SENSITIVITY OF NUMERATE INDIVIDUALS TO LARGE ASYMMETRY IN OUTCOMES: A REGISTERED REPLICATION OF TRACZYK ET AL. (2018)

Supratik Mondal*
SWPS University of Social Sciences and Humanities

Abstract: *The main aim of this study is to replicate the effect shown by Traczyk et al. (2018), where individuals with higher statistical numeracy, compared to individuals with lower statistical numeracy, employed a more effortful choice strategy when outcomes were meaningful. I hypothesize that participants with higher numeracy will be more likely to make choices predicted by Cumulative Prospect Theory and Expected Value theory (CPT/EV) in high-payoff problems than in low-payoff problems. Data collection was done online by appointing 73 participants. Participants' preference, fluid intelligence, objective and subjective numeracy were measured using thirteen high and eleven low payoff choice problems, International Cognitive Ability Resource (ICAR), Berlin Numeracy Test (BNT), and Subjective Numeracy Scale (SNS), respectively. All the measures mentioned above were presented randomly. Results showed that all participants, in high-payoff condition, on average maximized EV; however, participants with high BNT scores were more likely to make choices consistent with CPT/EV predictions than individuals with low BNT scores. Furthermore, compared to less numerate participants, highly numerate participants were less likely to make choices consistent with CPT/EV predictions in low-payoff condition. Highly numerate individuals adjusted their choice strategy by modulating their response time, indicating their discernible sensitivity towards large asymmetry in payoff. In conclusion, the effect shown by Traczyk et al. (2018) was successfully replicated.*

Key words: *Numeracy; Strategy selection; Risky choice; Priority heuristic; EV maximization strategy; Cumulative prospect theory.*

---

* Supratik Mondal, Center for Research on Improving Decision Making (CRIDM), Faculty of Psychology in Wroclaw, SWPS University of Social Sciences and Humanities, at ul. Ostrowskiego 30b, 53-238 Wrocław. E-mail: smondal@st.swps.edu.pl
The current research was funded by the National Science Centre, Poland (Grant No. 2019/35/O/HS6/04026).

 DOI: 10.7206/DEC.1733-0092.150

WRAŻLIWOŚĆ NA ASYMETRIĘ W WYPŁATACH WŚRÓD OSÓB Z WYSOKIM POZIOMEM ZDOLNOŚCI NUMERYCZNYCH. PREREJESTROWANA REPLIKACJA BADANIA TRACZYKA I IN. (2018)

Streszczenie: *Głównym celem tego badania była próba zreplikowania efektu wykazanego przez Traczyka i in. (2018), zgodnie z którym osoby z wyższym poziomem statystycznych zdolności numerycznych, w porównaniu do osób z niższym poziomem statystycznych zdolności numerycznych, angażują wymagające poznawczo strategie decyzyjne, gdy potencjalne konsekwencje wyboru są znaczące. Postawiłem hipotezę, że osoby z wysokim poziomem statystycznych zdolności numerycznych będą częściej dokonywały wyborów przewidywanych przez skumulowaną teorię perspektywy i model wartości oczekiwanej (CPT/EV) w problemach decyzyjnych z wysokimi wypłatami (tj. znaczącymi konsekwencjami) niż w problemach decyzyjnych z niskimi wypłatami. W badaniu online 73 ochotników podejmowało decyzje w 13 problemach z wysokimi wypłatami oraz w 11 problemach z niskimi wypłatami. Badani rozwiązywali testy mierzące inteligencję płynną, statystyczne zdolności numeryczne oraz subiektywne zdolności numeryczne. Wszystkie miary były prezentowane w losowej kolejności. Wyniki pokazały, że w warunkach wysokiej wypłaty osoby badane dokonywały wyborów maksymalizujących wartość oczekiwaną. Osoby z wysokimi wynikami w teście mierzącym statystyczne zdolności numeryczne częściej dokonywały jednak wyborów zgodnych z przewidywaniami CPT/EV niż osoby z niskimi wynikami w tym teście. Ponadto osoby z wysokim poziomem statystycznych zdolności numerycznych były mniej skłonne do dokonywania wyborów zgodnych z przewidywaniami CPT/EV w warunku niskiej wypłaty. Osoby te dostosowały swoją strategię wyboru do problemu decyzyjnego poprzez zarządzanie czasem przeznaczanym na podjęcie decyzji, co wskazuje, że mogą one charakteryzować się większą wrażliwością na asymetrię w wypłatach. Podsumowując, efekt opisany w badaniu Traczyka i in. (2018) został pomyślnie zreplikowany.*

Słowa kluczowe: *zdolności numeryczne; strategie decyzyjne; ryzykowny wybór; heurystyka pierwszeństwa; strategia maksymalizacji wartości oczekiwanej; skumulowana teoria perspektywy.*

Every moment of our life is bombarded with information condensed in a statistical shell (Rothman et al., 2006). In order to make informed decisions, from giving tips to buying dogecoin, one on a daily basis, needs to comprehend and calculate various kinds of statistical information. Out of numerous decisions we make every day, seldom do we come across decisions that can have a momentous impact (Cirillo & Taleb, 2016; Taleb, 2020). The quality of our judgments in those crucial moments is highly dependent on each individual's level of expertise. For example, a statistician would be much less hopeful (assuming the person has done the math) about their prospect of winning a lottery compared to a person who lacks knowledge in the field and probably would continuously buy lottery tickets year after year with the hope of being a millionaire one day. Hence, expertise modulates human preferences as well as expectations concerning those choices (Reyna, Nelson, Han, & Dieckmann, 2009).

Many studies have examined the effect of individual differences (i.e., numeracy, intelligence, personality traits, and so on) on human preferences (Becker, Deckers, Dohmen, Falk, & Kosse, 2012; Sobkow, Garrido, & Garcia-Retamero, 2020; Traczyk & Fulawka, 2016). For example, patients with low numeracy often do not have accurate perception of benefits and risks associated with unproven medical treatments and interventions. Moreover, less numerate patients fail to accurately consider the reported prevalence rate of diseases, which skews their perception of personal risk of suffering several diseases compared to individuals with high numeracy (Davids, Schapira, McAuliffe, & Nattinger, 2004; Gurmankin, Baron, & Armstrong, 2004).

A significant amount of evidence has also been accumulated regarding the role of numeracy in the context of financial decision-making in the last decade (Jasper, Bhattacharya, & Corser, 2017; Lusardi, 2012; Sobkow, Garrido, & Garcia-Retamero, 2020). In economic theory, optimal behavior under risk and uncertainty is interpreted by variants of expected value or expected utility models. These theories were proposed as a normative rational choice theory, where a rational agent should select an action that is expected to maximize its outcome (for an introduction, see Małecka, 2020). Although, when the normative theory was put to the test, it revealed that humans did not follow normative standards all the time. Therefore, a positive theory of behavior was proposed (i.e., prospect theory and later the cumulative representation of prospect theory) (Kahneman & Tversky, 1979; Thaler, 1980). It is a modification of expected utility theory while keeping the framework of expected utility theory. Prospect theory used another set of psychological variables (i.e., reference point and non-linear weighting function) to address the discrepancies between normative models and human preferences but adhering to the assumption that human preference can be successfully modeled by weighting and summing operations inherited from expected value calculation (Hands, 2015; Tversky & Kahneman, 1992).

Recent results point out that objectively numerate individuals are more sensitive to changes in expected value compared to less numerate individuals (Jasper, Bhattacharya, Levin, Jones, & Bossard, 2013). Furthermore, highly numerate individuals are also more consistent in their preferences regardless of how information is presented compared to individuals with low statistical knowledge. Numerate participants consistently choose riskier options in both decisions-from-description and decisions-from-experience task, providing evidence for their consistency (Ashby, 2017). Interestingly, Cokely and Kelley (2009) showed that despite the positive relationship between numeracy and choices maximizing Expected Value (EV; Bernoulli, 1954; Russell & Norvig, 2002), protocol analyses revealed that individuals with high numeracy did not commonly use EV calculations to arrive at those choices. Instead, retrospective verbalization revealed that participants used elaborative heuristic search processes to make their decisions. Hence, the authors concluded that superior decisions could also be made with simple heuristic processes instead of energy-intensive weighting and summing operations (Gigerenzer, 2007; Gigerenzer & Goldstein, 1996). In addition, one can also interpolate that numerate individuals have a wider repertoire of decision strategies given that numerate individuals' decisions resemble EV maximization strategy even though they are implementing heuristic processes. Put differently, numerate individuals are equipped with the toolkit necessary to use both an energy-intensive EV calculation strategy and can also rely on simple heuristic processes.

In order to test the aforementioned conjecture, Traczyk et al. (2018) conducted a study examining whether people with high objective numeracy modulate their strategy to the consequence of the decision or whether they simply make normatively superior decisions regardless of the magnitude of the outcome. The authors observed that individuals with higher objective numeracy maximized EV and made choices consistent with the predictions of Cumulative Prospect Theory (CPT; Tversky & Kahneman, 1992) when the EV ratio difference between gambles were high. However, in problems where the EV ratio between gambles were low and the potential outcomes were comparable, highly numerate participants adapted their strategy and made choices consistent with the predictions of the Priority Heuristic (PH; Brandstätter, Gigerenzer, & Hertwig, 2006) and, on average, did not maximize EV or made decisions predicted by CPT compered to less numerate participants.

The main aim of the current study is to replicate the effect studied by Traczyk et al. (2018) where people with higher statistical numeracy, in comparison to people with lower statistical numeracy, strategically employ a more effortful choice strategy to make adaptive choices when the choice problem is meaningful. That is, I seek to replicate the effect where participants with high statistical numeracy (i.e., the ability to understand and process numerical and statistical information) will be more likely to make choices consistent with the prediction made by CPT/EV, compared to participants with low statistical numeracy, in high-payoff choice problems but not in low-payoff choice problems.

Method

The current study is a pre-registered close replication study. Complete pre-registration, experimental procedure, sample size estimation (R scripts), data used for analysis, complete analysis (R markdown file), and supplementary materials have been posted on the Open Science Framework (https://osf.io/cje9b/).

Procedure and materials

The current study is investigating the relationship between statistical numeracy and choices under asymmetric payoff conditions. Participants' fluid intelligence, objective numeracy, and subjective numeracy was measured using the International Cognitive Ability Resource (ICAR; Condon & Revelle, 2014), Berlin Numeracy Test (BNT; Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero, 2012), and Subjective Numeracy Scale (SNS; Fagerlin et al., 2007), respectively. Participants also responded to, randomly presented, thirteen high-payoff (outcome difference between two gambles is high) and eleven low-payoff (outcome difference between two gambles is low) choices in binary two-outcome gambles framed as gains.

Participants were instructed to complete the procedure individually during one session. They were further asked not to use a calculator and turn off any devices that might cause inattentiveness during the session. Tasks were designed in Inquisit Web (2016) software and ran in the Prolific platform. Texts were displayed in black font on a light gray background. During one session, after the demographic questionnaire, participants were asked to answer BNT, SNS, ICAR, and choice problems all presented in a random order.[1] The entire procedure was presented in English and took 25 minutes on average (although there were no time constraints) to complete.

Objective statistical numeracy. In the article, objective statistical numeracy is defined as a metric to differentiate individuals proficient in probabilistic and statistical computations (Cokely et al., 2012; Lipkus, Samsa, & Rimer, 2001; Schwartz, Woloshin, Black, & Welch, 1997). The Berlin Numeracy Test was used to measure objective statistical numeracy, risk literacy, and comprehension of probabilistic concepts. A computerized version of the Berlin Numeracy Test was used

[1] Presentation of each measure were done based on a sequence generated randomly.

in the current study consisting of four items presented to participants in a predefined order. Possible scores ranged from 0 to 4 points, with higher scores indicating higher objective statistical numeracy.

Subjective numeracy. Subjective numeracy measures an individual's perception of their numeracy (Fagerlin et al., 2007). In the study, subjective numeracy was measured using an 8-item self-assessment Subjective Numeracy Scale that includes two sub-scales referring to perceived numerical abilities (e.g., "How good are you at calculating a 15% tip?") and preference for numerical and statistical information in daily life (e.g., "How often do you find numerical information to be useful?"). Participants were instructed to choose options that best represent their beliefs about themselves. Possible scores ranged from 0 to 48 points, with higher scores indicating higher subjective numeracy.

Fluid intelligence. Fluid intelligence can be defined by an individual's ability to use reasoning to solve abstract problems without or minimally using prior learning (McGrew, 2021). Four matrix reasoning items from ICAR were used to measure fluid intelligence (Condon & Revelle, 2014). Reasoning problems were presented in the form of three-by-three matrices of elements with one missing element. Participants were instructed to identify the rule underlying the matrix and select one of the six response elements that satisfied the rule. Possible scores ranged from 0 to 4 points, with higher scores indicating higher fluid intelligence.

Choice problems. Being a replication study, almost the same 24 (except one)[2] choice problems were used from the original study. Each choice problem was classified either as a low or high payoff problem based on the EV ratio between gambles. When EV ratio between gambles are low (i.e., 1.5-1.6), choices are considered low-payoff choice problems because playing them repeatedly, on average, would lead to relatively small differences in payoffs irrespective of the chosen gambles. Hence, it is assumed that low-payoff choice problems are trivial because much less consequence is attached when participants are choosing between options. Notwithstanding, when EV ratios between gambles are high (i.e., 5.56-5.87), choices are considered high-payoff choice problems because the EV of each gamble differs significantly. Therefore, it is assumed that high-payoff choice problems are meaningful because choosing any gamble with the higher EV, on average, will lead to much higher payoffs.

These choice problems were explicitly selected to distinguish between the strategy predicted by heuristic strategy (PH) and weighting and summing operation (CPT/

[2] One choice problem differed from the original study because of inappropriate translation of research materials from Polish to English.

DOI: 10.7206/DEC.1733-0092.150

EV). Put differently, choice problems were designed specifically to distinguished between weighting and summing operations embodied by compensatory expectation models (i.e., CPT/EV), and heuristics non-compensatory simple processes relying on trade-offs (i.e., PH). For example,

Gamble A: $5.40 with 29%; $0 with 71%

Gamble B: $9.70 with 17%; $0 with 83%

PH predicts that a decision-maker will choose Gamble A because the difference in minimum gain in probabilities is larger than 10% of the probability scale (i.e., 0.71 vs. 0.83). In contrast, CPT with standard parameters from Tversky and Kahneman (1992) predicts that a decision-maker will choose Gamble B because of its greater CPT value (i.e., 1.38 vs. 1.78). Under the current experimental procedure, CPT predictions are to be the same as EV maximization strategy.[3] Therefore, whatever participants decide, it will match with either PH theory's prediction or will resemble EV maximization strategy/CPT theory's prediction (for more elaboration, see Pachur, Hertwig, Gigerenzer, & Brandstätter, 2013). Regardless of the participant's choice, it does not imply that participants conform to either theory. Instead, it was intended to examine and track changes (if any) in strategy (compensatory to non-compensatory) corresponding to changes in the payoff structure.

Differences between the original and replication study. Unlike the original study (Traczyk et al., 2018), the current study is not using the Need for Cognition Scale (NCS; Cacioppo & Petty, 1982) and Raven's Advanced Progressive Matrices (RAPM; John & Raven, 2003). Instead, the current study uses International Cognitive Ability Resource (ICAR; Condon & Revelle, 2014) as a replacement for RAPM. Second, the mode of instruction between the original and the current replication study is different. The original study was in Polish, but the current replication study is in English. As a consequence, participants in the original study belonged mostly from Poland, but anyone proficient in English can partake in the replication study. This might include a more heterogeneous sample, which may have an influence on the effect. Notwithstanding, to mitigate the effect of instruction difference, I have used the help of Google Translator and bilinguals (proficient with both Polish and English) to make the translation as accurate to the original as feasible without distorting the meaning. Second, to control the potential effect of ICAR introduction, I have randomized presentations of each block to counterbalance the effect.

[3] CPT prediction is aligned with EV maximization strategy for the current set of choice problems. All the gambles CPT predicts also have higher EV (i.e., EV of 1.57 vs. 1.65 for Gamble A vs. Gamble B). Therefore, when participants choose any gamble predicted by CPT, it means choosing a gamble with a higher EV value.

 DOI: 10.7206/DEC.1733-0092.150

Participants

The sample size was determined by simulation using the data collected in the original study. I used the Generalized Linear Mixed Model framework to estimate sample size to obtain 95% statistical power.

Sample size estimation model.

$$log\left[\frac{p(choice=1)}{1-p(choice=1)}\right] = \beta_0 + \beta_1(BNT) + \beta_2(SNS) + \beta_3(NCS) + \beta_4(RAPM) +$$
$$\beta_5(Payoff) + \beta_6(Payoff:BNT) + \beta_7(Payoff:SNS)$$
$$+ subject_{0s} + e_{si} \quad (1)$$

Where,

$$subject_{0s} \sim N(0, \tau_{00}^2),$$
$$e_{si} \sim N(0, \sigma^2).$$

Using the aforementioned model, I calculated the effect size for the interaction term between statistical numeracy (measured by the Berlin Numeracy Test) and payoff (high vs. low). Considering the effect size estimated in the original study (R^2 = .012, d = 0.442; Brysbaert & Stevens, 2018), simulation with 1000 random data points suggest that 75 participants would be sufficient to obtain a significant (p < .05) interaction effect between BNT and payoff with 95% statistical power. (Arnold, Hogan, Colford, & Hubbard, 2011; Johnson, Barry, Ferguson, & Müller, 2015).

Out of seventy-five participants, only two did not finish the entire study; hence their data is eliminated (in accordance with the disclosure made in the pre-registration form). Seventy-three adult volunteers (*age range*: 19-57 years; *mean* = 27 years) participated in an online study for a half-hourly compensation of £4.00 GBP (equivalent to approximately $5.5 USD). Participants were recruited via the Prolific platform, where they were explicitly told that the current study only examines their cognitive abilities, and compensation was by no means based on their performance in the study. Participation in the study was voluntary, and participants could quit the study at any time without any consequences. Participants gave informed consent before starting the study. The departmental ethics committee of SWPS University of Social Sciences and Humanities approved the study protocol.

DOI: 10.7206/DEC.1733-0092.150

Results

In order to examine how measures of individual differences[4] interact with varied payoff structure, Pearson's correlation were calculated and presented in Figure 1. ICAR (a measure of fluid intelligence) is significantly correlated with both BNT (a measure of objective numeracy) and SNS (a measure of subjective numeracy), $r = .32$ ($p = .006$), and $r = .30$ ($p = .011$), respectively; however, BNT and SNS themselves have a negligible correlation of $r = .04$ ($p = .744$) unlike previous studies. Due to this unusual result, Cronbach's α was calculated for both SNS ($\alpha = 0.78$) and BNT ($\alpha = 0.62$). In high-payoff choice problems (EV ratio is relatively high), both SNS and ICAR have a positive correlation of $r = .34$ ($p = .003$), and $r = .32$ ($p = .005$), with choices predicted by CPT/EV, respectively. However, higher scores in BNT are negatively correlated with CPT/EV consistent choices in low-payoff conditions (EV

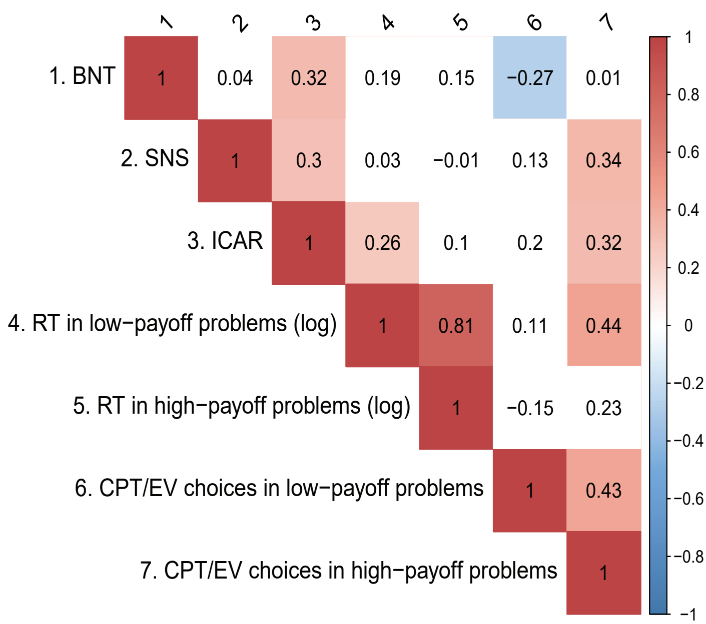

***Figure 1*. Pearson's zero-order correlation coefficient matrix illustrating the relationships between measures used in the study. Significant correlations are marked with color. Here, BNT – Berlin Numeracy Test; SNS – Subjective Numeracy Scale; ICAR – International Cognitive Ability Resource; CPT/EV choices and RT in low and high payoff problems refer to response time and choices consistent with expected value predictions.**

[4] The descriptive table for individual difference measures are in the supplementary material section (https://osf.io/65xdq/).

 DOI: 10.7206/DEC.1733-0092.150

ratio is relatively low) with a coefficient value of r = -.27 (p = .021). Lastly, there is a significant correlation of r = .26 (p = .027) between ICAR and Response Time (RT; the time participants spend in each trial before making a decision.) in low-payoff choice problems, whereas there is negligible correlation of r = .1 (p = .4) between ICAR and RT in high-payoff choice problems.

Next, Mann-Whitney test was conducted to illustrate the difference in choice strategy between participants with varying levels of numeracy. There is a significant difference in choice strategy (W = 72242.5, p = 0.003) between participants with high BNT scores compared to participants with low BNT scores in low-payoff condition with an effect size of -0.09. Similarly, participants with high BNT scores also followed significantly different choice strategy than participants with low BNT scores in high-payoff condition (W = 118852.5, p = 0.006) with an effect size of 0.07.

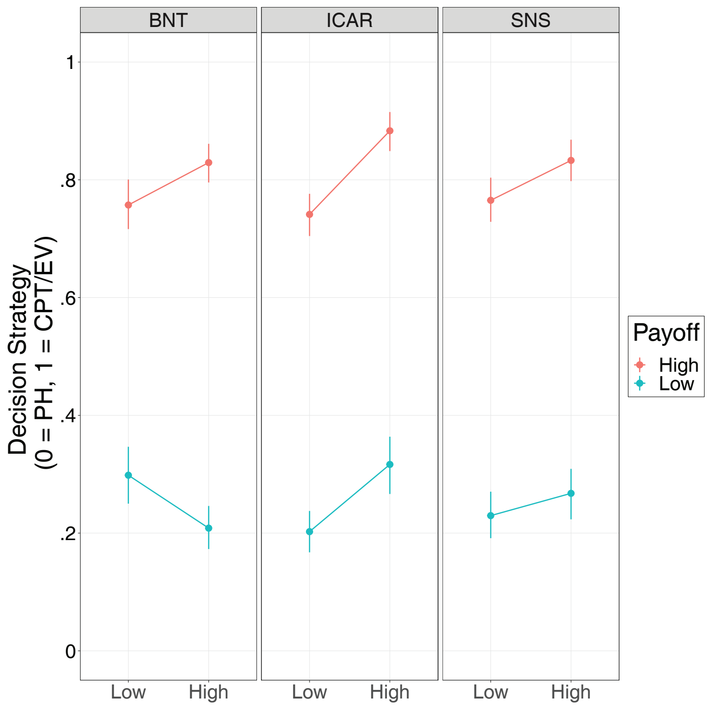

Figure 2. **Decision strategy as a function of BNT, ICAR, and SNS scores. Changes in decision strategy under varied payoff condition illustrated using different colours. Here, BNT – Berlin Numeracy Test; SNS – Subjective Numeracy Scale; ICAR – International Cognitive Ability Resource; 0 = PH refers to choices consistent with Priority Heuristic; 1 = CPT/EV refers to choices consistent with Cumulative Prospect Theory/Expected Value.**

 DOI: 10.7206/DEC.1733-0092.150

Figure 2 affirms the aforementoned results and further communicates how subjective numeracy, fluid intelligence, and objective numeracy predict different aspects of human decision-making. These differences are especially apparent in low-payoff conditions, where participants with higher BNT scores made decisions more consistent with the strategy predicted by PH as opposed to participants with higher SNS and ICAR scores.

This contrasting strategy selection is what motivated me to use a multivariate analysis technique such as Canonical Correlation Analysis (CCA) to test the relationships between variables (i.e., BNT, SNS, ICAR, CPT/EV consistent choices, and payoff) without committing, or minimizing the probability of committing, a Type I error (Sherry & Henson, 2005). I performed bivariate correlation (Pearson r) between Canonical Variate 1 (CV1) and Canonical Variate 2 (CV2). CV1 is a synthetic predictor variable consist of linear combination of BNT, SNS, & ICAR. In contrary, CV2 is a synthetic criterion consist of linear combination of CPT/EV consistent choices, and payoff sensitivity.

Table 1
Loadings on CV1, & CV2

	CV 1	CV 2
BNT	-0.31	0.847
SNS	0.65	0.551
ICAR	0.705	0.311
	CV 1	**CV 2**
Choice	0.817	-0.577
Payoff	-0.005	0.999

Table 1 shows the weights for all three variables that formulate synthetic predictor and two variables that formulate synthetic criterion. The two variables "SNS" and "ICAR" load mostly on CV1. CV1 is also strongly related to the variable "choice". On the other hand, both "BNT" and "Payoff" variable loads mostly on CV2. There is a sign difference between SNS, ICAR, and BNT corroborating earlier evidences from the correlational matrix and Figure 2. In addition, BNT has a negative structure coefficient on CV1 and a high positive structure coefficient on CV2. The sign difference indicates opposite relationships with "Choice" variables.

Table 2
Wilcoxon signed-rank test was conducted for participants with high BNT score and Student's t-test was conducted for participants with low BNT scores

BNT score	Condition 1		Condition 2	t	df	p
High	RT in High-Payoff (log)	-	RT in Low-Payoff (log)	256.500		0.025
Low	RT in High-Payoff (log)	-	RT in Low-Payoff (log)	-0.845	31	0.405

Note. Paired Samples T-Test.

Next, exploratory analysis were performed to observe the relationships between response time, payoff, and numeracy. There is a significant difference in RT between low and high-payoff conditions[5]. Participants' RT was longer (*Mean* = 8.233, *SD* = 0.769) in low-payoff conditions compared to high-payoff conditions

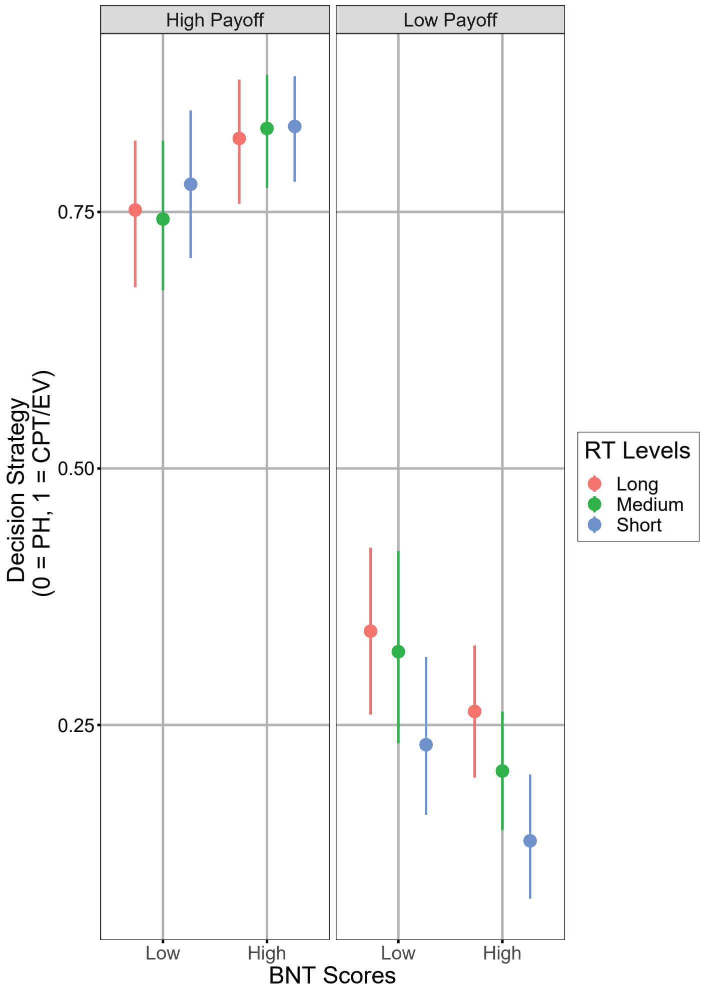

***Figure 3*. Decision strategy as a function of RT and BNT scores under varied payoff condition. Here, BNT – Berlin Numeracy Test; RT – Response time; 0 = PH refers to choices consistent with Priority heuristic; 1 = CPT/EV refers to choices consistent with Cumulative Prospect Theory/Expected Value.**

[5] Tables are in Supplementary Materials section.

(*Mean* = 8.131, *SD* = 0.725). Put differently, participants spent comparatively more time over choice problems when the outcome difference between gambles is low regardless of participants' numeracy level. Furthermore, participants with higher BNT scores have significantly longer (*Mean* = 8.241, *SD* = 0.700) RT compared to participants with low BNT scores (*Mean* = 8.096, *SD* = 0.796). Previous results (i.e., Mann-Whitney test results and Figure 2) point out a significant difference in participants' choice strategy based on their numeracy levels in both high and low payoff condition. Analysis of RT data shows a similar trend but only for highly numerate participants. As Table 2 indicates, there is a significant difference in RT between high and low payoff conditions for only participants with high BNT scores, whereas there is no such difference for participants with low BNT scores. Figure 3 effectively corroborates earlier results while validating the interaction effect between payoff, numeracy, and RT across participants.

Notwithstanding, drawing robust conclusion from the aforementioned results is not ideal due to weaker conditional independence. In order to make effective conclusion from the choice data at hand, multi-level regression analysis using Generalised Linear Mixed Model (GLMM) framework was performed (see, McElreath, 2018). In the model, logit function was used as the link function with four fixed effect parameters, two interaction terms, and one random factor. This basic model was declared in the pre-registered form.

Model 1:

$$log\left[\frac{p(choice=1)}{1-p(choice=1)}\right] = \beta_0 + \beta_1(BNT) + \beta_2(SNS) + \beta_3(ICAR) + \beta_4(Payoff) + \beta_5(Payoff:BNT) + \beta_6(Payoff:SNS) + subject_{0s} + e_{si} \quad (2)$$

Where,

$$subject_{0s} \sim N(0, \tau_{00}^2),$$
$$e_{si} \sim N(0, \sigma^2).$$

Here, β_0, β_1,...,β_4 are fixed effect parameters, while β_5, β_6 capture interaction effects. Lastly, error term (e_{si}) and random effect ($subject_{0s}$) is modeled under normal distribution with mean 0 and variance σ^2, and ${\tau_{00}}^2$, respectively. Model 1 has a Nakagawa marginal and conditional R^2 value of 0.41 and .59, respectively with an *AUC* of *ROC* value of 89.69% (Nakagawa & Schielzeth, 2013).

However, Model 1 fails to account for all possible by-subject dependencies. The experiment has multiple observations per combination of participant and payoff conditions, so this variability in the population will also create clustering in the sample, and $subject_{s0}$ alone cannot capture all this variability because it only allows partici

Table 3
Fixed effects of Model 1

	Estimate	Std. Error	z value	Pr(>| z |)
(Intercept)	0.426	0.238	1.790	0.073
Payoff	2.626	0.208	12.656	<0.001
ICAR	0.374	0.137	2.733	0.006
SNS	0.046	0.025	1.845	0.065
BNT	-0.420	0.330	-1.274	0.203
Payoff:BNT	1.285	0.307	4.191	<0.001
Payoff:SNS	0.036	0.023	1.540	0.124

Table 4
Random effects of Model 1

Groups	Name	Variance	Std.Dev.
Subject	(Intercept)	1.387	1.178

Number of obs: 1752
Groups: Subject = 73

pants to vary around β_0. Hence, random slope was added to allow participants to vary with respect to β_4, our treatment effect. Lastly it is assumed that each participant to have varied preferences among a gamble set; hence Model 1 also lacked a second random effect intercept.

Model 2:

$$log\left[\frac{p(choice = 1)}{1 - p(choice = 1)}\right] = \beta_0 + \beta_1(BNT) + \beta_2(SNS) + \beta_3(ICAR) + \beta_5(RT) + \beta_6(Payoff : BNT : RT) + subject_{0s} + Gamble_{0i} + (\beta_4 + subject_{1s})Payoff_i + e_{si} \quad (3)$$

Where,

$$(subject_{0s}, subject_{1s}) \sim N\left(0, \left[\begin{pmatrix} \tau_{00}^2 & \rho\tau_{00}\tau_{11} \\ \rho\tau_{00}\tau_{11} & \tau_{11}^2 \end{pmatrix}\right]\right),$$

$$gamble_{0i} \sim N(0, \eta_{00}^2),$$

$$e_{si} \sim N(0, \sigma^2).$$

 DOI: 10.7206/DEC.1733-0092.150

Here, as seen in line 2 of Equation 2, I follow standard assumptions in taking this distribution as a bi-variate normal distribution with a mean of (0, 0) and three free parameters: ${\tau_{00}}^2$ (random intercept variance), ${\tau_{11}}^2$ (random slope variance), and $\rho\tau_{00}\tau_{11}$ (the intercept/slope co-variance). Lastly, the intercept of $gamble_{0i}$ is also drawn from a normal distribution with a mean of 0 and variance of σ^2.

Model 2, compared to Bayes Factor (BF) of 0.005 for Model 1, has a higher BF of 200.87 (Schönbrodt & Wagenmakers, 2018). In addition, Model 2 also has a higher Nakagawa marginal and conditional R^2 value of 0.414 and 0.68 respectively, with a higher *AUC* of *ROC* value of 92.44%. Model 1 has a *RMSE*, and log loss scores of .36 and .404 whereas Model 2 has lower *RMSE*, and log loss scores of .33 and .351, respectively. Moreover, Model 2 has a lower deviance score of 1542.1 compared to the deviance score of 1597.5 of Model 1. In light of the above information, Model 2 is much better at explaining variance with much lower *BIC* and *AIC* scores of 1646.7, and 1570.1 respectively, compared to *BIC*, and *AIC* scores of Model 1 1657.3, and 1613.5, respectively.

Table 5
Fixed effects of Model 2

	Estimate	Std. Error	z value	Pr(>| z |)
(Intercept)	0.746	0.323	2.308	0.021
Payoff	3.026	0.434	6.971	<0.001
Medium RT	-0.177	0.170	-1.043	0.297
Short RT	-0.373	0.178	-2.099	0.036
BNT	-0.374	0.381	-0.980	0.327
SNS	0.042	0.029	1.463	0.143
ICAR	0.403	0.154	2.618	0.009
Payoff:Long RT:BNT	1.625	0.537	3.027	0.003
Payoff:Medium RT:BNT	1.580	0.537	2.940	0.003
Payoff:Short RT:BNT	1.371	0.549	2.497	0.013

Table 6
Random effects of Model 2

Groups	Name	Variance	Std.Dev.	Corr
Subject	(Intercept)	1.789	1.338	
	Payoff	1.682	1.297	0.340
Gamble ID	(Intercept)	0.439	0.662	

Number of obs: 1752
Groups: Subject = 73
Gamble ID = 24.

 DOI: 10.7206/DEC.1733-0092.150

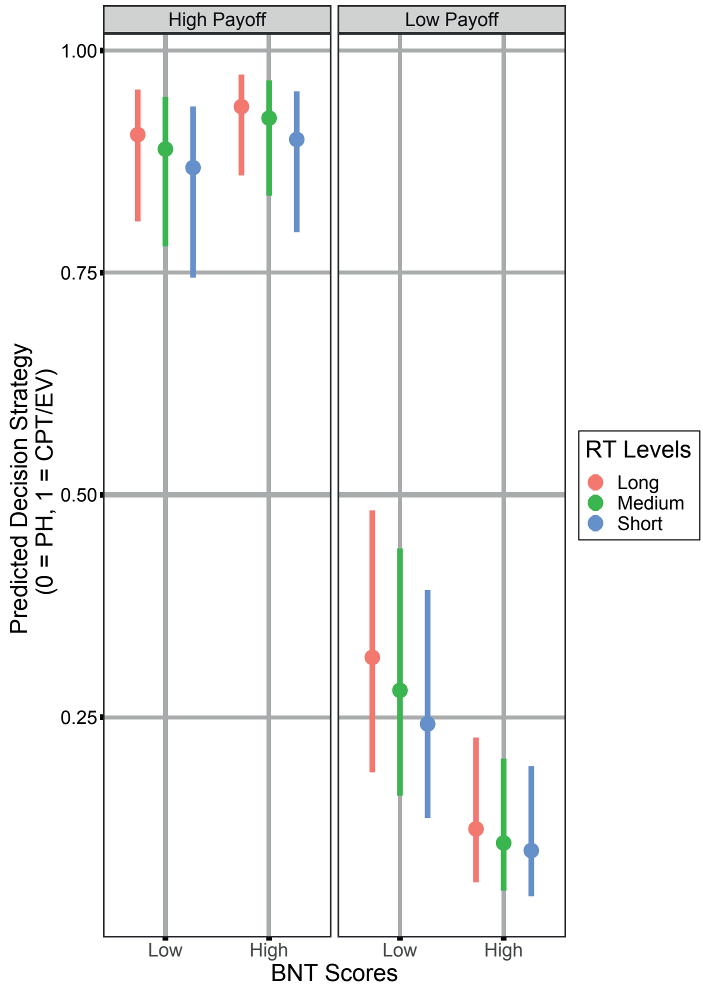

Figure 4. **Predicted decision strategy as a function of RT and BNT scores under varied payoff condition. Here, BNT – Berlin Numeracy Test; RT – Response time; 0 = PH refers to choices consistent with Priority heuristic; 1 = CPT/EV refers to choices consistent with Cumulative Prospect Theory/Expected Value.**

As Table 5 indicates, there is signi cant interaction between participants' objective numeracy and response time in varied payoff conditions. Figure 4 illustrates this interaction more prominently. It was generated by estimating marginal means (predicted values) from Model 2 using *ggeffects* R package (Lüdecke, 2018). High resemblance between simulated data (i.e., Figure 4) and observed data (i.e., Figure 3) indicates robustness of the data collected; at the same time, it attests to the capability of Model 2 to successfully model current data and predict future observations.

 DOI: 10.7206/DEC.1733-0092.150

Discussion

The present study examined whether people with higher statistical numeracy, in comparison to people with lower statistical numeracy, strategically employ a more effortful choice strategy to make adaptive choices when the choice problems are meaningful.

Current finding shows that highly numerate individuals seem to follow compensatory decision strategy embodied by CPT/EV significantly more times when the outcome difference between gambles is high compared to less numerate individuals. However, in low-payoff condition, highly numerate individuals change their strategy and opt for a non-compensatory policy that resembles predictions from PH significantly more times than less numerate individuals. This modulation in strategy between two payoff conditions is present for all participants, but the shift in strategy is substantially distinct for highly numerate individuals than individuals with low numeracy, attesting to highly numerate individuals' acuity to changes in payoff structure. The result sufficiently replicates the finding from the original study and the disclosure made in the pre-registration form. This result is also consistent with earlier work (Estrada-Mejia, de Vries, & Zeelenberg, 2016; Ghazal, Cokely, & Garcia-Retamero, 2014; Horn & Freund, 2021; Pachur et al., 2013; Traczyk et al., 2018).

Furthermore, highly numerate individuals did not only made changes in their decision strategy, but also modulated other aspects (i.e., response time) of decision making in accordance with the environment. Current exploratory analysis indicates that highly numerate individuals significantly modulated the amount of time (RT) they spent on each choice problem based on payoff condition; however, individuals with low statistical numeracy did not adjust their response time in relation to payoff condition. Consequently, highly numerate participants strategically employ a more effortful choice strategy to make adaptive choices when the choice problem is meaningful but choose to opt for a heuristics strategy when choices are less meaningful. The current result corroborates with choice data and attests to highly numerate individuals' discernible sensitivity to payoff structure changes.

Apart from measuring objective numeracy, two other scales were also used to measure subjective numeracy and fluid intelligence. Results show that objective and subjective numeracy explains different aspects of human decision-making (Peters & Bjalkebring, 2015). Individuals with high BNT scores, on average, opted for a strategy predicted by PH in low-payoff conditions, opposite of individuals with high SNS scores. On the other hand, on average, individuals with low SNS scores opted for a strategy predicted by PH in low-payoff conditions, opposite of individuals with high BNT scores. Existing literature suggests that individuals with high objective numer-

 DOI: 10.7206/DEC.1733-0092.150

acy are better equipped to do number comparisons, operations, and calculations, whereas subjective numeracy has been linked to emotional reactions to numbers. Individuals with higher subjective numeracy, unsurprisingly, have more confidence in their ability to perform effectively in numeric tasks and follow EV maximization policy irrespective of the payoff structure (Peters & Bjalkebring, 2015; Traczyk et al., 2018). On the contrary, numerate participants are more sensitive to changes in the environment and make normatively superior decisions adaptively. This contrast helps to explain quantitative differences in predictions from subjective and objective numeracy measures. Although contrary to earlier studies, there is a negligible correlation between SNS and BNT in our study (Sobkow, Olszewska, & Traczyk, 2020; Traczyk et al., 2018). Authors of SNS argued that SNS could replace BNT or could be used as a proxy of BNT (Fagerlin et al., 2007). However, results from the original study and the current study indicate that both scales predict different outcomes hence can not be replaced or be used interchangeably.

Nevertheless, there are some limitations I need to acknowledge. From the data, I could not conclude whether less numerate individuals were making choices that are more consistent with predictions made by CPT/EV theory or they made random choices, given choices are less meaningful in low-payoff conditions. Put differently, for choices in which outcome differences between gambles are low, less numerate participants could have been more inconsistent and switched between strategies (i.e., CPT/EV, PH, random), but such questions are beyond the current experimental purview. Future research can look into this matter. The current replication study used a within-participant design. In future studies, I intend to conduct further experiments with between-participant design to establish the causal relationship between adaptive behavior and numeracy. The current study was conducted in the gain domain. Hence current gambles used in the study may not capture risk attitude of participants adequately. Future work should use gambles from the mixed domain. Lastly, participants had the luxury to spend as much time as they wished for each problem, but in reality, there are always costs associated with time. Hence, I intend to further explore whether numerate individuals continue to follow EV maximization strategy in meaningful circumstances under time pressure.

Conclusion

The current study sufficiently demonstrated that subjective and objective numeracy made quantitatively different predictions under risk. Importantly, I successfully replicated the effect where objectively numerate decision-makers are more sensitive to changes in payoff structure and modulate their strategy to an effortful choice strategy in order to make adaptive choices when the choice problem is meaningful. In summary, I demonstrated that people with higher statistical numeracy, compared to people with lower statistical numeracy, strategically employ more energy-intensive choice strategies to make adaptive choices when the choice problem is meaningful; otherwise, numerate individuals use less effortful heuristic strategies.

References

Arnold, B. F., Hogan, D. R., Colford, J. M., & Hubbard, A. E. (2011). Simulation methods to estimate design power: An overview for applied research. *BMC Medical Research Methodology*, *11*(1), 94. doi: 10.1186/1471-2288-11-94

Ashby, N. J. S. (2017). Numeracy predicts preference consistency: Deliberative search heuristics increase choice consistency for choices from description and experience. *Judgment and Decision Making*, *12*(2), 128–139.

Becker, A., Deckers, T., Dohmen, T., Falk, A., & Kosse, F. (2012). The Relationship Between Economic Preferences and Psychological Personality Measures. *Annual Review of Economics*, *4*(1), 453–478. doi: 10.1146/annurev-economics-080511-110922

Bernoulli, D. (1954). Exposition of a New Theory on the Measurement of Risk. *Econometrica*, *22*(1), 23–36. doi: 10.2307/1909829

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review*, *113*(2), 409–432. doi: 10.1037/0033-295X.113.2.409

Brysbaert, M., & Stevens, M. (2018). Power Analysis and Effect Size in Mixed Effects Models: A Tutorial. *Journal of Cognition*, *1*(1), 9. doi: 10.5334/joc.10

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*(1), 116–131. doi: 10.1037/0022-3514.42.1.116

Cirillo, P., & Taleb, N. N. (2016). On the statistical properties and tail risk of violent conflicts. *Physica A: Statistical Mechanics and its Applications*, *452*, 29–45. doi: https://doi.org/10.1016/j.physa.2016.01.050

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), 23.

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.

Condon, D. M., & Revelle, W. (2014). The international cognitive ability resource: Development and initial validation of a public-domain measure. *Intelligence*, *43*, 52–64. doi: 10.1016/j.intell.2014.01.004

Davids, S. L., Schapira, M. M., McAuliffe, T. L., & Nattinger, A. B. (2004). Predictors of pessimistic breast cancer risk perceptions in a primary care population. *Journal of General Internal Medicine*, *19*(4), 310–315. doi: 10.1111/j.1525-1497.2004.20801.x

Estrada-Mejia, C., de Vries, M., & Zeelenberg, M. (2016). Numeracy and wealth. *Journal of Economic Psychology*, *54*, 53–63. doi: 10.1016/j.joep.2016.02.011

Fagerlin, A., Zikmund-Fisher, B. J., Ubel, P. A., Jankovic, A., Derry, H. A., & Smith, D. M. (2007). Measuring Numeracy without a Math Test: Development of the Subjective Numeracy Scale. *Medical Decision Making*, *27*(5), 672–680. doi: 10.1177/0272989X07304449

Ghazal, S., Cokely, E. T., & Garcia-Retamero, R. (2014). Predicting biases in very highly educated samples: Numeracy and metacognition. *Judgment and Decision Making*, *9*(1), 15–34.

Gigerenzer, G. (2007). *Gut feelings: The intelligence of the unconscious*. Penguin.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*(4), 650–669. doi: 10.1037/0033-295X.103.4.650

Gurmankin, A. D., Baron, J., & Armstrong, K. (2004). The Effect of Numerical Statements of Risk on Trust and Comfort with Hypothetical Physician Risk Communication. *Medical Decision Making*, *24*(3), 265–271. doi: 10.1177/0272989X04265482

Hands, D. W. (2015). *Normative Rational Choice Theory: Past, Present, and Future* (SSRN Scholarly Paper No. ID 1738671). Rochester, NY. doi: 10.2139/ssrn.1738671

Horn, S., & Freund, A. M. (2021). Adult age differences in monetary decisions with real and hypothetical reward. *Journal of Behavioral Decision Making*. doi: 10.1002/bdm.2253

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. doi: 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. doi: 10.1002/bdm.1748

John, & Raven, J. (2003). Raven Progressive Matrices. In R. S. McCallum (Ed.), *Handbook of Nonverbal Assessment* (pp. 223–237). Boston, MA: Springer US. doi: 10.1007/978-1-4615-0153-4_11

Johnson, P. C. D., Barry, S. J. E., Ferguson, H. M., & Müller, P. (2015). Power analysis for generalized linear mixed models in ecology and evolution. *Methods in Ecology and Evolution*, *6*(2), 133–142. doi: 10.1111/2041-210X.12306

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–291. doi: 10.2307/1914185

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. doi: 10.1177/0272989X0102100105

Lüdecke, D. (2018). Ggeffects: Tidy Data Frames of Marginal Effects from Regression Models. *Journal of Open Source Software*, *3*(26), 772. doi: 10.21105/joss.00772

Lusardi, A. (2012). *Numeracy, financial literacy, and financial decision-making* (Tech. Rep. No. w17821). National Bureau of Economic Research. doi: 10.3386/w17821

Małecka, M. (2020). The normative decision theory in economics: A philosophy of science perspective. The case of the expected utility theory. *Journal of Economic Methodology*, *27*(1), 36–50. doi: 10.1080/1350178X.2019.1640891

McElreath, R. (2018). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. Chapman and Hall/CRC. doi: 10.1201/9781315372495

McGrew, K. S. (2021). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, *37*(1), 1–10. doi: 10.1016/j.intell.2008.08.004

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133–142. doi: 10.1111/j.2041-210x.2012.00261.x

Pachur, T., Hertwig, R., Gigerenzer, G., & Brandstätter, E. (2013). Testing process predictions of models of risky choice: A quantitative model comparison approach. *Frontiers in Psychology*, *4*. doi: 10.3389/fpsyg.2013.00646

Peters, E., & Bjalkebring, P. (2015). Multiple numeric competencies: When a number is not just a number. *Journal of Personality and Social Psychology*, *108*(5), 802–822. doi: 10.1037/pspp0000019

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*(6), 943–973. doi: 10.1037/a0017327

Rothman, R. L., Housam, R., Weiss, H., Davis, D., Gregory, R., Gebretsadik, T., ... Elasy, T. A. (2006). Patient Understanding of Food Labels: The Role of Literacy and Numeracy. *American Journal of Preventive Medicine*, *31*(5), 391–398. doi: 10.1016/j.amepre.2006.07.025

Russell, S., & Norvig, P. (2002). Artificial intelligence: A modern approach.

Schönbrodt, F. D., & Wagenmakers, E.-J. (2018). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin & Review*, *25*(1), 128–142. doi: 10.3758/s13423-017-1230-y

Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H. G. (1997, December). The Role of Numeracy in Understanding the Benefit of Screening Mammography. *Annals of Internal Medicine*, *127*(11), 966–972. doi: 10.7326/0003-4819-127-11-199712010-00003

Sherry, A., & Henson, R. K. (2005). Conducting and Interpreting Canonical Correlation Analysis in Personality Research: A User-Friendly Primer. *Journal of Personality Assessment*, *84*(1), 37–48. doi: 10.1207/s15327752jpa8401_09

Sobkow, A., Garrido, D., & Garcia-Retamero, R. (2020). Cognitive Abilities and Financial Decision Making. In T. Zaleskiewicz & J. Traczyk (Eds.), *Psychological Perspectives on Financial Decision Making* (pp. 71–87). Springer, Cham. doi: 10.1007/978-3-030-45500-2_4

Sobkow, A., Olszewska, A., & Traczyk, J. (2020). Multiple numeric competencies predict decision outcomes beyond fluid intelligence and cognitive reflection. *Intelligence*, *80*, 101452. doi: 10.1016/j.intell.2020.101452

Taleb, N. N. (2020). Statistical Consequences of Fat Tails: Real World preasymptotics, epistemology, and applications. *arXiv:2001.10488*.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, *1*(1), 39–60. doi: 10.1016/0167-2681(80)90051-7

Traczyk, J., & Fulawka, K. (2016). Numeracy moderates the influence of task-irrelevant affect on probability weighting. *Cognition*, *151*, 37–41. doi: 10.1016/j.cognition.2016.03.002

Traczyk, J., Sobkow, A., Fulawka, K., Kus, J., Petrova, D., & Garcia-Retamero, R. (2018). Numerate decision makers don't use more effortful strategies unless it pays: A process tracing investigation of skilled and adaptive strategy selection in risky decision making. *Judgment and Decision Making*, *13*(4), 372–381.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323. doi: 10.1007/BF00122574