Note: This is a preprint of a manuscript that has been submitted for publication. Version number: 2, November 2024

The role of numeracy in judgment and decision making: the replication of eleven effects across various numeracy scales

Authors:

Agata Sobkow*, Kacper Jurewicz, Tomasz Zaleskiewicz, Anna Prończuk, Supratik Mondal, Justyna Badecka, Urszula Bartkiewicz, Wiktoria Bobowiec, Michał Bochniak, Karina Chwalibóg, Łukasz Czech, Joanna Daniluk-Malcherek, Maja Domańska, Nikola Dziwusz, Patrycja Falborska, Zuzanna Gazda, Łukasz Gruszczyński, Sara Grzejdziak, Adam Kalus, Katarzyna Knotz, Aleksandra Kocik, Adrian Krzysiek, Tomasz Kuriata, Wiktor Kuriata, Karolina Kusiak, Patrycja Litwinczuk, Daniel Łaciok, Justyna Marek, Laura Mnich, Aleksander Mrowiec, Iga Myślikowska, Aleksandra Ossowska, Paula Panek, Agnieszka Pawelec, Jakub Pozaroszczyk, Alicja Rogowska, Anna Seredyuk, Miłosz Seweryn, Ivan Shakhraichuk, Michał Sielicki, Maria Siemieniec, Aleksandra Stec, Aleksandra Szczepaniak, Oksana Trufanova, Anna Zgajewska, Jakub Traczyk

SWPS University; Faculty of Psychology in Wroclaw, Poland

*Corresponding author:

Agata Sobkow, PhD
SWPS University
Faculty of Psychology in Wroclaw
Ostrowskiego 30b
53-238 Wroclaw
Poland
asobkow@swps.edu.pl

Abstract

Numeracy, representing the ability to understand and process information related to probabilities and numbers, is crucial for accurate decision making. This study evaluates the replicability of eleven effects that underscore the pivotal role of numeracy in judgment and decision making, with successful replication defined as a statistically significant effect in the same direction as reported in the original study. Furthermore, the study explores the potential impact of employing diverse objective numeracy measures on the replicability of the tested effects. Additionally, we investigated correlations with various numeric competencies beyond objective numeracy, including subjective numeracy and approximate numeracy. We ran an online study on Polish-speaking Prolific users (*N* = 209). Six correlational effects were successfully replicated using the same numeracy measures as the original studies (decision rules, financial knowledge, consistency in risk perception, medical risk comprehension, maximizing the expected value, and preference for optimal options). Another two correlational effects were replicated using subjective instead of objective numeracy measures (resistance to sunk cost and financial well-being). Findings regarding the role of numeracy as a moderator variable were mixed. We found no evidence that numeracy influenced the strength of the framing effect, the attractiveness ratings of bets, and performance in the diagnostic inference task. Nevertheless, our exploratory analyses revealed that individuals with higher numeracy found the loss bet more attractive compared to the no-loss bet. Additionally, visual aids improved diagnostic inferences among participants with lower objective numeracy. This comprehensive examination provides valuable insights into the multifaceted interplay between numeracy and decision-making processes. Materials, data, and scripts are available at: https://osf.io/927bx/.

Keywords: numeracy, replication, decision making, judgment, objective numeracy

The role of numeracy in judgment and decision making: the replication of eleven effects across various numeracy scales

Numeracy reflects one's ability to understand and process basic information about probabilities and numbers, which is essential for making informed decisions in everyday situations. Prior studies have consistently highlighted numeracy as a robust predictor of superior decision outcomes (Garcia-Retamero et al., 2019; Peters, 2020; Reyna et al., 2009; Reyna & Brust-Renck, 2014). Individuals with higher numeracy demonstrate more accurate estimations of treatment efficacy, risks associated with diseases or side effects, and a comprehensive understanding of health-related numerical information (Lipkus & Peters, 2009; Nelson et al., 2008; Peters, 2012; Peters & Levin, 2008). The disparities between individuals with higher vs. lower numeracy translate into tangible real-life outcomes. Higher numeracy is linked to a higher quality of life, reduced morbidity, a lower risk of hospitalization, decreased prevalence of comorbidity (Apter et al., 2006, 2009; Garcia-Retamero, Andrade, et al., 2015; Ginde et al., 2008), and enhanced adherence to treatment regimens (Estrada et al., 2004; Waldrop-Valverde et al., 2010). Additionally, high numeracy is associated with superior financial decision making in paradigmatic risk tasks such as making choices in monetary lotteries resulting in higher financial gains and/or lower losses (Ghazal et al., 2014; Mondal et al., 2024; Mondal & Traczyk, 2023) and real-life situations, which translates into concrete financial outcomes such as personal wealth (Estrada-Mejia et al., 2016) or subjective financial well-being (for a review, see Garcia-Retamero et al., 2019; Peters, 2020; Sobkow, Garrido, et al., 2020).

Multiple numeric competencies

However, numeracy is not a singular, unitary concept. Peters and Bjalkebring (2015) expanded the scope of numerical abilities by incorporating subjective and approximate numeracy alongside objective numeracy. Objective numeracy can be defined as the ability to comprehend and effectively use statistical and probabilistic information in everyday situations. In contrast, approximate numeracy involves an intuitive capability to perceive and manipulate numbers and to map symbolic representations to magnitudes. Finally, subjective numeracy captures a decision maker's own confidence in their numerical abilities and preference for numerical format of information. Different numeric competencies have been shown to affect decision making in various ways. Specifically, people with high subjective numeracy use more complex decision-making strategies regardless of the importance of the problem (Traczyk et al., 2018) and feel less negative emotions when thinking about mathematical tasks (Peters & Bjalkebring, 2015). However, such individuals sometimes also make less optimal everyday decisions due to being overly confident (Peters, Tompkins, et al., 2019; Sobkow, Olszewska, et al., 2020).

Replications of numeracy effects

Despite the increasing popularity of numeracy-related research, the examination into the replicability of these effects remains limited (see also Cipora & Soltanlou, 2021 for the replications in numerical cognition; and Sobkow et al., in press for the review of the currect state of replications in JDM). This limitation hinders the self-correcting process initiated in response to the replication crisis, as observed in social and personality psychology (Vazire & Nosek, 2023). To our knowledge, only a handful of replication projects have revisited the relationship between numeracy and different aspects of decision making (Mondal, 2021; Persson et al., 2021; Sobkow et al., 2022; Zhu & Feldman, 2023). For example, Sobkow et al. (2022) and Mondal (2021) undertook preregistered replications of effects initially observed by Traczyk et al. (2018). These studies confirmed that individuals with higher objective numeracy adjusted their decision strategy based on the importance of a decision problem, but the relationship between subjective numeracy and strategy selection was only partially replicated.

Another previous project focused on a direct replication of seminal studies by Peters et al. (2006), pioneering the incorporation of numeracy into the field of judgment and decision making. In this endeavor, Zhu and Feldman (2023) found consistent support for the original findings regarding numeracy's interactions with three decision-making effects: weaker framing effect (Study 1), weaker ratio bias (Study 3), and stronger bets effect (Study 4). However, the exploration of the interaction between objective numeracy and the frequency-percentage effect (Study 2) yielded mixed results, making the understanding of numeracy's role in decision-making dynamics more complex.

The current study

Due to the scarcity of studies attempting to replicate the effects of numeracy on decision making, coupled with mixed observed results, our study aims to broaden the understanding of this topic. We intend to investigate the replicability of eleven effects (see the method section for the description of selection criteria) by utilizing various numeracy scales.

In the current replication study, we focused on the original studies that demonstrated associations between objective numeracy and different factors related to judgment and decision making. In particular, these studies documented that individuals with higher numeracy exhibited more accurate application of decision rules (Skagerlund et al., 2021), greater resistance to sunk cost (Skagerlund et al., 2021), enhanced financial knowledge (Peters, Tompkins, et al., 2019), highly consistent risk perception (Skagerlund et al., 2021), a propensity for choices maximizing expected value (Sobkow et al., 2020), superior medical risk comprehension (Rolison et al., 2020; Study 1), and a tendency to perceive the loss bet more attractive than the no-loss bet (Peters et al., 2006; Study 4). Moreover, they declared higher financial well-being (Peters, Tompkins, et al., 2019). Conversely, individuals with lower numeracy demonstrated a stronger framing effect (Peters et al., 2006; Study 1) and a preference for suboptimal options (Peters et al., 2006; Study 3). Additionally, previous research found that visual aids improved the interpretation of the predictive value of diagnostic tests among individuals with lower numeracy (Garcia-Retamero, Cokely, et al., 2015).

Furthermore, we expanded our project by investigating whether the use of objective numeracy measures other than those used originally could affect the replicability of the chosen effects. In the original studies, despite the researchers' intent to measure a similar concept—objective numeracy—various measures were employed, such as the Berlin Numeracy Test (BNT: Cokely et al., 2012), the scale developed by Lipkus et al. (2001), or the Abbreviated Numeracy Scale (ANS: Weller et al., 2013). These measures vary in difficulty and item content. For instance, the Berlin Numeracy Test involves only four challenging items (e.g., "*Out of 1,000 people in a small town 500 are members of a choir. Out of these 500 members in the choir 100 are men. Out of the 500 inhabitants that are not in the choir 300 are men. What is the probability that a randomly drawn man is a member of the choir?*") while the Lipkus et al. scale (2001) incorporates relatively easy ones (e.g., "*The chance of getting a viral infection is .0005. Out of 10,000 people, about how many of them are expected to get infected?*"). In addition, the Abbreviated Numeracy Scale was developed using the item response theory approach and consists of items from commonly used numeracy tests (including the scale developed by Lipkus et al., 2001), as well as the Cognitive Reflection Test (e.g., "*In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?*" Frederick, 2005). As a result, these measures may tap into slightly different cognitive mechanisms or be more suitable for studying different populations (i.e., general vs. highly educated populations).

Finally, recognizing that numeracy is not a singular entity and that various types of numeric competencies may complement each other (Patalano et al., 2020; Peters, 2020; Peters & Bjalkebring, 2015; Peters & Shoots-Reinhard, 2022; Schley & Peters, 2014; Sobkow, Olszewska, et al., 2020), we also opted to delve into the effects of subjective numeracy (measured using self-report questions such as: "*How good are you at working with percentages?*") and approximate numeracy (measured using the symbolic-number mapping task in which participants are asked to locate numbers on a number line anchored with 0 and 1000).

The effects selected for replication

Below, we provide a concise overview of the effects selected for replication, accompanied by a brief outline of the original studies' methodology and the primary results that served as the focus of our replication efforts.

Numeracy and application of decision rules, resistance to sunk cost, and consistency in risk perception (Skagerlund et al., 2021)

In the original study, Skagerlund Forsblad, Tinghög, and Västfjäll (2021) investigated the relationship between objective numeracy (measured by the BNT; Cokely et al., 2012) and application of decision rules, resistance to sunk cost, and consistency in risk perception (measured by the Adult Decision Making Competence; A-DMC; Bruine de Bruin et al., 2007). The main aim of the study was to investigate which cognitive abilities predicted general decision-making competence. Over three sessions utilizing both computer-based and paper-pencil methods, 182 university students from Sweden underwent a comprehensive battery of tests encompassing numeracy, decision-making competence, executive functions, and cognitive abilities. The study revealed a positive association between higher objective numeracy and superior application of decision rules, increased resistance to sunk cost, and greater consistency in risk perception. Our replication seeks to validate and extend these observed effects.

S01[1]: People with higher numeracy apply decision rules more correctly. In the "applying decision rules" subscale of A-DMC, participants were asked to complete ten tasks requiring applying specified decision rules to indicate for each hypothetical consumer which of the DVD players they would buy (e.g., "*LaToya only wants a DVD player that got a "Very High" rating on Reliability of Brand. Which one of the presented DVD players would LaToya prefer? Select one of the five DVD players presented*").

[1] S01-S11 are the numbers of replicated effects as per our preregistration (https://osf.io/tz2yn).

The outcome variable indicating a superior application of decision rules was the percentage of correct responses. The results indicated that higher objective numeracy was related to a higher percentage of correct answers, $r = .41$; $p < .001$.

S02: People with higher numeracy are more resistant to sunk cost. In the "resistance to sunk cost" subscale of A-DMC, participants were asked to answer ten items on a six-point scale, judging the likelihood that they would ignore prior investments when making decisions (e.g., "*After a large meal at a restaurant, you order a big dessert with chocolate and ice cream. After a few bites, you find you are full, and you would rather not eat any more of it. Would you be more likely to eat more or to stop eating it?*"). The outcome variable indicating higher resistance to sunk cost was the average rating across items. The results indicated that higher objective numeracy was related to higher resistance to sunk cost, $r = .18$; $p < .05$.

S04: People with higher numeracy are more consistent in their risk perception. In the "consistency in risk perception" subscale of A-DMC, participants were asked to rate the percentage probabilities of ten sets of opposing events (e.g., "*What is the probability that you will get into a car accident while driving during the next year?*" vs. "*What is the probability that your driving will be accident-free during the next year?*"). The outcome variable indicating higher consistency in risk perception was the percentage of consistent risk judgments. The results indicated that higher statistical numeracy was related to higher consistency in risk perception, $r = .23$; $p < .05$.

***Higher objective numeracy predicts superior decision making in monetary lotteries* (Sobkow, Olszewska, et al., 2020)**

S05: People with higher numeracy make more choices maximizing expected value. In the original study, Sobkow, Olszewska, and Traczyk (2020) investigated the relationship between objective numeracy (measured by the BNT) and superior decision making under risk (using a set of 10 monetary lotteries). In an online study, Polish-speaking participants from the general population ($N = 581$) were asked to select an alternative they preferred in ten monetary lotteries (e.g., "*Win PLN50 for sure vs. 50% chance to win PLN400 and 50% to win nothing*"). The outcome variable indicating superior decision making was the number of choices that maximized expected value (EV). The results indicated that higher objective numeracy was related to more choices consistent with the EV maximization principle, $r = .295$, $p < .001$.

Numeracy and medical risk comprehension (Rolison et al., 2020)

S06: People with higher numeracy demonstrate better medical risk comprehension. In the original study, Rolison, Morsanyi, and Peters (2020) explored how objective numeracy, subjective numeracy, and math anxiety are associated with risk comprehension. Objective numeracy was measured by a scale consisting of 14 items taken from Schwartz et al. (1997), the Expanded Numeracy Scale (Lipkus et al., 2001), and the Cognitive Reflection Test (Primi et al., 2016). Medical risk comprehension was assessed through 12 items covering the understanding of absolute risk, relative risk, the lifetime risk of cancer based on genetic testing, ratios, misconceptions about the likelihood of random events in a hospital setting, proportions indicating the percentage of individuals at increased risk of developing a serious health condition, and comprehension of comparative information related to multiple hospital performance indicators. The outcome variable for medical risk comprehension was quantified as the number of correct answers. This online study involved 1,011 participants, with 705 from the USA or Canada, 244 from the UK or Ireland, and 59 from other countries. The results indicated that people with higher objective numeracy demonstrated better medical risk comprehension, $r = .70$; $p < .001$.

Numeracy, financial knowledge, and financial well-being (Peters, Tompkins, et al., 2019; Study 1)

In the initial investigation conducted by Peters, Tompkins et al. (2019), the primary objective was to explore the interplay between objective numeracy, numeric self-efficacy (linked to subjective numeracy), and their joint effect on financial outcomes. This large-scale online study involved participants from the Understanding America Study (*N* = 4,572), who completed various assessments, including measures of numeracy and financial outcomes. The specific effects we chose for our replication project were the positive associations between objective numeracy (Weller et al., 2013), financial knowledge, and financial well-being.

S03: People with higher numeracy have better financial knowledge. As a measure of financial knowledge, participants responded to 20 true-false (e.g., "*Bonds are normally riskier than stocks*") and multiple-choice (e.g., "*Assume a friend inherits $10,000 today and his sibling inherits $10,000 but 3 years from now. Who is richer today because of the inheritance?*") questions. The outcome variable was the sum of correct answers converted to an estimated item response theory score. Higher scores indicated greater financial knowledge. The results indicated that higher objective numeracy was related to greater financial knowledge, $r = .57$; $p < .01$.

S07: People with higher numeracy declare higher financial well-being. As a measure of financial well-being, participants answered ten questions assessing their perceived financial well-being on a 5-point scale (e.g., "*I can enjoy life because of the way I'm managing my money*"). The outcome variable was the sum of responses on a scale. Higher scores indicated superior financial well-being. The results indicated that higher objective numeracy was related to greater financial well-being, $r = .25$; $p < .01$.

Numeracy, the loss-bet paradox, preference for suboptimal options, and vulnerability to framing (Peters et al., 2006)

Peters et al. (2006) conducted four studies that delved into the moderating role of objective numeracy (measured using the scale developed by Lipkus et al., 2001; in all studies, the authors used a median split of the numeracy variable) in decision making, particularly in attribute framing, risk representation, preference for suboptimal options, and the attractiveness of simple gambles. For our investigation, we focused on Studies 1, 3, and 4, which are explained below.

S08: Peters et al. (2006) Study 1 - People with lower numeracy show a stronger framing effect. In Study 1, Peters et al. (2006) investigated whether the level of objective numeracy moderates the susceptibility to the attribute framing effect—the cognitive bias where people's decisions are influenced by how the information is presented, particularly whether it is framed positively or negatively. The study was conducted in an experimental design as a part of a series of paper-and-pencil experiments on a sample consisting of volunteers who responded to an announcement in a campus newsletter in the United States (*N* = 100). The framing effects were tested by presenting participants with the exam scores of five psychology students and asking subjects to rate each student's quality of work on a 7-point scale ranging from -3 (very poor) to +3 (very good). The framing of the exam scores was manipulated between the subjects. Depending on the group to which participants were assigned, the results were presented in the form of either the percentage of correct answers or the percentage of incorrect answers. For example, depending on condition, a student was described as receiving either 74% correct or 26% incorrect on their exam. The results indicated that less numerate participants showed a stronger framing effect than highly numerate participants, $F(1, 96) = 5.6$; $p < .05$, $\eta^2 = .11$.

S11: Peters et al. (2006) Study 3 - People with lower numeracy have a stronger preference for suboptimal options than people with higher numeracy (jellybeans task). In Study 3, Peters et al. (2006) investigated whether both the representation of numerical information (i.e., frequency vs. percentage) and numeracy had an effect on decisions. The study was conducted among volunteers from a psychology course (*N* = 46) who were presented with two drawings of bowls filled with colored and white jellybeans. Bowl A was larger and contained 100 jellybeans, 9 of which were colored. Bowl B contained ten jellybeans, 1 of which was colored. Both bowls were labeled with the percentage of colored jellybeans (Bowl A: 9% colored jellybeans vs. Bowl B: 10% colored jellybeans). The participants were asked which bowl they would choose on a 13-point scale ranging from 6 (strong preference for A), zero in the middle, to 6 (strong preference for B)[2]. The results indicated that people with lower numeracy preferred the suboptimal option (i.e., Bowl A with a larger number of colored jellybeans but with a lower objective probability to draw a colored jellybean), $t(44) = -2.5$, $p < .05$; $d = .75$.

S09: Study 4 - People with higher numeracy find the loss bet more attractive than the no-loss bet. In their Study 4, Peters et al. (2006) investigated whether objective numeracy moderated the effect of a loss bet vs. a no-loss bet on the evaluation of the attractiveness of the bet. A paper-and-pencil experiment was conducted using a between-subjects design among volunteers from a psychology department's subject pool on a college campus in the United States ($N = 171$). Participants were instructed to use a 21-point scale ranging from 0 (not at all an attractive bet) to 20 (extremely attractive bet) to evaluate the attractiveness of a simple monetary gamble, either in a no-loss condition (i.e., 7 chances out of 36 to win $9.00) or in a loss condition (i.e., 7 chances out of 36 to win $9.00 and 29 chances out of 36 to lose ¢5).

The results indicated that people with higher numeracy rated the no-loss bet as less attractive than the loss bet, $t(89) = 3.1$, $p < .01$, $d = 0.66$. No differences were observed between the ratings of loss and no-loss bets in individuals with lower numeracy. Additionally, the regression analysis revealed that the interaction between numeracy and the bet condition (with and without a small loss) was statistically significant, suggesting the presence of the loss-bet paradox, $t(1) = 2.2$; $p < .05$, $d = 0.34$.[3]

---

[2] In the original study, after providing their preferences, participants were asked about their affective precision ("*How clear a feeling do you have about the goodness or badness of Bowl A's 9% chance of winning?*") on a scale from 0 (completely unclear) to 6 (completely clear). In the last step, they were asked about their affect ("*How good or bad does Bowl A's 9% chance make you feel?*") on a 7-point scale ranging from -3 (very bad) to 3 (bad). However, in the current study we decided not to replicate this effect.

[3] According to the description of the analysis in the original article, degrees of freedom for the interaction term should be df = 167 (the number of observations, $N = 171$, minus four estimated coefficients: intercept, two predictors and their interaction). Therefore, we decided to perform power analysis based on group means and ANOVA results.

Visual aids, numeracy, and risk comprehension(Garcia-Retamero, Cokely, et al., 2015)

S10: Visual aids improve the interpretation of the predictive value of the diagnostic test among people with lower numeracy. In their original study, Garcia-Retamero, Cokely, and Hoffrage (2015) investigated whether visual aids displaying numerical information could improve the interpretation of the predictive value of the diagnostic test among patients with lower numeracy. In a paper-pencil experiment, 108 Spanish adult patients (M_{age} = 52 years) completed a two-part questionnaire, where the first part included three tasks requiring them to make inferences about breast cancer from a positive mammogram, colon cancer from a positive hemoccult test, and insulin-dependent diabetes from a genetic test. Half of the patients were given information about the sensitivity, false-positive rate, and prevalence of the diseases in numerical format only (natural frequencies), without any visual aids. The other half received the same numerical information along with a grid visually representing the information. The objective accuracy score was calculated as the percentage of correct inferences. Additionally, participants estimated their accuracy in the inference task and completed the numeracy scale (12 items taken from Lipkus et al., 2001 and Schwartz et al., 1997). The participants were split into two groups (low vs. high numeracy) according to the median numeracy score. The results indicated that visual aids improved the interpretation of the predictive value of the diagnostic test (they provided more correct responses) in people with lower numeracy, $F(1, 104) = 3.82$, $p = .05$, $\eta^2 = 0.04$.

Methods

The study was preregistered on the Open Science Framework (OSF; https://osf.io/z2agf) platform. All materials, datasets, and scripts used can be obtained under this link: (https://osf.io/927bx/?view_only=f143e1b75ca64f249968984f9f6f4725. In the preregistration form, we described our aims, replicated effects, known differences between the replication and original studies, sample size justification, and a plan of the analyses. Additionally, we have added to our OSF project exploratory analyses (e.g., response times) and tables comparing sample characteristics, procedural details, setting, and other relevant aspects of the original studies with those in the present replication. Based on the criteria proposed by LeBel et al. (2017), we have classified our study as a close or very close replication.

Selection criteria

The study was conducted in cooperation with undergraduate students as part of a decision-making course at the SWPS University. Eleven effects were selected based on the following criteria: 1) study reporting a simple correlation between objective numeracy and a measure related to decision making or objective numeracy used as a moderator, 2) the brevity of the procedure, 3) feasibility of conducting an online study in adults from the general population. We have intentionally selected studies varying in effect sizes and the type of decision-making measures.

Participants

Two hundred and nine Polish-speaking Prolific users aged 18 to 53 (*M* = 23.6; *SD* = 6.3) participated in the study, out of which 106 were males and 94 were females, 52 had higher education, and 148 had secondary education at the time the study was conducted. Participants were reimbursed for their participation with £6.88 for a 55-minute study. The study was approved by the Ethics Committee at the Faculty of Psychology in Wroclaw at SWPS University (06/P/12/2021).

Procedure

The study was conducted in Polish. Participants completed two blocks in a randomized order. The Numeracy block contained the Berlin Numeracy Test (Cokely et al., 2012), the Abbreviated Numeracy Scale (Weller et al., 2012), the numeracy scale developed by Lipkus et al. (2001), three items from the Cognitive Reflection Test (Primi et al., 2016), the Subjective Numeracy Scale (Fagerlin et al., 2007), the Symbolic-number mapping task (Sobkow, Olszewska, et al., 2020). The order of presentation of numeracy scales was counterbalanced.

The judgment and decision making (JDM) block contained three scales from the Adult-Decision Making Competence (Bruine de Bruin et al., 2007): sunk cost, decision rules, and consistency in risk perception; 12 items measuring medical risk comprehension (Rolison et al., 2020); a set of ten monetary lotteries (Sobkow, Olszewska, et al., 2020); ten questions measuring financial well-being (Peters, Tompkins, et al., 2019); 20 questions measuring financial knowledge (Houts & Knoll, 2020, as cited in Peters, Tompkins, et al., 2019); the attribute framing task (Peters et al., 2006); the loss-bet lottery task (Peters et al., 2006); the diagnostic inferences task (Garcia-Retamero, Cokely, et al., 2015); and the suboptimal choice task (jellybeans task; Peters et al., 2006). The order of all JDM tasks was counterbalanced. All study materials were translated into Polish based on those published in the original articles or provided by the authors (if not available online). The exact experimental procedure used in the study can be found at the following link: https://osf.io/927bx/?view_only=f143e1b75ca64f249968984f9f6f4725.

Power analysis

We used G*Power (for correlation coefficients and t-test; Faul et al., 2007) to calculate the required sample size for each correlational effect (one-tailed) based on the original effect size (we used effect sizes reported in original). Conventionally, we assumed $\alpha = .05$ and $1 - \beta = .80$. Since we did not have access to the raw data used to analyze interaction effects in the experimental studies (i.e., S08 – framing, S09 – loss vs. no-loos bets, and S10 – visual aids), we extracted the group means for the 2x2 design and estimated standard deviations (assuming equal sample sizes across conditions) based on the available information[4]. Using this data, we simulated datasets with the *Superpower* package (Lakens &

[4] We used WebPlotDigitizer https://automeris.io/ to extract data regarding means and standard deviations.

Caldwell, 2021) and conducted power analysis to detect the simulated interaction effects, with $\alpha = .05$ and $1 - \beta = .80$. The minimum number of participants to test the smallest effect size included in the replication project was $N = 197$. We collected data from a larger number of participants to account for potential issues with data quality in an online study, including participants who might resign or withdraw. Documentation of our sample size calculations, including scripts and software printouts, is available on the OSF project platform.

Results

Main analyses

Our initial aim was to assess the replicability of the selected effects with the numeracy scales used in the original studies. We assumed in our preregistration that the effect is successfully replicated if it is statistically significant ($p < .05$) and its direction is consistent with the one in the original study. In the case of moderation effects (i.e., S08 – framing, S09 – loss vs. no-loos bets, and S10 – visual aids), we converted F-values for interaction effects to Pearson's *r* coefficients. All effects were transformed into Pearson's r (with 95% CIs) using the *correlation* (Makowski et al., 2020) and *effectsize* (Ben-Shachar et al., 2020) packages in the R statistical computing environment (R Core Team, 2022) to allow for comparisons among effect sizes. For the preregistered main analyses, we used one-tailed statistical tests, while for additional exploratory analyses, we used two-tailed tests. For each effect size, we report or plot the 95% CI for the two-sided tests.

Using the abovementioned criterion, six correlational effects were replicated successfully (Table 1 and Figure 1).

Figure 1

Replicated effects with the numeracy scales used in the original research. Effects sizes (Pearson's r) in the original and replication studies are represented with blue circles and orange triangles, respectively. For moderation effects (i.e., framing, loss vs. no-loos bets, and visual aids), we converted F-values for interaction effects to Pearson's r coefficients. Error bars represent a 95% confidence interval for the original and replicated effect sizes.

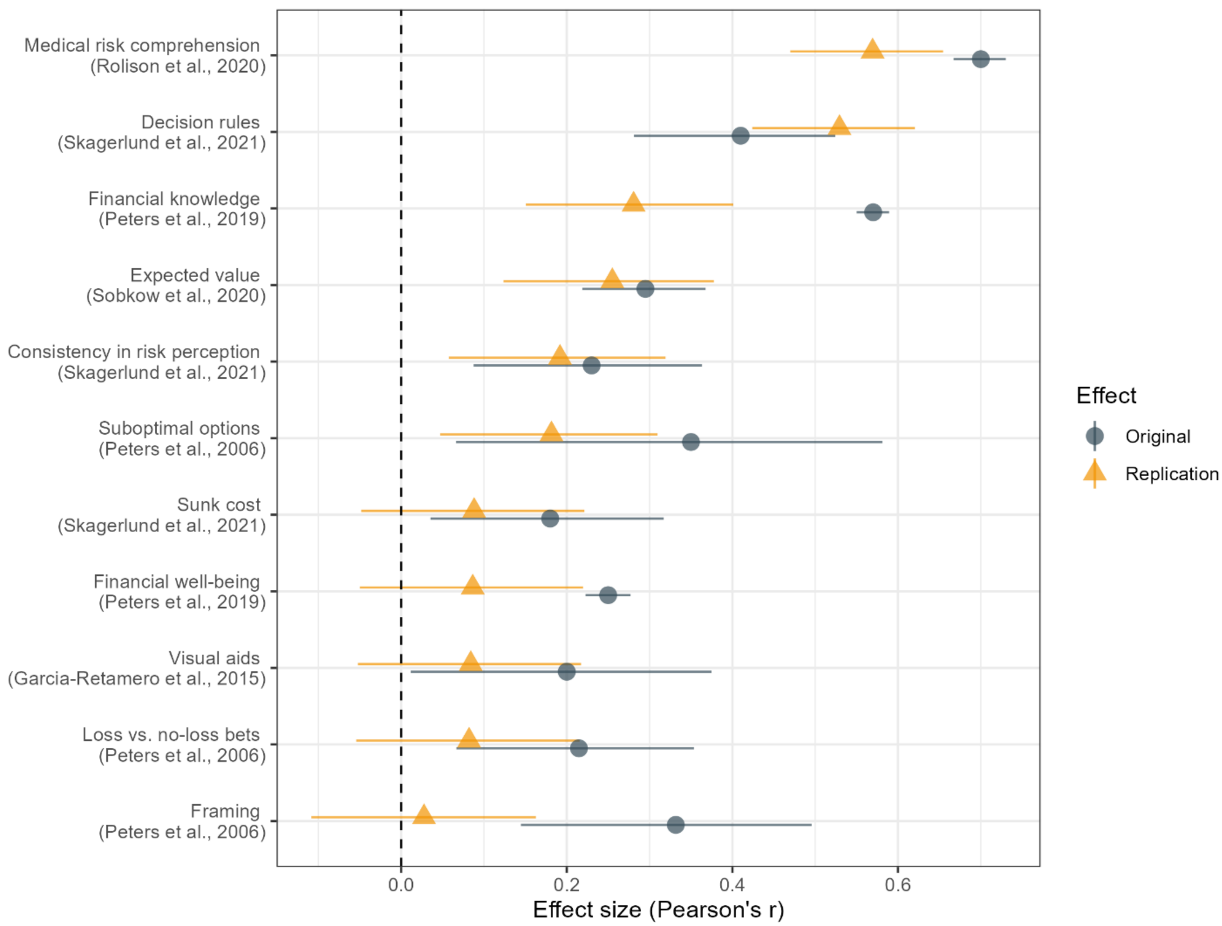


Table 1

Replicated effects with varied objective numeracy scales.

Effect	BNT	Peters & Lipkus	Garcia-Retamero	Rolison	Weller
S01 Decision rules	**.53*** **	.45***	.46***	.54***	.56***
S02 Sunk cost	**.09**	-.02	-.04	.05	.003
S03 Financial knowledge	.14	.22**	.21**	.28***	**.28*** **
S04 Consistency in risk perception	**.19** **	.27***	.27***	.30***	.31***
S05 Expected value	**.26*** **	.27***	.27***	.29***	.34***
S06 Medical risk comprehension	.46***	.52***	.52***	**.57*** **	.55***
S07 Financial well-being	.03	.10	.09	.07	**.09**
S08 Framing	.03	**.03**	.03	.01	.10
S09 Loss vs. No loss bets	.01	**.08**	.06	.02	.02
S10 Visual aids	.07	.10	**.08**	.08	.01
S11 Suboptimal options	.13	**.18** **	.18**	.20**	.18**

Note. Pearson's r correlation coefficients for the relationships among replicated effects and various objective numeracy scales. Numeracy scales used in original studies are in bold. The columns represent different numeracy measures used in the replicated studies: BNT - Berlin Numeracy Test (Cokely et al., 2012); Peters & Lipkus - numeracy scale used by Peters et al. (2006); Garcia-Retamero - numeracy scale used by Garcia-Retamero, Cokely et al. (2015); Rolison - numeracy scale used by Rolison et al. (2020); Weller - Abbreviated Numeracy Scale (Weller et al., 2012).
* p < .05, ** p < .01, *** p < .001 (two-tailed)

In line with the original findings, we found that people with higher objective numeracy applied decision rules more correctly (S01; $r = .53$, $p < .001$), made more EV-consistent choices in lottery gambles (S05; $r = .26$, $p < .001$), better understood medical risks (S06; $r = .57$, $p < .001$), had better financial knowledge (S03; $r = .28$, $p < .001$), were more consistent in risk perception (S04, $r = .192$; $p = .003$), and were more likely to make optimal choices, preferring the bowl with a higher probability of drawing a winning jellybean (S11; $r = .18$, $p = .004$). However, we did not replicate the relationships between objective numeracy and financial well-being (S07; $r = .09$, $p = .107$) and resistance to sunk cost (S02, $r = .09$; $p = .102$).

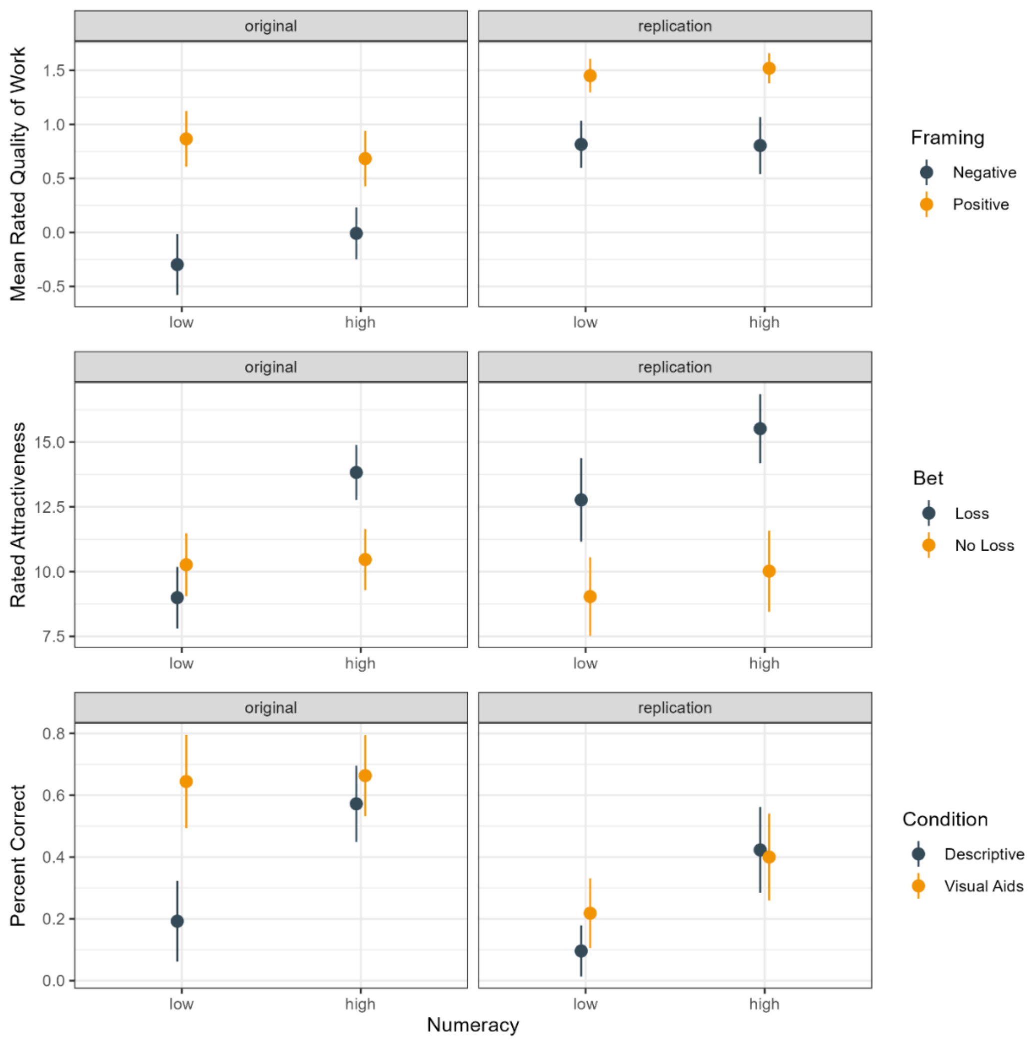
None of the effects in which objective numeracy served as a moderator were replicated (see Figure 2). Despite the significant main effect of attribute framing (S08), $F(1, 205) = 45.185$, $p < .001$, $\eta^2_p = .181$, we did not find a significant interaction effect with objective numeracy, $F(1, 205) = 0.157$, $p = .693$, $\eta^2_p = .001$ (as in the original studies, participants were categorized into low and high-numeracy groups based on a median split). Similarly, although we observed the main effect of the loss/no-loss bet, $F(1, 205) = 37.841$, $p < .001$, $\eta^2_p = .156$, and objective numeracy, $F(1, 205) = 6.182$, $p = .014$, $\eta^2_p = .029$, on the attractiveness of a gamble, we did not find a significant interaction effect, $F(1, 205) = 1.389$, $p = .240$, $\eta^2_p = .007$ (S09). Nevertheless, an exploratory independent t-test revealed that people with higher numeracy rated the loss bet as more attractive ($M = 15.52$, $SD = 4.89$) than the no-loss bet ($M = 10.02$, $SD = 5.56$), $t(103) = 5.387$, $p < .001$, $d = 1.052$. Additionally, contrary to original findings, there was no significant interaction between information format (i.e., numerical only vs. numerical and visual) and numeracy, $F(1, 205) = 1.459$; $p = .228$, $\eta^2_p = .007$ (S10). In line with our preregistered brief description of the effect, we also decided to explore the difference between experimental conditions in individuals with lower numeracy. The results revealed a significant difference in the mean number of correct inferences. Participants with lower numeracy who were presented with a visual aid ($M = 0.22$, $SD = 0.42$) made significantly more correct inferences than those who received only numerical information ($M =$

0.10, SD = 0.30), *t*(105) = -1.734, *p* = .043, *d* = -0.335. Also, we found that individuals with higher numeracy generally exhibited a better understanding of the results of this diagnostic test, $F(1, 205) = 17.937$, $p < .001$, $\eta^2_p = .080$.

In summary, employing the preregistered criteria (i.e., the significance and direction of the effect), we successfully replicated six correlational effects. Two correlational effects and three moderation effects were not replicated.

Figure 2

Differences between conditions for experimental effects: attribute framing (S08), loss/no-loss bet (S09), and visual aids (S10) in original and replication studies. Error bars represent 95% CI. Participants were divided into low/high numeracy groups based on a median split. Means and standard deviations in the original studies were simulated based on the available data and analyses provided in the articles.



Additional analyses

As part of our exploratory analyses, we aimed to assess the comparative predictive efficacy of various objective numeracy scales utilized in replicated studies, along with investigating the roles of two complementary types of numeracy—subjective and approximate. Last but not least, we compared the original and replication effects using different methods to evaluate replicability success.

Relationships among numeracy scales

Overall, positive correlations (Table 2) were observed among all objective numeracy scales, with a moderate correlation (*r* ranging from .42 to .53) observed with the BNT and a very strong correlation among other scales (*r* ranging from .80 to .99). This pattern was expected, given the similarity or overlap in items among some of the scales used in the replicated studies (Garcia-Retamero, Cokely, et al., 2015; Peters et al., 2006; Rolison et al., 2020; Weller et al., 2012). In contrast, the BNT is a relatively independent scale designed for a highly educated sample, and the strength of correlation between BNT and other objective numeracy scales was similar to the original study validating this test (Cokely et al., 2012).

The correlation coefficients between objective numeracy scales and subjective numeracy are moderate, ranging from .39 to .49, while the correlation between objective numeracy scales and approximate numeracy ranges from .30 to .46. This suggests that these constructs partially overlap but may also tap into complementary mechanisms related to motivation/numeric self-efficacy (subjective numeracy) or intuitive number representation (approximate numeracy). This pattern of findings aligns with previous research results (Peters & Bjalkebring, 2015; Sobkow et al., 2019; Sobkow, Olszewska, et al., 2020).

Impact of numeracy measures on replication success

To examine the impact of numeracy measures on replication success, we conducted Pearson's correlations for all numeracy measures employed in our study (refer to Table 1 for results on objective numeracy scales and Table 3 for subjective and approximate numeracy).

Generally, the results of replication were consistent across all objective numeracy scales. However, a distinction arose in the effect of financial knowledge (S03). Specifically, the strongest effect was observed when numeracy scales used by Weller et al. (2013) and Rolison et al. (2020) were employed, and for the BNT there were no significant effects (Table 1). This may be due to the BNT's higher difficulty level, which can limit the strength of correlations, potentially reducing its sensitivity to detect effects in some contexts.

Regarding the two complementary types of numeracy, namely subjective and approximate, we identified subtle yet significant relationships with the consistency in risk perception (r = .15 and r = .18, respectively; see Table 3). Furthermore, approximate and subjective numeracy also correlated with the resistance to the sunk cost effect (r = .12 and r = .20, respectively). Finally, subjective numeracy emerged as a significant predictor of financial well-being (r = .20) despite no effects being found for objective numeracy scales. Neither SNS nor SMap were related to performance in the framing (S08), the loss bet task (S09), and inference (S10) tasks.

We argue that the abovementioned effect may be more driven by the motivational or intuitive aspects of numeracy than by the objective ability to perform calculations. To further explore the role of different numeracy measures on the replicated effects, we fitted linear regression models (Table 4). First, only an objective numeracy measure used in the original study was introduced to the model. In the next step, we added measures of subjective (SNS) and approximate (SMap) numeracy. We found that subjective numeracy contributed significantly to several effects (S01, S02, S03, S05, S06, S07, S09, S10, S11), explaining the variance of outcome variables beyond measures of objective numeracy.

Approximate numeracy significantly increased the model fit for the decision rules (S01) and medical risk comprehension (S06) effects.

Table 2

Descriptive statistics and Pearson's correlations among various numeracy measures.

	Numeracy measure	Scale	*M*	*Median*	*SD*	Min	Max	1	2	3	4	5	6	7
1	objective numeracy	BNT	1.79	2.0	1.24	0.00	4.00	-						
2	objective numeracy	Peters & Lipkus	9.82	11.0	1.73	0.00	11.00	.415	-					
3	objective numeracy	Garcia-Retemero	10.74	11.0	1.84	0.00	12.00	.420	.990	-				
4	objective numeracy	Rolison	11.68	12.0	2.37	1.00	14.00	.526	.929	.924	-			
5	objective numeracy	Weller	6.05	6.0	1.84	0.00	8.00	.533	.805	.801	.850	-		
6	subjective numeracy	SNS	33.19	34.0	7.80	13.00	48.00	.433	.410	.392	.483	.491	-	
7	approximate numeracy	SMap	-3.02	-3.0	0.42	-5.00	-1.33	.368	.420	.398	.464	.365	.299	-

Note. Note. Pearson's r correlation coefficients. BNT - Berlin Numeracy Test (Cokely et al., 2012); Peters & Lipkus - numeracy scale used in Peters et al. (2006); Garcia-Retamero - numeracy scale used in Garcia-Retamero, Cokely et al (2015); Rolison - numeracy scale used in Rolison et al., (2020); Weller - Abbreviated Numeracy Scale (Weller et al., 2012); SNS - Subjective Numeracy Scale (Fagerlin et al., 2007); SMap - Symbolic-number mapping (Sobkow, Olszewska, et al., 2020); all *p*s <.001

Table 3

Replicated effects when subjective and approximate numeracy measures were used.

Effect	SNS	SMap
S01 Decision rules	.40*** [.27, .50]	.38*** [.26, .49]
S02 Sunk cost	.20** [.07, .33]	.12 [-.02, .25]
S03 Financial knowledge	.31*** [.19, .43]	.12 [-.01, .25]
S04 Consistency in risk perception	.15* [0.02, .28]	.18** [.04, .31]
S05 Expected value	.23*** [.10, .35]	.16* [.02, .29]
S06 Medical risk comprehension	.39*** [.27, .50]	.38*** [.25, .49]
S07 Financial well-being	.20** [.07, .33]	-.06 [-.19, .08]
S08 Framing	.13 [-.01, .26]	.05 [-.09, .18]
S09 Loss vs. No loss bets	.10 [-.04, .23]	.07 [-.07, .20]
S10 Visual aids	.01 [-.13, .14]	.09 [-.05, .22]
S11 Suboptimal options	.23*** [0.09, .35]	.15* [0.01, .28]

Note. Pearson's r correlation coefficients [95% CI]. * $p < .05$, ** $p < .01$, *** $p < .001$ (two-tailed)

Table 4

Linear regression models predicting replicated effects with the original objective numeracy scales and subjective/approximate numeracy measures.

Effect	Coefficient	*b*	*SE*	*t*	*p*		R^2	Coefficient	*b*	*SE*	*t*	*p*		R^2
S01 Decision rules	Intercept	5.53	0.22	24.95	<.001	***		Intercept	7.26	1.22	5.96	<.001	***	
	Objective numeracy	0.91	0.10	8.98	<.001	***		Objective numeracy	0.66	0.11	5.91	<.001	***	
								Subjective numeracy	0.05	0.02	2.71	.007	**	
								Approximate numeracy	0.95	0.31	3.03	.003	**	
							.28							.34
S02 Sunk cost	Intercept	39.94	0.79	50.42	<.001	***		Intercept	38.95	4.48	8.70	<.001	***	
	Objective numeracy	0.46	0.36	1.27	.200			Objective numeracy	-0.11	0.41	-0.27	.790		
								Subjective numeracy	0.16	0.06	2.50	.013	*	
								Approximate numeracy	1.09	1.15	0.95	.342		
							.01							.05
S03 Financial knowledge	Intercept	7.16	0.64	11.18	<.001	***		Intercept	5.22	1.94	2.69	.008	**	
	Objective numeracy	0.43	0.10	4.21	<.001	***		Objective numeracy	0.26	0.12	2.19	.030	*	
								Subjective numeracy	0.08	0.03	3.04	.003	**	
								Approximate numeracy	-0.06	0.47	-0.14	.891		
							.08							.12
S04 Consistency in risk perception	Intercept	15.90	0.28	56.78	<.001	***		Intercept	17.34	1.60	10.84	<.001	***	
	Objective numeracy	0.36	0.13	2.81	.005	**		Objective numeracy	0.23	0.15	1.54	.125		
								Subjective numeracy	0.02	0.02	0.89	.372		
								Approximate numeracy	0.62	0.41	1.51	.132		
							.04							.05
S05 Expected value	Intercept	7.37	0.18	40.60	<.001	***		Intercept	7.20	1.03	6.95	<.001	***	
	Objective numeracy	0.32	0.08	3.80	<.001	***		Objective numeracy	0.22	0.10	2.29	.023	*	
								Subjective numeracy	0.03	0.01	1.83	.069		
								Approximate numeracy	0.18	0.27	0.69	.490		
							.07							.08
S06 Medical risk comprehension	Intercept	4.39	0.47	9.27	<.001	***		Intercept	5.97	1.13	5.28	<.001	***	
	Objective numeracy	0.40	0.04	9.96	<.001	***		Objective numeracy	0.31	0.05	6.40	<.001	***	

								Subjective numeracy	0.03	0.01	2.11	.036	*	
								Approximate numeracy	0.50	0.25	2.01	.046	*	
							.32							.35
S07 Financial well-being	Intercept	20.85	1.68	12.44	<.001	***		Intercept	8.48	5.06	1.68	.095		
	Objective numeracy	0.33	0.26	1.25	.213			Objective numeracy	0.08	0.31	0.27	.789		
								Subjective numeracy	0.21	0.07	2.97	.003	**	
								Approximate numeracy	-2.27	1.23	-1.85	.066		
							.01							.06
S08 Framing	Intercept	1.15	0.05	22.83	<.001	***		Intercept	0.68	0.50	1.36	.175	***	
	Objective numeracy	0.03	0.10	0.28	.777			Objective numeracy	-0.05	0.11	-0.43	.665		
	Condition	0.68	0.10	6.72	<.001	***		Condition	0.67	0.10	6.67	<.001	***	
	Objective numeracy:Condition	0.08	0.20	0.40	.693			Objective numeracy:Condition	0.06	0.20	0.31	.757		
								Subjective numeracy	0.01	0.01	1.89	.060		
								Approximate numeracy	-0.01	0.13	-0.06	.951		
							.18							.20
S09 Loss vs. No loss bets	Intercept	11.84	0.38	31.56	<.001	***		Intercept	10.29	3.58	2.88	.004	**	
	Objective numeracy	1.87	0.75	2.49	0014	*		Objective numeracy	0.36	0.79	0.45	.653		
	Condition	4.61	0.75	6.15	<.001	***		Condition	4.64	0.72	6.46	<.001	***	
	Objective numeracy:Condition	1.77	1.50	1.18	.240			Objective numeracy:Condition	1.76	1.45	1.22	.224		
								Subjective numeracy	0.19	0.05	3.71	<.001	***	
								Approximate numeracy	1.57	0.92	1.70	.090		
							.18							.26
S10 Visual aids	Intercept	0.28	0.03	9.47	<.001	***		Intercept	0.19	0.29	0.66	.511		
	Objective numeracy	0.25	0.06	4.24	<.001	***		Objective numeracy	0.17	0.07	2.65	.009	**	
	Condition	0.05	0.06	0.82	.411			Condition	0.06	0.06	1.00	.320		
	Objective numeracy:Condition	-0.15	0.12	-1.21	.228			Objective numeracy:Condition	-0.13	0.12	-1.09	.278		
								Subjective numeracy	0.01	0.00	2.52	.013	*	
								Approximate numeracy	0.09	0.08	1.14	.257		
							.09							.13
S11 Suboptimal options	Intercept	5.77	1.56	3.70	.000	***		Intercept	6.69	3.36	1.99	.048	*	
	Objective numeracy	0.41	0.16	2.66	.009	**		Objective numeracy	0.20	0.18	1.09	.278		

	Subjective numeracy	0.09	0.04	2.30	.022	*	
	Approximate numeracy	0.56	0.71	0.79	.432		
.03							.06

Note. * $p < .05$, ** $p < .01$, *** $p < .001$

Comparison of different methods used to evaluate the replication success

There is no universally accepted robust method for evaluating the success or failure of replication attempts (Errington et al., 2021; Mathur & VanderWeele, 2019). Therefore, we employed additional four dichotomous methods to assess successful or unsuccessful replication by comparing the effect reported in the original study with the effect observed in the replication attempt (Table 5).

Table 5

Evaluation of the replication attempts using additional criteria (based on Errington et al., 2021)

Effect	Same direction and significance	Same direction	Original ES in replication 95%CI	Replication ES in original 95% CI	Replication ES ≥ original ES
S01 Decision rules	✓	✓			✓
S02 Sunk cost		✓	✓	✓	
S03 Financial knowledge	✓	✓			
S04 Consistency in risk perception	✓	✓	✓	✓	
S05 Expected value	✓	✓	✓	✓	
S06 Medical risk comprehension	✓	✓			
S07 Financial well-being		✓			
S08 Framing					
S09 Loss vs. no-loss bets		✓	✓	✓	
S10 Visual aids		✓	✓	✓	
S11 Suboptimal options	✓	✓		✓	
Summary	6 effects (55%)	10 effects (91%)	5 effects (46%)	6 effects (55%)	1 effect (9%)

Note: ES – Effect Size

According to our first replication criterion (preregistered for the purpose of this project), a study was considered successfully replicated when both the original and replication effects were in the same direction and statistically significant ($p < 0.05$). Applying this measure, we found that 6 out of 11 effects (55%) were successfully replicated. The second, more basic method involved evaluating replicability through analyzing only the direction of the effect. According to this method, 10 effects (91%) were successfully replicated (i.e., for the correlational effects the direction of the relationship was the same as in the original studies, while for the moderation effects, the profile of means in simple effects was consistent with the pattern of interaction effects of the original studies). The next criterion assumed that replication was deemed successful if the original effect size fell within the 95% CI of the replication effect size. Using this approach, 5 out of 11 effects (46%) were successfully replicated. The fourth criterion adopted was a complementary approach, wherein replication was considered successful if the replication effect size was contained within the 95% CI of the original effect size. Based on this method, 6 out of 11 effects (55%) were deemed successfully replicated. Lastly, we also applied the most stringent criterion, requiring that the replication effect size be equal to or greater than the original effect size for a replication to be considered successful. Under this criterion, only 1 out of 11 studies (9%) achieved successful replication.

Discussion

In the present study, we sought to replicate 11 effects concerning the role of numeracy in judgment and decision making. While our approach may be considered limited or controversial (Brandt et al., 2014; Errington et al., 2021), for the sake of clarity and simplicity, we defined a successful replication in our preregistration as achieving a statistically significant effect in the same direction as reported in the original study. Using this criterion, we successfully replicated six effects, and found no support for five effects. Notably, correlational effects were more likely to replicate than those where numeracy was a moderator.

Nevertheless, we also examined alternative replication criteria (Errington et al., 2021), which in some instances might lead to conclusions that differ from those based on our preregistered definition of success. Furthermore, our analysis revealed that different numeracy scales had a comparable impact on the outcome variables and incorporating measures of subjective and approximate numeracy enhanced the model fit for some effects. This suggests that a comprehensive understanding of effects in the field of judgment and decision making may necessitate analyzing multiple components of numeracy.

The effects of objective numeracy on decision making

Using the objective numeracy scales from the original studies, we replicated six correlational effects out of eight (S01, S03, S04, S05, S06, S11), achieving a 75% replication success rate. The effects we successfully replicated were slightly weaker than the original ones, which is consistent with other replication projects, and the effect sizes were comparable regardless of the objective numeracy scale used. This suggests that objective numeracy is a stable and robust predictor for several outcomes, such as applying decision rules, maximizing expected value, understanding medical risks, making optimal choices, having consistent risk preferences, and possessing better financial knowledge.

According to our preregistered criteria, we did not find support for two correlational effects regarding resistance to sunk cost and financial well-being (S02, S07). However, in both cases, the observed effects were in the same direction as in the original studies. For the sunk cost replication, the effect size fell within the 95% confidence interval (CI) of the original study, and vice versa (the original effect size was within the 95% CI of the replication). Notably, the effect of numeracy on resistance to sunk cost in the original study was relatively small (r = .18 vs. r = .09 in our replication), and the relationship between numeracy and financial well-being in the original study was based on a large and diversified sample of participants. The lack of sufficient power to detect a small effect size of r = .09 could explain our unsuccessful replication.

A more complex pattern of relationships emerged for the three effects where objective numeracy served as a moderator variable. We found only partial evidence for the role of numeracy in the diagnostic inference task (S10) and the ratings of the attractiveness of simple gambles (S09). Specifically, people with lower numeracy benefited from visual aids and numerical information compared to numerical information alone, though a smaller but significant effect was also observed in people with higher numeracy. Additionally, people with higher numeracy rated the attractiveness of a loss bet higher compared to a no-loss bet, but this difference, although smaller, was also present in people with lower numeracy. When we applied alternative replication criteria, we found that these two effects (S09 and S10) were in the same direction as in the original studies, and their effect sizes fell within the 95% confidence intervals (CIs) of the original studies and vice versa (the original effect sizes were within the 95% CI of the replication).

Finally, we found no evidence for the interaction effect between numeracy and framing (S08). While our participants rated students more positively in a positive frame compared to a negative frame, numeracy did not moderate this effect. This finding was not replicated even when with alternative criteria.

The mixed or unsuccessful replication results, where numeracy was a moderator, may be attributed to several factors in our study. These include procedural differences between the original studies and our replication (see the supplementary table in the OSF project for details), variations in sample characteristics, and effect sizes.

In particular, our study included a large number of diverse tasks that were not present in the original studies. While previous research using similar designs (e.g., the Many Labs 2 project; Klein et al., 2018) found little evidence for order effects, we counterbalanced task order to mitigate potential systematic position biases. However, it is possible that repeated exposure to mathematical tasks may have influenced responses in subsequent tasks—such as by encouraging additional math operations—potentially reducing the power of the moderator. A similar order effect for the bet task was observed in Peters, Fennema, & Tiede (2019).

Additionally, our sample—comprised primarily of younger, Polish-speaking Prolific users—tended to have relatively higher numeracy levels (e.g., a median score of 11 on the Lipkus scale) than those in some of the original studies. For example, in Garcia-Retamero, Cokely et al. (2015), participants were hospital patients with an average age of 52 and a median numeracy score of 8. Similarly, in the two moderation effects reported by Peters et al. (2006), our sample scored higher on the Lipkus numeracy scale than the original sample (median = 9). This narrower range of numeracy in our sample may have limited the appearance of moderation effects.

Nevertheless, similarly to our replication attempt, several studies (e.g., Lohre et al., 2019; Mandel & Kapler, 2018; Peters et al., 2011) found no significant moderation effect of numeracy on modified framing tasks, including both objective and subjective numeracy. Finally, recent replication by Zhu and Feldman (2023), conducted with a large sample of MTurk users (*N* = 860), found evidence for significant framing and numeracy effects, but the interaction effect was negligible (η^2_p = .01).

Similarly, regarding loss/no-loss bets, previous research has produced inconsistent results. Notably, when only highly numerate participants (e.g., certified public accountants) rated the attractiveness of loss versus no-loss bets (Peters, Fennema, et al., 2019), no interaction effect was observed. However, the main effects of the condition and numeracy aligned with our findings. Likewise, Reyna and Brust-Renck (2020) observed no interaction, even with two large samples (*N* > 900 each) were used (main effects of both the condition and numeracy were significant). Lastly, in a large replication study by Zhu and Feldman, the interaction effect was significant, though the effect size remained small (η^2_p = .02).

One could argue that despite conducting power calculations prior to our study, it may still have been underpowered, particularly for detecting moderation effects. We based our calculations on effect sizes reported in the original studies; however, as shown in replication studies by Zhu and Feldman (2023), the true effect sizes may be considerably smaller. An alternative approach to estimating sample size for replication studies involves multiplying the original sample size by 2.5 (Simonsohn, 2015). Nevertheless, implementing this method systematically in our case presented significant challenging. Some original studies had exceptionally large sample sizes, and using this multiplier across all effects would have been prohibitively expensive and could significantly increase the risk of false-positive results.

In summary, while the majority of correlational effects were successfully replicated, the phenomena where numeracy acted as a moderator proved to be more nuanced, possibly requiring larger or more diverse samples and/or greater control over procedural details.

The complementary effects of subjective and approximate numeracy on decision making

One of the secondary aims of our replication project was to evaluate the complementary roles of subjective and approximate numeracy scales in various aspects of judgment and decision making. When subjective and approximate numeracy were taken into account, the pattern of replication effects was generally similar to that observed with objective numeracy. For instance, replacing the objective numeracy scales with the subjective numeracy scale resulted in the successful replication of nine out of eleven decision outcomes, while using the approximate numeracy scale in place of objective numeracy was associated with the replication of six effects.

Nevertheless, previous studies (Patalano et al., 2020; Peters & Bjalkebring, 2015; Schley & Peters, 2014; Sobkow, Olszewska, et al., 2020) suggest that objective, subjective, and approximate numeracy may lead to distinct predictions, as each taps into different cognitive or motivational mechanisms. Below, we explore these mechanisms in the context of the effects, for which we observed unique contributions from different numeracy facets, either by leading to divergent conclusions about replication success or by explaining additional variance in the outcomes.

Interestingly, despite not replicating two effects (namely, resistance to sunk cost and financial well-being) with original or alternative objective numeracy scales, we found that higher scores in the subjective numeracy scale were related to greater resistance to sunk cost and superior financial well-being.

It has been argued that subjective numeracy can be viewed as numeric self-efficacy or numeric confidence (Peters & Shoots-Reinhard, 2022). Individuals with higher scores on this numeracy scale may be more strongly motivated to work with numbers and more adept at handling the challenges associated with using quantitative information in everyday situations. Numeric self-efficacy encourages persistence in understanding and managing financial matters. This persistence is crucial for maintaining beneficial financial habits, such as budgeting, saving, and investing, which together contribute to long-term financial well-being. Furthermore, we argue that higher numeric self-efficacy helps individuals resist common cognitive biases like the sunk cost fallacy, where past investments unduly influence future financial decisions. By avoiding such biases, individuals make more rational and beneficial financial choices.

Interestingly, our study revealed that a better intuitive understanding of numbers, measured by the SMap task, was correlated with better performance in some decision-making tasks, such as consistency in risk perception, maximizing expected value (EV), and making fewer suboptimal decisions. However, when we controlled our regression models for objective and subjective numeracy measures, approximate numeracy explained an independent portion of the outcome variable's variance in two effects: decision rules and medical risk comprehension. These findings suggest that while basic intuitive numerical skills (approximate numeracy) may not significantly contribute to explaining all judgment and decision making variables measured in this study, they appear to be particularly important for certain types of tasks. However, the limited number of significant findings for approximate numeracy may also be attributed to random variation or measurement limitations.

This pattern suggests that for tasks like decision rules and medical risk comprehension, beyond possessing adequate levels of objective numeracy and numerical self-efficacy, individuals who make superior choices also need to incorporate an intuitive understanding of numbers and magnitudes into their decision-making process. Even though these explanations are promising, they should be rigorously tested in future studies to better understand the contribution of various numeracy components to improved decision making.

Limitations and Future Directions

Several limitations should be acknowledged. First, our sample was primarily composed of Polish-speaking Prolific users, which may limit the generalizability of our findings to other populations. Nevertheless, choosing this group of participants was aimed at testing and replicating the effects across different populations compared to the original studies. Second, while we employed rigorous preregistered criteria for replication, variations in study design and measurement tools could have influenced the outcomes. For example, cultural or language differences may impact the success of replication, particularly for questions regarding financial well-being or resistance to sunk cost. Future research should aim to replicate these findings across diverse cultural and demographic groups and explore the underlying mechanisms driving the observed effects.

Moreover, the mixed results regarding the moderation effects of numeracy suggest that the interaction between numeracy and decision-making processes is complex and warrants further investigation. Our sample consisted of highly numerate participants (i.e., those scoring relatively high on numeracy scales). While some studies focused on highly educated samples (Ghazal et al., 2014), others aimed to estimate the effects in the general and more diversified population (Peters, Tompkins, et al., 2019). Although we classified participants into high- and low-numeracy groups based on a median split as in the original studies, our division may have primarily distinguished between individuals with moderate and high numeracy. Additionally, we did not control for other cognitive abilities, such as fluid or crystallized intelligence. Although previous research has demonstrated that objective numeracy predicts decision outcomes beyond intelligence, it is crucial to control for other measures of individual differences to provide clearer and more robust evidence of numeracy's role.

Conclusion

In conclusion, our replication study provides valuable insights into the interplay between numeracy and decision making. We successfully replicated several critical effects, demonstrating that objective numeracy predicts different decision-making variables. However, the partial and non-replications, particularly where numeracy served as a moderator variable, underscore areas that warrant further investigation. By embracing a comprehensive approach that includes various dimensions of numeracy—subjective numeric competence and intuitive approximate numeracy—future research can deepen our understanding of the psychological processes underpinning people's performance in decision-making tasks.

References

Apter, A. J., Cheng, J., Small, D., Bennett, I. M., Albert, C., Fein, D. G., George, M., & Van Horne, S. (2006). Asthma Numeracy Skill and Health Literacy. *Journal of Asthma*, *43*(9), 705–710. https://doi.org/10.1080/02770900600925585

Apter, A. J., Wang, X., Bogen, D., Bennett, I. M., Jennings, R. M., Garcia, L., Sharpe, T., Frazier, C., & Ten Have, T. (2009). Linking numeracy and asthma-related quality of life. *Patient Education and Counseling*, *75*(3), 386–391. https://doi.org/10.1016/j.pec.2009.01.003

Ben-Shachar, M. S., Lüdecke, D., & Makowski, D. (2020). effectsize: Estimation of effect size indices and standardized parameters. *Journal of Open Source Software*, *5*(56), 2815. https://doi.org/10.21105/joss.02815

Brandt, M. J., IJzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., Grange, J. A., Perugini, M., Spies, J. R., & van ’t Veer, A. (2014). The Replication Recipe: What makes for a convincing replication? *Journal of Experimental Social Psychology*, *50*(1), 217–224. https://doi.org/10.1016/j.jesp.2013.10.005

Bruine de Bruin, W., Parker, A. M., & Fischhoff, B. (2007). Individual differences in adult decision-making competence. *Journal of Personality and Social Psychology*, *92*(5), 938–956. https://doi.org/10.1037/0022-3514.92.5.938

Cipora, K., & Soltanlou, M. (2021). Direct and conceptual replication in numerical cognition. *Journal of Numerical Cognition*, *7*(3), 240–247. https://doi.org/10.5964/jnc.7799

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.

Errington, T. M., Mathur, M., Soderberg, C. K., Denis, A., Perfito, N., Iorns, E., & Nosek, B. A. (2021). Investigating the replicability of preclinical cancer biology. *ELife*, *10*, e71601. https://doi.org/10.7554/eLife.71601

Estrada, C. A., Martin-Hryniewicz, M., Collins, C., Byrd, J. C., & Peek, B. T. (2004). Literacy and Numeracy Skills and Anticoagulation Control. *The American Journal of the Medical Sciences*, *328*(2), 88–93. https://doi.org/10.1097/00000441-200408000-00004

Estrada-Mejia, C., de Vries, M., & Zeelenberg, M. (2016). Numeracy and wealth. *Journal of Economic Psychology*, *54*, 53–63. https://doi.org/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. https://doi.org/10.1177/0272989X07304449

Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

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

Garcia-Retamero, R., Andrade, A., Sharit, J., & Ruiz, J. G. (2015). Is Patients' Numeracy Related to Physical and Mental Health? *Medical Decision Making*, *35*(4), 501–511. https://doi.org/10.1177/0272989X15578126

Garcia-Retamero, R., Cokely, E. T., & Hoffrage, U. (2015). Visual aids improve diagnostic inferences and metacognitive judgment calibration. *Frontiers in Psychology*, *6*, 1–12. https://doi.org/10.3389/fpsyg.2015.00932

Garcia-Retamero, R., Sobkow, A., Petrova, D., Garrido, D., & Traczyk, J. (2019). Numeracy and Risk Literacy: What Have We Learned so Far? *The Spanish Journal of Psychology*, *22*, E10. https://doi.org/10.1017/sjp.2019.16

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. https://doi.org/10.1017/S1930297500004952

Ginde, A. A., Clark, S., Goldstein, J. N., & Camargo, C. A. (2008). Demographic disparities in numeracy among emergency department patients: Evidence from two multicenter studies. *Patient Education and Counseling*, *72*(2), 350–356. https://doi.org/10.1016/j.pec.2008.03.012

Houts, C. R., & Knoll, M. A. Z. (2020). The Financial Knowledge Scale: New Analyses, Findings, and Development of a Short Form. *Journal of Consumer Affairs*, *54*(2), 775–800. https://doi.org/10.1111/joca.12288

Klein, R. A., Vianello, M., Hasselman, F., Adams, B. G., Adams, R. B., Alper, S., Aveyard, M., Axt, J. R., Babalola, M. T., Bahník, Š., Batra, R., Berkics, M., Bernstein, M. J., Berry, D. R., Bialobrzeska, O., Binan, E. D., Bocian, K., Brandt, M. J., Busching, R., … Nosek, B. A. (2018). Many Labs 2: Investigating Variation in Replicability Across Samples and Settings. *Advances in Methods and Practices in Psychological Science*, *1*(4), 443–490. https://doi.org/10.1177/2515245918810225

Lakens, D., & Caldwell, A. R. (2021). Simulation-Based Power Analysis for Factorial Analysis of Variance Designs. *Advances in Methods and Practices in Psychological Science*, *4*(1), 251524592095150. https://doi.org/10.1177/2515245920951503

LeBel, E. P., Berger, D., Campbell, L., & Loving, T. J. (2017). Falsifiability is not optional. *Journal of Personality and Social Psychology*, *113*(2), 254–261. https://doi.org/10.1037/pspi0000106

Lipkus, I. M., & Peters, E. (2009). Understanding the Role of Numeracy in Health: Proposed Theoretical Framework and Practical Insights. *Health Education & Behavior*, *36*(6), 1065–1081. https://doi.org/10.1177/1090198109341533

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

Lohre, E., Sobkow, A., Hohle, S. M., & Teigen, K. H. (2019). Framing experts' (dis)agreements about uncertain environmental events. *Journal of Behavioral Decision Making*, *32*(5), 564–578. https://doi.org/10.1002/bdm.2132

Makowski, D., Ben-Shachar, M., Patil, I., & Lüdecke, D. (2020). Methods and Algorithms for Correlation Analysis in R. *Journal of Open Source Software*, *5*(51), 2306. https://doi.org/10.21105/joss.02306

Mandel, D. R., & Kapler, I. V. (2018). Cognitive Style and Frame Susceptibility in Decision-Making. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01461

Mathur, M. B., & VanderWeele, T. J. (2019). Challenges and suggestions for defining replication "success" when effects may be heterogeneous: Comment on Hedges and Schauer (2019). *Psychological Methods*, *24*(5), 571–575. https://doi.org/10.1037/met0000223

Mondal, S. (2021). Sensitivity of numerate individuals to large asymmetry in outcomes: A registered replication of Traczyk et al. (2018). *Decyzje*, *2021*(35), 5–25. https://doi.org/10.7206/DEC.1733-0092.150a

Mondal, S., Lenda, D., & Traczyk, J. (2024). Recurring Suboptimal Choices Result in Superior Decision Making. *Decision*. https://doi.org/10.1037/dec0000240

Mondal, S., & Traczyk, J. (2023). Conditionality of adaptiveness: Investigating the relationship between numeracy and adaptive behavior. *Journal of Economic Psychology*, *96*, 102611. https://doi.org/10.1016/j.joep.2023.102611

Nelson, W., Reyna, V. F., Fagerlin, A., Lipkus, I., & Peters, E. (2008). Clinical implications of numeracy: Theory and practice. *Annals of Behavioral Medicine*, *35*(3), 261–274. https://doi.org/10.1007/s12160-008-9037-8

Patalano, A. L., Zax, A., Williams, K., Mathias, L., Cordes, S., & Barth, H. (2020). Intuitive symbolic magnitude judgments and decision making under risk in adults. *Cognitive Psychology*, *118*(October 2019), 101273. https://doi.org/10.1016/j.cogpsych.2020.101273

Persson, E., Andersson, D., Koppel, L., Västfjäll, D., & Tinghög, G. (2021). A preregistered replication of motivated numeracy. *Cognition*, *214*, 104768. https://doi.org/10.1016/j.cognition.2021.104768

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. (2020). *Innumeracy in the Wild: Misunderstanding and Misusing Numbers* (1st ed.). Oxford University Press. https://doi.org/10.1093/oso/9780190861094.001.0001

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. https://doi.org/10.1037/pspp0000019

Peters, E., Fennema, M. G., & Tiede, K. E. (2019). The loss-bet paradox: Actuaries, accountants, and other numerate people rate numerically inferior gambles as superior. *Journal of Behavioral Decision Making*, *32*(1), 15–29. https://doi.org/10.1002/bdm.2085

Peters, E., Hart, P. S., & Fraenkel, L. (2011). Informing Patients: The Influence of Numeracy, Framing, and Format of Side Effect Information on Risk Perceptions. *Medical Decision Making*, *31*(3), 432–436. https://doi.org/10.1177/0272989X10391672

Peters, E., & Levin, I. P. (2008). Dissecting the risky-choice framing effect: Numeracy as an individual-difference factor in weighting risky and riskless options. *Judgment and Decision Making*, *3*(6), 435–448.

Peters, E., & Shoots-Reinhard, B. (2022). Numeracy and the Motivational Mind: The Power of Numeric Self-efficacy. *Medical Decision Making*, 0272989X221099904. https://doi.org/10.1177/0272989X221099904

Peters, E., Tompkins, M. K., Knoll, M. A. Z., Ardoin, S. P., Shoots-Reinhard, B., & Meara, A. S. (2019). Despite high objective numeracy, lower numeric confidence relates to worse financial and medical outcomes. *Proceedings of the National Academy of Sciences*, *116*(39), 19386–19391. https://doi.org/10.1073/pnas.1903126116

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

Primi, C., Morsanyi, K., Chiesi, F., Donati, M. A., & Hamilton, J. (2016). The Development and Testing of a New Version of the Cognitive Reflection Test Applying Item Response Theory (IRT). *Journal of Behavioral Decision Making*, *29*(5), 453–469. https://doi.org/10.1002/bdm.1883

R Core Team. (2022). *R: a language and environment for statistical computing* [Manual]. R Foundation for Statistical Computing. https://www.R-project.org/

Reyna, V. F., & Brust-Renck, P. G. (2014). A review of theories of numeracy: Psychological mechanisms and implications for medical decision making. *Numerical Reasoning in Judgments and Decision Making about Health*, 215–251. https://doi.org/10.1017/cbo9781139644358.011

Reyna, V. F., & Brust-Renck, P. G. (2020). How representations of number and numeracy predict decision paradoxes: A fuzzy-trace theory approach. *Journal of Behavioral Decision Making*, *33*(5), 606–628. https://doi.org/10.1002/bdm.2179

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. https://doi.org/10.1037/a0017327

Rolison, J. J., Morsanyi, K., & Peters, E. (2020). Understanding Health Risk Comprehension: The Role of Math Anxiety, Subjective Numeracy, and Objective Numeracy. *Medical Decision Making*, *40*(2), 222–234. https://doi.org/10.1177/0272989X20904725

Schley, D. R., & Peters, E. (2014). Assessing “Economic Value” Symbolic-Number Mappings Predict Risky and Riskless Valuations. *Psychological Science*, *25*(3), 753–761. https://doi.org/10.1177/0956797613515485

Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H. G. (1997). The role of numeracy in understanding the benefit of screening mammography. *Annals of Internal Medicine*, *127*(11), 966–972. https://doi.org/10.7326/0003-4819-127-11-199712010-00003

Simonsohn, U. (2015). Small Telescopes: Detectability and the Evaluation of Replication Results. *Psychological Science*, *26*(5), 559–569. https://doi.org/10.1177/0956797614567341

Skagerlund, K., Forsblad, M., Tinghög, G., & Västfjäll, D. (2021). Decision-making competence and cognitive abilities: Which abilities matter? *Journal of Behavioral Decision Making*, *35*(1), e2242. https://doi.org/10.1002/bdm.2242

Sobkow, A., Fulawka, K., Tomczak, P., Zjawiony, P., & Traczyk, J. (2019). Does mental number line training work? The effects of cognitive training on real-life mathematics, numeracy, and decision making. *Journal of Experimental Psychology: Applied*, *25*(3), 372–385. https://doi.org/10.1037/xap0000207

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.

Sobkow, A., Jurewicz, K., & Traczyk, J. (in press). Replications in Judgment and Decision Making: A Critical Review and Future Pathways. In *Decision-Making in Life and Work: Foundations, Strategies, and Current Neuroscience.* Springer Nature.

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

Sobkow, A., Surowski, M., Olszewska, A., Antoniewska, N., Barcik, K., Bartkiewicz, U., Brzeska, A., Brzozowska, A., Budrewicz, O., Choja, J., Choma, K., Chorbotowicz, P., Filimoniak, M., Filip, Ł., Gambuś, P., Gierlik, W., Gonczar, T., Goryczka, K., Góra, M., … Traczyk, J. (2022). Conceptual replication study of fifteen JDM effects: Insights from the Polish sample. *Polish Psychological Bulletin*, 138–151. https://doi.org/10.24425/ppb.2022.141862

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. https://doi.org/10.1017/S1930297500009244

Vazire, S., & Nosek, B. (2023). Introduction to special topic "Is psychology self-correcting? Reflections on the credibility revolution in social and personality psychology." *Social Psychological Bulletin*, *18*, e12927. https://doi.org/10.32872/spb.12927

Waldrop-Valverde, D., Jones, D. L., Gould, F., Kumar, M., & Ownby, R. L. (2010). Neurocognition, Health-Related Reading Literacy, and Numeracy in Medication Management for HIV Infection. *AIDS Patient Care and STDs*, *24*(8), 477–484. https://doi.org/10.1089/apc.2009.0300

Weller, J. A., Dieckmann, 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

Zhu, M., & Feldman, G. (2023). Revisiting the Links Between Numeracy and Decision Making: Replication Registered Report of Peters et al. (2006) With an Extension Examining Confidence. *Collabra: Psychology*, *9*(1), 77608. https://doi.org/10.1525/collabra.77608

Contributions

- Contributed to conception and design: AS and JT (leading), rest of the co-authors (supporting)
- Contributed to acquisition of data: all authors equal
- Contributed to analysis and interpretation of data: AS, JT, KJ, TZ, SM, AP (leading), rest of the co-authors (supporting)
- Drafted and/or revised the article: AS, JT, KJ, TZ, SM, AP (leading), rest of the co-authors (supporting)
- Approved the submitted version for publication: all authors equal

Acknowledgments

We want to thank the Dean of the Faculty of Psychology in Wroclaw (SWPS University) for the support for this project.

Funding information

This work was financially supported by the SWPS University and by the National Science Centre, Poland (grant number 2018/31/D/HS6/02899).

Competing interests

The funding sources exerted no influence on the formulation of the study design, the process of data collection, analysis, and interpretation, the composition of the paper, or the decision to submit it for publication. No additional competing interests are to be disclosed.

Data accessibility statement

All the stimuli, presentation materials, participant data, and analysis scripts can be found on this paper's project page on the Open Science Framework (https://osf.io/927bx/?view_only=f143e1b75ca64f249968984f9f6f4725).