitive rather than deliberate thinking. *Cognition, 204*, Article 104381. https://doi.org/10.1016/j.cognition.2020.104381

Rapan, K., & Valerjev, P. (2021). Is automation of statistical reasoning a suitable mindware in a base-rate neglect task? *Psihologijske Teme, 30*(3), 447–466. https://doi.org/10.31820/pt.30.3.3

Reyna, V. F., & Brainerd, C. J. (2007). The importance of mathematics in health and human judgment: Numeracy, risk communication, and medical decision making. *Learning and Individual Differences, 17*(2), 147–159. https://doi.org/10.1016/j.lindif.2007.03.010

Rubinstein, A. (2013). Response time and decision making: An experimental study. *Judgment and Decision Making, 8*(5), 540–551. https://doi.org/10.1017/S1930297500003648

PsychOpen GOLD

Sinayev, A., & Peters, E. (2015). Cognitive reflection vs. calculation in decision making. *Frontiers in Psychology, 6*, Article 532. https://doi.org/10.3389/fpsyg.2015.00532

Stanovich, K. E. (2016). The comprehensive assessment of rational thinking. *Educational Psychologist, 51*(1), 23–34. https://doi.org/10.1080/00461520.2015.1125787

Stanovich, K. E., & West, R. F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General, 127*(2), 161–188. https://doi.org/10.1037/0096-3445.127.2.161

Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology, 94*(4), 672–695. https://doi.org/10.1037/0022-3514.94.4.672

Stanovich, K. E., West, R. F., & Toplak, M. E. (2011). The complexity of developmental predictions from dual process models. *Developmental Review, 31*(2-3), 103–118. https://doi.org/10.1016/j.dr.2011.07.003

Thompson, V. A., & Johnson, S. C. (2014). Conflict, metacognition, and analytic thinking. *Thinking & Reasoning, 20*(2), 215–244. https://doi.org/10.1080/13546783.2013.869763

Thompson, V. A., Turner, J. A. P., & Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology, 63*(3), 107–140. https://doi.org/10.1016/j.cogpsych.2011.06.001

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition, 39*(7), 1275–1289. https://doi.org/10.3758/s13421-011-0104-1

Weller, J. A., Diekmann, N. F., Tusler, M., Mertz, C. K., Burns, W. J., & Peters, E. (2013). Development and testing of an abbreviated numeracy scale: A Rasch analysis approach. *Journal of Behavioral Decision Making, 26*(2), 198–212. https://doi.org/10.1002/bdm.1751

White, R., & Nygren, T. (2009). *The decision making styles inventory: Analysis of psychometric properties* [Poster presentation]. 30th annual conference for the Society of Judgment and Decision Making, Boston, MA, USA.

Yoon, H., Scopelliti, I., & Morewedge, C. K. (2021). Decision making can be improved through observational learning. *Organizational Behavior and Human Decision Processes, 162*, 155–188. https://doi.org/10.1016/j.obhdp.2020.10.011