The Relations of Objective Numeracy and Subjective Numeracy to Financial Outcomes over Time

THESIS

Presented in Partial Fulfillment of the Requirements for the Degree Master of Arts in the Graduate School of The Ohio State University

By

Mary Kathleen Tompkins

Graduate Program in Psychology

The Ohio State University

2015

Master's Examination Committee:

Ellen Peters, Advisor; Kentaro Fujita; Ian Krajbich

Abstract

No known research has examined the possible separable impacts of numeric competencies on financial decision outcomes, and it is unclear which skills might matter most. In a longitudinal study using undergraduate participants (N=749), we examined the abilities of objective (ONS) and subjective numeracy (SNS) to predict financial outcomes. Data collection points were the beginning and end of a semester, as well as a one-year follow-up. We hypothesized that being more proficient in ONS and/or SNS would be associated with better financial outcomes at baseline due to ONS's relation to performing number comparisons and calculations and SNS's motivational and emotional associations with numbers. Additionally, in a quasi-experimental design, we manipulated numeracy (taking a statistics course vs. not) to investigate possible improvements, in ONS, SNS, and financial outcomes. We hypothesized that taking a statistics course (vs. not) would be associated with positive changes in ONS and SNS from the beginning to end of the semester, and these ONS/SNS changes would predict financial outcome changes during that time. Lastly, we hypothesized that taking a statistics course would improve financial outcomes over a year later. Results indicated that greater SNS (but not ONS) predicted financial outcomes at baseline and changes in financial outcomes from the beginning to the end of the semester. Additionally, taking the statistics course appeared to protect students from detrimental changes in SNS over the semester and from

detrimental changes in financial outcomes over a year later. The present findings should be taken into consideration when developing strategies to improve financial outcomes.

Acknowledgments

Funding for this project was provided from the National Science Foundation (SES-1155924) and The Ohio State University's Decision Sciences Collaborative. I am grateful to many researchers and undergraduate research assistants for their help with this study. Thank you to Ellen Peters, Martin Tusler, Dan Schley, Louise Meilleur, Alex Sinayev, Joe Macisco, David Weiner, Niraj Patel, Kara Ruder, Kelsey Louise, Elizabeth Seguin, Nick Larbes, Leayh Abel, Qijia Chen, and Tomas Moreno-Vasquez.

Vita

2011..B. S. Psychology, The Ohio State University

2013 to 2014 ....University Graduate Fellow, The Ohio State University

2014..Graduate Research Associate, Department of Psychology, The Ohio State University

2015 to present ....Graduate Teaching Associate, Department of Psychology, The Ohio State University

Publications

Peters, E., Meilleur, L., Tompkins, M. K. (2014). Numeracy and the Affordable Care Act: Opportunities and challenges. Institute of Medicine (IOM). Roundtable on Health Numeracy. Washington, DC: The National Academies Press.

Fields of Study

Major Field: Psychology

Table of Contents

List of Tables

List of Figures

Introduction

Numbers are ubiquitous in financial decisions, from simple, everyday decisions (e.g., balance keeping) to complex decisions with long-term consequences (e.g., credit card and mortgage borrowing, saving and investing for retirement). However, individuals differ in their ability to process and use numbers and many Americans experience negative financial outcomes in life. The national statistics on Americans' debt for 2014 were alarming. In particular, roughly 910,000 American individuals filed for non-business bankruptcy (acquired for personal, family, or household purposes; U.S. Courts, 2014). Additionally, Americans' outstanding credit debt (mostly comprised of credit card loans) totaled to $889.5 billion in 2014 (U.S. Federal Reserve, 2015). The ability to make good financial decisions is important for individuals and society because many governments and employers are putting more and more responsibilities such as saving, investing, and borrowing onto individuals (Lusardi, 2012).

Prior studies have examined the importance of greater numeric competence in experiencing better financial outcomes. For example, greater objective numeracy (defined as the ability to understand and use probabilistic and mathematical concepts) has been found to be associated with greater retirement savings and investment portfolios, fewer debt-repayment errors, and greater avoidance of predatory loans (Banks & Oldfield, 2007; Sinayev & Peters, 2015; Soll, Keeney, & Larrick, 2013). Although objective numeracy is clearly important to financial outcomes, having motivation (a state of energizing and impelling behavior) and confidence (belief in one's ability) to deal with numbers may also be important. One way to assess these constructs is through subjective numeracy (measured as perceptions of one's objective ability and a preference for numbers). If subjective numeracy also matters to experiencing good financial outcomes, which numeric competence matters most?

The present research examined the impact of both objective and subjective numeracy on financial outcomes. Additionally, we attempted to manipulate objective and subjective numeracy to improve financial outcomes over time. First, I will provide an overview of objective numeracy followed by subjective numeracy (e.g., what is it? how is it measured? how has it been used?). Next, I will summarize recent studies illustrating objective and subjective numeracy's dissociable roles in judgment and decision processes. Lastly, I will discuss why both objective and subjective numeracy may matter in financial domains, specifically, and I will outline the hypotheses of the current research.

Objective Numeracy and Decision-Making Competence

Many Americans have poor numeric skills and even the highly educated can be innumerate (Lipkus, Samsa, Rimer, 2001). Objective numeracy is defined as the ability to understand and use basic probability and numerical concepts and it is measured with a math test (e.g., "If the chance of getting a disease is 10% how many people would be expected to get the disease out of 100?"; Lipkus, Samsa, Rimer, 2001; Weller, Dieckmann, Tusler, Mertz, Burns, & Peters, 2013). Based on the latest estimates from the

Organization for Economic Co-operation and Development (OECD), roughly 29% of American adults (approximately 68 million) lack the basic number skills needed to locate easily identifiable quantitative information and use it to solve simple, one-step problems, such as adding two numbers to complete an ATM deposit slip (OECD, 2013; Peters, Tompkins, Meilleur, & Bjalkebring, *in review*).

Most numeracy research has shown that individuals who are more objectively numerate tend to have better decision-making competence (Peters et al., 2006; Peters, 2012). For example, individuals higher in objective numeracy (ONS) tend to be less susceptible to framing effects (e.g., exam scores labeled 74% correct perceived as higher quality vs. scores labeled 26% incorrect), show greater sensitivity to numbers, and draw more affective meaning from numbers compared to less numerate individuals (Lipkus et al., 2010; Peters et al., 2006). Additionally, individuals lower in objective numeracy are more sensitive to non-numeric information such as narratives and mood states (Peters, Dieckmann, et al., 2009; Vastfjall, Peters, and Starmer, in preparation). These associations hold even after controlling for a variety of intelligence proxies (e.g., SAT scores, working memory, vocabulary; Peters et al., 2006; Peters & Bjalkebring, 2015).

Subjective Numeracy

Numeracy has also been assessed with subjective numeracy measures. Subjective numeracy measures were developed as proxies for ONS because they are easier to administer and are less frustrating and stressful for participants to complete (Fagerlin, Zikmund-Fisher, Ubel, Jankovic, Derry, & Smith, 2007). They are useful for testing in medical environments (e.g., hospitals) where administration is burdensome and test taking can be especially aversive. The most commonly used measure, the subjective numeracy scale (SNS; Fagerlin et al., 2007), consists of two subscales: subjective ability and number preference. The ability subscale measures an individual's beliefs about his/her objective numeracy skills (e.g., "How good are you at working with percentages?"; Fagerlin et al., 2007). The number preference subscale measures an individual's preference for numeric information (e.g., "When people tell you the chance of something happening, do you prefer that they use words ('it rarely happens') or numbers ('there's a 1% chance')?"). The developers of the SNS scale have encouraged its use as a proxy for objective numeracy because its predictive ability approached that of objective numeracy for numeric comprehension in some studies (Zikmund-Fisher, Smith, et al., 2007). Some researchers, however, have found subjective measures to be poor diagnostic indicators of objective numeracy (Dolan, Cherkasky, Li, Chin, & Veazie, 2014; Liberali, Reyna, et al., 2012; Nelson, Moser, & Han, 2012).

Dissociations Among Numeric Competencies in Judgments and Decisions

In the literature, researchers have typically treated objective and subjective numeracy measures as the same numeric competence (e.g., Glasgow et al., 2014; Hess, Visschers, Siegrist, 2011;Hanoch, Miron-Shatz, Himmelstein, 2010; Keller, Siegrist, & Visschers, 2009; Schley & Fujita, 2014; Tait, Voepel-Lewis, et al., 2010, Zikmund-Fisher, Smith et al., 2007; Zikmund-Fisher, Fagerlin et al., 2008). However, objective and subjective numeracy appear to be related, but separable constructs. Recent research revealed differences between objective and subjective numeracy in their ability to predict the processing and use of numbers in judgments and decisions (Peters & Bjalkebring, 2015). In particular, Peters and Bjalkebring found that subjective numeracy was associated with emotional reactions to math. Additionally, higher- vs. lower-SNS individuals appeared more motivated and/or confident in numeric tasks. In comparison, objective numeracy was associated with performing number operations, such as number comparisons and calculations, across a series of tasks.

In a memory task from Peters and Bjalkebring, for example, participants were asked to memorize a series of numbers and objects in a specific location (e.g., "20 skulls on a shelf"). Later, they were asked to recall the number and object when given the location. Results indicated that SNS significantly predicted non-responses to the number recall question. Specifically, lower-SNS participants were less likely to provide a response than higher-SNS participants. Objective numeracy did not provide predictive power for non-responses. The authors interpreted these findings as being due to those higher in subjective numeracy being more willing to provide a guess compared to those lower in subjective numeracy because they had greater confidence and/or motivation when dealing with numbers. Additionally, Peters and Bjalkebring did not find a relation between subjective numeracy and non-responses on a non-numeric vocabulary task, suggesting subjective numeracy's influence may be specific to numeric tasks.

Not only do higher-SNS individuals appear more confident and/or motivated in numeric domains, but they also appear to have more positive emotional reactions to numbers. Peters and Bjalkebring had participants self-report their emotions towards math and found that positive math emotions were a stronger predictor of subjective numeracy than was objective numeracy. Furthermore, they also demonstrated an association between subjective numeracy and affective responses in a numeric bet task. Participants rated the attractiveness of a bet in either a loss (7/36 to win $9; 29/36 chances to lose 5₵) or no loss (7/36 to win $9; otherwise win nothing) condition. Results indicated that higher-SNS individuals rated the bet as more attractive than lower-SNS individuals across both conditions (controlling for ONS, general intelligence, and a positive-rating tendency across tasks), suggesting that SNS was associated with more affective responses to numbers.

Prior research also suggests that individuals with greater SNS may value numeric information more than lower-SNS individuals due to affective responses to numbers. Specifically, greater subjective numeracy was associated with a greater willingness to pay for BRCA1/2 genetic testing among high-risk individuals with a family history of breast cancer (Miron-Shatz, Hanoch, Doniger, Omer, & Ozanne, 2014). In comparison, ONS was not associated with willingness to pay for the genetic test, even though both ONS and SNS correlated in this study. Lower-SNS individuals may have perceived less value for the genetic test due to the negative emotional reactions of receiving numeric test results.

Additional research has shown that subjective numeracy also predicts preferences for providing and receiving numeric information. For example, lower-SNS individuals preferred receiving risk disclosure information with words rather than numbers (Couper & Singer, 2009). Also, lower vs. higher-SNS physicians were less likely to use numbers to explain health screening tests results to patients (Anderson, Obrecht, Chapman, Driscoll, & Schulkin, 2011). In both studies, ONS was less predictive of these preferences in separate models. These preferences for receiving and providing words versus numeric information among lower-SNS individuals may be due to negative emotional reactions or a lack of confidence when working with numbers.

In comparison to subjective numeracy, Peters and Bjalkebring demonstrated that objective numeracy was associated with performing simple number comparisons. For example, in the previously described numeric bet task, higher-ONS participants rated the attractiveness of the bet as greater when there was an added loss (29/36 chances to lose 5₵) than when there was no loss (otherwise lose nothing), illustrating ONS's relation to performing simple number comparisons (comparing the $9 with the 5-cent loss). In addition to performing simple number comparisons, Peters and Bjalkebring found that ONS was associated with performing more complex number operations. For example, in a risky-gamble valuation task, participants indicated how much money they needed to win for sure to be indifferent between a 50% chance of winning some amount of money and a 50% chance of winning a larger amount. Controlling for subjective numeracy and other variables, higher-ONS individuals provided more valuations close to and equal to expected value than lower-ONS individuals. No effects of subjective numeracy existed. These results demonstrated the association between objective numeracy and performing number comparisons and explicit calculations.

Overall, this body of research provided a theoretical shift from the existence of a single numeric competence (numeracy) to multiple numeric competencies (e.g., objective and subjective numeracy). Objective and subjective numeracy appear to be related, but separable constructs with unique roles in judgment and decision-making processes. No known research has examined the possible separable impacts of objective and subjective numeracy on financial outcomes, and it is unclear which skill might matter most.

The Importance of Objective Numeracy in Financial-Decision Outcomes

The previously discussed literature suggests that objective numeracy skills may be essential to obtaining good financial outcomes due to superior numerical reasoning, such as performing number comparisons and calculations. A number of studies have shown the importance of objective numeric skills over and above other cognitive abilities (e.g., literacy, memory, verbal fluency, attention, visual search) in financial domains, although objective numeracy was the only numeric competence assessed (Banks & Oldfield, 2007; Banks, O'Dea, & Oldfield, 2010; Gerardi, Goette, & Meier, 2013; Sinayev & Peters, 2015; Smith, McArdle, & Willis, 2010; Soll, Keeney, & Larrick, 2013). In particular, lower objective numeracy has been associated with poorer knowledge and understanding of pension accrual rates, fewer retirement savings and investment portfolios, and lower net financial wealth (defined as the value of all financial assets, excluding private and state pensions and housing, minus any outstanding non-mortgage debt; Banks & Oldfield, 2007; Banks, O'Dea, & Oldfield, 2010; Smith, McArdle, & Willis, 2010). For example, for families in which both spouses got all objective numeracy questions correct, total household wealth was eight times higher (1.7 million dollars in total wealth) compared to families in which both spouses got all objective numeracy questions incorrect (approximately $200,000 in total wealth; Smith, McArdle, & Willis, 2010).

Poorer objective numeracy has also been associated with a propensity to default on mortgages, even after controlling for a broad set of sociodemographic variables and cognitive skills (Gerardi, Goette, & Meier, 2013). Interestingly, there was no evidence in this study that objective ability impacted mortgage outcomes due to the initial choice of the mortgage contract. Specifically, objective ability did not predict various mortgage attributes that were also predictors of mortgage default (e.g., mortgage origination amount, loan-to-value ratio). Additionally, the inclusion of various mortgage attributes that were predictors of mortgage default in the model did not change the relationship between objective numeracy and mortgage default. These findings suggested that the relationship between objective numeracy and mortgage defaults was due to behaviors unrelated to the initial mortgage choice. Perhaps individuals with limited objective numeracy were more likely to default on their mortgages because they underestimated other financial obligations and they overestimated their ability to pay off the mortgage amount. Prior research has shown that individuals with lower-ONS scores made more errors when estimating debt repayments compared to those with higher scores (Soll, Keeney, & Larrick, 2013). Specifically, lower-ONS individuals tended to underestimate the monthly payments required to pay off credit card debt in a three year time period.

These studies provide evidence for the importance of objective numeracy skills in financial decision outcomes, although objective numeracy was the only numeric competence assessed. However, based on the results of Peters and Bjalkebring (2015), we suspect the financial knowledge and judgment findings are primarily due to objective abilities being related to the calculation components of these questions (e.g., pension accrual rates, debt repayment estimates).

The Potential Importance of Subjective Numeracy in Financial-Decision Outcomes

Subjective numeracy, on the other hand, may be crucial for experiencing good financial outcomes due to its emotional associations to numeric tasks and the behavioral consequences of self-efficacy beliefs. Drawing from self-efficacy theory, self-efficacy beliefs have behavioral consequences, such as approach or avoidance behavior and persistence in the face of obstacles (Bandura, 1977; Betz, 2013). As a consequence, lower-SNS individuals may be less confident to approach number-heavy financial decisions (e.g., selecting retirement funds) and may be more likely to avoid and less likely to persist on difficult or boring financial tasks (e.g., budgeting and monitoring bank statements). Additionally, lower-SNS individuals may be less likely to seek out financial information or financial advice due to more negative emotional reactions to numbers (e.g., Peters & Bjalkebring, 2015), less valuation of numeric information (e.g., Miron-Shatz et al., 2014), or a decreased preference for receiving numeric information (e.g., Couper & Singer, 2009). Furthermore, greater self-efficacy beliefs have been shown to be an important predictor of information seeking behaviors in other domains (Hong, 2006).

Another behavioral consequence of greater subjective numeracy may simply be making a decision (vs. not). Experimental studies in consumer psychology found that greater subjective knowledge (independent of objective knowledge) was associated with making a choice vs. not (Hadar, Sood, & Fox, 2013). In a clever design, Hadar and colleagues provided participants with a descriptive paragraph about a hypothetical investment and manipulated participants' subjective knowledge for the investment by crossing out a sentence in the description. The crossed-out sentence was excluded in the control condition; thus, objective knowledge was held constant because both conditions received the same information about the investment. Results indicated that participants in the high-subjective knowledge condition were more likely to make an investment choice compared to those in the control condition. Using various subjective knowledge manipulations, subsequent studies found that participants were more willing to enroll in a hypothetical retirement program and were more likely to select a hypothetical mutual fund if they were higher in subjective knowledge. Thus, individuals more confident in their numerical abilities may be more likely to make financial choices vs. take no action. Simply making choices may be beneficial as some financial mistakes appear to be the result of inaction (e.g., a lack of 401(k) savings behaviors). Such patterns have motivated research on various kinds of automatic enrollment programs (Madrian & Shea, 2001).

In summary, both objective and/or subjective numeracy skills may be important to financial decision outcomes. Improved outcomes could be due to superior number calculations (ONS) and/or greater motivation, confidence, and positive reactions to numbers (SNS). The first aim of the current research was to examine the predictive power of objective and subjective numeracy in financial outcomes.

Hypothesis 1:

Participants lower in one or both numeric competencies would report experiencing worse financial outcomes.

We predicted that individuals lower in one or both numeric competencies would report experiencing worse financial outcomes, controlling for intelligence proxies and demographic variables (Hypothesis 1).

The second aim of the current research was to attempt to manipulate numeric competencies with statistics training. We predicted that taking a statistic course would improve objective numeracy (Hypothesis 2). Statistical reasoning can be improved. For example, Fong, Krantz, and Nisbett (1986) examined the influence of statistical training on the use of the law of large numbers to make statistical responses in a variety of hypothetical scenarios. Results indicated that statistical training helped participants make better quality and a greater quantity of statistical responses. Additionally, the frequency and quality of statistical responses increased as levels of statistics training increased (e.g., no training, undergraduate training, or graduate training).

We predicted that taking a statistics course would also improve subjective numeracy (Hypothesis 2), as prior studies have shown that training courses increased individuals' subjective beliefs about their ability, even if the courses did not lead to improved objective performance (Gibbs, 1981; Main, 1980; as cited in Ross, 1989, p.348). For example, university students who took a study skills course reported significantly greater improvement in their study skills than students waitlisted for the course, even though academic grades were not affected by the program (Ross, 1989).

Hypothesis 2:

Taking a statistics course (vs. not) would improve objective and/or subjective numeracy from the beginning to the end of the semester.

The third aim was to examine if changes in numeric competencies were associated with changes in financial outcomes over time (Hypothesis 3). If objective and/or subjective numeracy predicted financial outcomes initially (Hypothesis 1), then we would also expect to see changes in numeric competencies being associated with changes in financial outcomes. Specifically, we predicted that improvements in objective and/or subjective numeracy from the beginning to the end of the semester would be associated with decreases in bad financial outcomes experienced. Stated differently, decrements in numeric competencies from the beginning to the end of the semester would be associated with increases in bad financial outcomes.

Hypothesis 3:

Positive changes in objective and/or subjective numeracy would predict better financial outcomes.

The fourth aim of the present research was to examine if taking a statistics course improved financial outcomes from the beginning to the end of the semester (controlling for intelligence proxies and demographics) through improvements in one or both numeric competencies (Hypothesis 4; illustrated in Figure 1).

Hypothesis 4:

Taking a statistics course (vs. not) would predict improved financial outcomes through changes in objective and/or subjective numeracy.

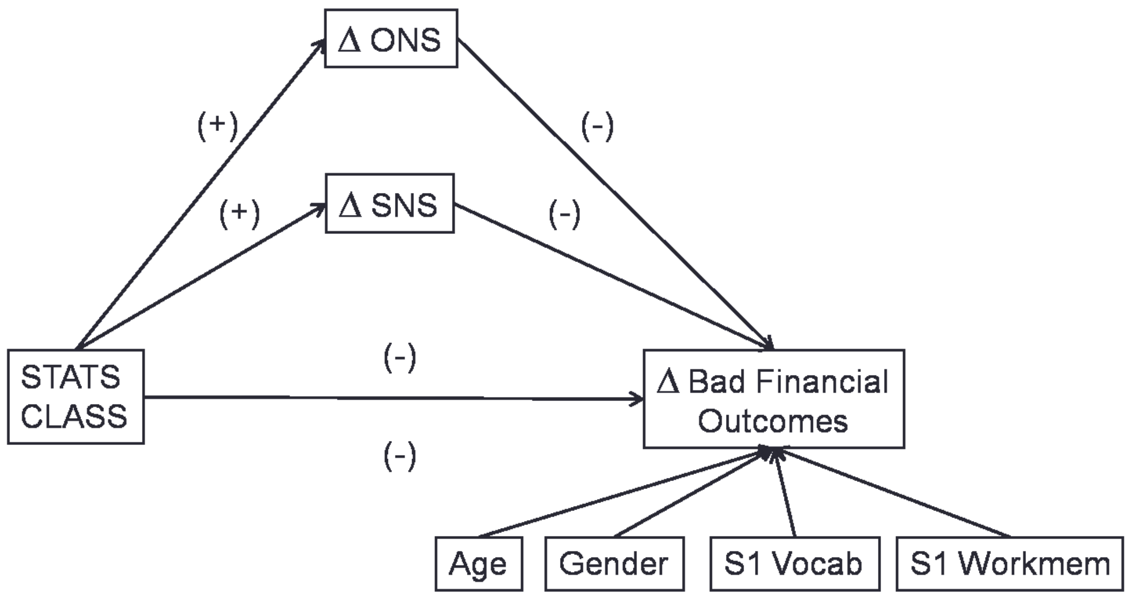

Figure 1. Hypothesis 4's predicted indirect effects controlling for general intelligence and demographics

Lastly, the fifth aim was to examine if taking a statistics course would be associated with better financial outcomes through changes in ONS or SNS over a year later, controlling for intelligence proxies and demographics (Hypothesis 5). Specifically, we predicted that taking the statistics course (vs. not) would be associated with decreases in bad financial outcomes through improvements in ONS or SNS from the beginning of the semester to over one year later. We were interested in this follow-up to examine any long-term effects of statistics training on financial outcomes.

Hypothesis 5:

Taking a statistics course (vs. not) would be associated with better financial outcomes through changes in objective and/or subjective numeracy over a year later.

Method

Participants

Eight-hundred thirty-six university undergraduates enrolled in a statistics or introductory psychology course were recruited to participate in the study. Students in the introductory psychology course were recruited through the Research Experience Program (REP).

Over the course of the study, Ohio State changed from a quarter system to a semester system. We recruited participants from three statistics courses. Two statistics courses were offered in the fall, lasting approximately eleven weeks on the quarter system. The third statistics course was offered in the winter, lasting approximately nine weeks on the quarter system. The statistics courses were taught by one of two instructors. We recruited REP participants in the same winter quarter and following spring quarter, lasting approximately nine and ten weeks, respectively. Lastly, we recruited REP participants the following fall semester, lasting approximately fifteen weeks, although we collected data over a period of thirteen weeks. Although there were two instructors for the statistics courses and the length of time varied for each academic term, academic term and instructor had no effects on the hypothesized results and thus will not be discussed further. [1] For simplicity, we will use the word semester going forward to refer to an academic term.

---

[1] All hypotheses tests were re-analyzed with academic term (fall1, winter, spring, fall2) and instructor (instructor 1, 2) included as covariates in the models. Academic term and instructor had no significant effects on results and were removed from the models.

Data collection points included the beginning and end of a semester, as well as a one-year follow-up in this longitudinal design. Participants enrolled in the statistics course received up to 5% extra credit in their course for participating in the study. Participants enrolled in a psychology course received course credit for participating in the study through the Research Experience Program. To improve follow-up retention, all participants had the chance to win one of five $100 cash prizes for completing the end of the semester follow-up. Lastly, participants received a $15 Amazon gift card for completing a one-year follow-up. Individuals from both courses had the option of completing paper assignments to earn the same amount of extra credit or course credit as an alternative to participating in the study.

Procedure

Participants completed measures of subjective numeracy, objective numeracy, financial outcomes, working memory, vocabulary, and demographics in a 2-hour lab session at the beginning of the semester (Session 1). The working memory and vocabulary tests were included as controls for general intelligence. At the end of the semester, participants completed subjective numeracy, objective numeracy, and financial outcomes in an online survey (Session 2). Lastly, participants completed subjective numeracy, objective numeracy, and financial outcomes in an online survey approximately a year following the end of the semester (Session 3). We accepted one-year follow-ups at a minimum of twelve months and a maximum of eighteen months from the end of the semester, but the bulk of participants responded to the follow-up about 13 months following the end of the semester. Other tasks were also conducted at each time point, but will not be included in this report.

Measures

Subjective Numeracy Scale (SNS; Fagerlin et al., 2007). Participants provided 8 ratings of their perceived abilities and preferences with numbers (e.g., "How good are you at figuring out how much a shirt will cost if it's 25% off?;" "How often do you find numerical information useful?" assessed on a 6-point scale). The subjective numeracy score was calculated by taking an average across the 8 items (possible range = 1-6, see Appendix A for full measure).

Objective Numeracy Scale (ONS; Weller et al., 2013). Participants completed 18 probabilistic and mathematic items (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?"). Two items failed to record correctly in at least one session, leaving us with 16 items. Each item was scored as correct or incorrect, and the correct items were summed (possible range = 0-16, see Appendix B for full measure).

Financial Outcomes. To assess financial outcomes, we measured participants' experiences with 10 negative outcomes (e.g., "In the past year, have you had an overdraft or bounced check from your account?") and knowledge questions ("e.g., Do you know your credit card balance?"). To compute an overall financial outcome score (FOS), bad financial outcomes were coded as 1, summed, divided by the total number of outcomes the participant had the opportunity to experience, and multiplied by 100 (possible range = 0-100%, see Appendix C for full measure). The absolute number of negative financial outcomes was not used because not all questions applied to all of the participants. For example, if a participant indicated that he/she did not own a credit card, then financial outcome items regarding credit cards were not relevant to the participant and did not count towards his/her financial outcome score. This scoring approach to handle variable experiences across participants is consistent with prior research (e.g., Decision Outcomes Inventory (DOI); Bruine de Bruin, Parker, & Fischhoff, 2007). Thus, the FOS is equal to the proportion of negative outcomes experienced out of those outcomes participants had the opportunity to experience. Higher scores indicated worse financial outcomes.

Working Memory Test. Participants completed a backwards letter sequence task as a measure of working memory (Wechsler, 1997). In this task, a list of letters were presented one at a time for 1 second each and participants had to type out the sequence in reverse order (e.g., for the first trial, AB was presented and participants typed BA). Participants had to correctly respond to two practice sequences to advance to the remainder of the test. At each level, an additional letter was added onto the sequence. Participants advanced through the levels up to a 9-letter sequence. The test ended after two failed attempts at a level or after completion of the 9-letter sequence. Completion of each sequence was coded as correct or incorrect and correct responses were summed to compute a working memory score (possible range = 2-9, see Appendix D for full measure).

Vocabulary Test. Participants responded to 36 items to test their vocabulary (Ekstrom et al., 1976). Participants were asked to select one of five word choices that had the same or nearly the same meaning as the target word. Correct answers were summed to compute a total vocabulary score (possible range= 0-36, see Appendix E for full measure).

Results

Correlation matrices of all study measures at each session and across sessions are available in the Appendix F-H. Table 1 shows tests of internal consistency and reliability for subjective numeracy, objective numeracy, and financial outcomes across the sessions. Cronbach's alpha, the mean of all split-half correlations, was used a measure of internal consistency for each measure (Cronbach, 1951). The observed alphas for subjective and objective numeracy were consistent with what is typically reported for these scales (Peters & Bjalkebring, 2015). For the sole purpose of computing the financial outcome alpha with complete data, we treated participants who did not have the opportunity to experience the outcome as having avoided the negative outcome (consistent with prior research; Bruine de Bruin, Parker, & Fischhoff, 2007). In all other analyses, we treated these questions as missing data for those participants. The financial outcome scale had the weakest internal consistency, but adequate test-retest reliability. All test-retest reliabilities were significant at the $p<.0001$ level. We computed Spearman-Brown reliabilities as an additional reliability test. It is similar to Cronbach's alpha, but a better tool to use when there are only two values (Eisinga, Grotenhuis, & Pelzer, 2012). Spearman-Brown reliabilities above .6 are considered acceptable. All the measures had acceptable to good Spearman-Brown reliabilities.

Table 1. Reliabilities of numeric competencies and financial outcomes

	(N)	SNS	ONS	FOS
Session 1 Cronbach's alpha	749	0.79	0.65	0.39
Session 2 Cronbach's alpha	587	0.84	0.78	0.48
Session 3 Cronbach's alpha	247	0.81	0.71	0.42
Session 1:2 test-retest	587	0.74***	0.59***	0.63***
Session 2:3 test-retest	247	0.72***	0.54***	0.52***
Session 1:3 test-retest	247	0.69***	0.52***	0.47***
Session 1:2 Spearman-Brown	587	0.85	0.74	0.77
Session 2:3 Spearman-Brown	247	0.84	0.70	0.68
Session 1:3 Spearman-Brown	247	0.82	0.68	0.64

Note. ***$p<.0001$, two sided.

Only participants who completed all relevant measures within a session were included in analyses for that session, leaving a final *N*=749 at the beginning of the semester [*N*=225 Statistics, *N*=524 REP, 64% female; mean (SD) age= 19.8(3.1)], *N*=587 at the end of the semester (78% retention from baseline; 88% of Statistics participants; 74% of REP participants), and *N*=247 over one year later (33% retention from baseline; 32% of Statistics participants; 33% of REP participants).

Questions related to major area of study were added during our second semester of running. Thus, we had major data for 63 out of the 225 Statistics participants and near complete major data for our REP participants (523 out of 524). Among the Statistic participants, the majority of participants indicated they were psychology majors (71%), 3% were social and behavioral sciences (excluding psychology), 2% were physical and natural sciences, 8% were medical, 2% were business and professional, 3% were arts and humanities, and the remaining 11% did not fit into a provided category. The REP participants were a diverse sample in terms of major areas of study. Among the REP participants, 9% were psychology, 5% were social and behavioral sciences (excluding psychology), 24% were physical and natural sciences, 15% were medical, 17% were business and professional, 4% were arts and humanities, 11% were undeclared, and the remaining 15% did not fit into a provided category.

Predictor measures were used in their continuous form for inferential statistics (i.e., SNS, ONS, Working Memory, and Vocabulary).

Session 1: Beginning of Semester

Objective and subjective numeracy were correlated as expected. Higher vs. lower ONS was associated with greater self-reported SNS ($r=.38$, $p<.0001$). Women had lower ONS ($r=-.20$, $p<.0001$) and SNS ($r=-.28$, $p<.0001$) compared to males, consistent with prior studies (e.g., Peters & Bjalkebring, 2015). Additionally, higher vs. lower vocabulary scores were associated with greater ONS ($r=.17$, $p<.0001$) and SNS ($r=.07$, $p=.04$). Similarly, higher vs. lower working memory scores were associated with greater ONS ($r=.12$, $p<.01$) and SNS ($r=.08$, $p=.03$). Descriptive data for each scale at Session 1 are shown in Table 2. On average, participants reported experiencing 18.52% of negative financial outcomes (see Appendix C for experienced outcome frequencies by item).

Table 2. Session 1 scale descriptive data (N=749)

Variable	Possible Range	Mean(SE)	Std Dev	Min	Max
SNS	1 to 6	4.47(.03)	0.78	1	6
ONS	0 to 16	11.01(.08)	2.33	2	16
Vocabulary	0 to 36	19.10(.20)	5.59	0	32
Working memory	2 to 9	5.43(.05)	1.30	2	9
FOS	0 to 100	18.52(.71)	19.56	0	100

No significant baseline difference existed between Statistics and REP participants on objective numeracy ($t=-1.06$, $p=.29$), but there were significant differences on subjective numeracy and financial outcomes. On average, Statistics participants had lower SNS scores compared to REP participants ($M_{Statistics}= 4.32$, $M_{REP}= 4.53$, $t=3.51$, $p<.001$). This baseline difference in subjective numeracy is not too surprising considering it is the beginning of the semester and some Statistics participants may be anxious about taking the course. Additionally, Statistics participants reported higher (or worse) financial outcome scores than REP participants ($M_{Statistics}= 20.76$, $M_{REP}=17.56$, $t= -2.06$, $p=.04$).[2] Given the significant differences in financial outcomes and subjective numeracy based on Statistics condition, we used change scores in analyses to examine condition effects over time. Details on computing change scores are discussed in Session 2 results.

[2] To examine if these differences were characteristic of psychology majors compared to non-psychology majors, we examined a subset of REP participants who indicated they were majoring in psychology (49 out of 524 REP participants). Among this subset of REP participants, we tested if the psychology majors had worse financial outcome and subjective numeracy scores compared to the non-psychology majors. Results were consistent with this pattern, but did not reach statistical significance (FOS: $M_{REPpsych}= 22.54$, $M_{REPother}= 16.88$, $t=-1.61$, $p=.11$, Satterthwaite method of unequal variances; SNS: $M_{REPpsych}=4.40$ $M_{REPother}= 4.55$, $t=1.27$, $p=.21$, Pooled method of equal variances).

Test of Hypothesis 1:

Participants lower in one or both numeric competencies would report experiencing worse financial outcomes (controlling for intelligence proxies) at baseline.

To test Hypothesis 1, we conducted multiple linear regression analysis with Session 1 ONS, SNS, vocabulary, working memory, age, and gender as predictors. Non-significant predictors (defined as $p>.05$) were removed one at a time, leaving SNS as the only remaining predictor ($b= -2.44$, $t= -2.66$, $p= .01$; see Table 3 for full and final model estimates).

Table 3. Multiple regression results for Session 1 financial outcomes

Variable	Full Model b(SE)	Final Model b(SE)
Intercept	**25.77 (7.99)**	**29.45(4.17)**
Subjective Numeracy	**-2.54 (1.03)**	**-2.44 (.92)**
Objective Numeracy	0.19 (0.34)	
Working memory	0.12 (0.55)	
Vocabulary	0.02 (0.13)	
Gender (1 Female, 0 Male)	0.93 (1.57)	
Age	0.02 (0.23)	
R^2	.01	.01
df	(6,741)	(1,747)
F	1.31, p=.25	7.08, p=.01

Note. Significant predictors are in bold for emphasis.

Consistent with Hypothesis 1, lower-SNS individuals experienced a greater proportion of negative financial outcomes than those higher in SNS (see Figure 2). ONS scores were not independently predictive (simple correlations with FOS were $r_{SNS}= -.10$, p= .01 and r_{ONS}= -.02, p= .58). Thus, subjective numeracy was more important than objective numeracy for predicting financial outcomes. For a person scoring lowest on SNS (score of 1), the predicted percentage of negative financial outcomes experienced was 27%, compared to only 15% for a person scoring highest on SNS (score of 6).

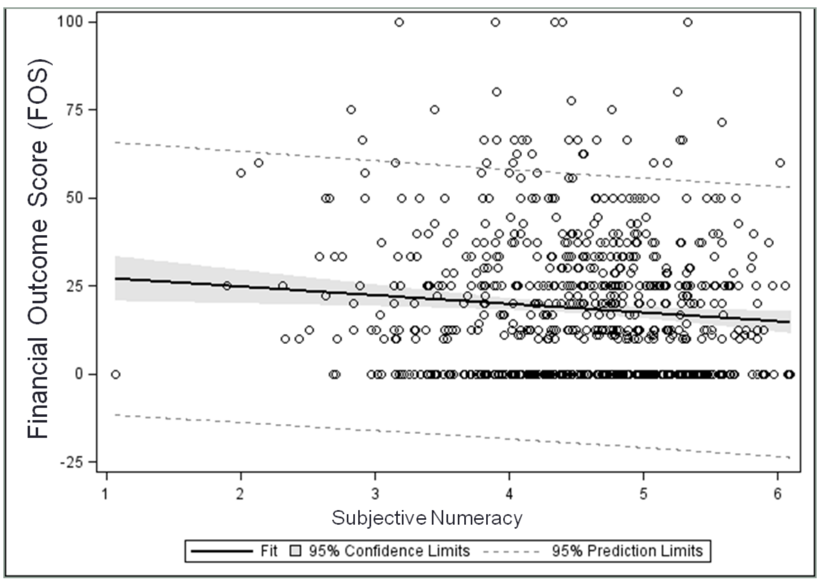

Figure 2. Session 1 subjective numeracy predicting financial outcomes.
Note. Data were jittered for visual purposes only.

Given that the subjective numeracy scale consisted of two subscales (ability and preference for numeric information), we tested if one subscale was more predictive in explaining the financial outcomes at baseline. We conducted an exploratory regression analysis using all the previously listed predictors, but separated SNS into its subscales for the model. After removing non-significant predictors (p>.05), the SNS ability subscale was the only remaining predictor (b= -1.76, t= -2.54, p= .01; see Table 4 for full and final model estimates). Lower vs. higher scores on the subjective ability (but not number preference) subscale predicted a greater proportion of negative financial outcomes experienced. Thus, subjective perceptions of ability appeared to be more important than preferences for numbers in predicting financial outcomes.

Table 4. Multiple regression results for Session 1 financial outcomes using SNS subscales

Variable	Full Model b(SE)	Final Model b(SE)
Intercept	**25.31 (8.05)**	**26.15 (3.08)**
SNS - Ability	**-1.57 (0.80)**	**-1.76 (0.69)**
SNS – Number Preference	-0.91 (0.90)	
Objective Numeracy	0.20 (0.34)	
Working memory	0.11 (0.56)	
Vocabulary	0.02 (0.13)	
Gender (1 Female, 0 Male)	0.84 (1.59)	
Age	0.03 (0.23)	
R^2	.01	.01
df	(7,740)	(1,747)
F	1.16, p=.32	6.47, p=.01

Note. Significant predictors are in bold for emphasis.

Session 2: End of Semester

To examine retention differences, we conducted a logistic regression using Session 1 FOS, ONS, SNS, Statistics condition, vocabulary, working memory, gender, age, and all two-way interactions with Statistics condition to predict participation in

Session 2. Non-significant predictors (defined as p>.05) were removed one at a time (see Table 5 for full and final model estimates).

Table 5. Logistic regression results for Session 2 retention (N=749)

Variable	Full Model *B*(SE)	Final Model *B*(SE)
Intercept	**1.35 (0.10)**	**1.38 (0.10)**
FOS S1	-0.09(0.10)	
Statistics Condition	**0.36 (0.12)**	**0.38 (0.11)**
SNS S1	-0.08 (0.12)	-0.07 (0.11)
ONS S1	**0.26 (0.11)**	**0.21 (0.10)**
Working memory S1	-0.19 (0.10)	
Vocabulary S1	**0.23 (0.10)**	**0.21 (0.09)**
Gender	**0.35 (0.11)**	**0.33 (0.09)**
Age S1	0.03 (0.12)	
FOS S1*Statistics Condition	0.05 (0.11)	
SNS S1*Statistics Condition	-0.27 (0.14)	**-0.30 (0.12)**
ONS S1*Statistics Condition	0.11 (0.13)	
Working memory S1*Statistics Condition	-0.15 (0.11)	
Vocabulary S1*Statistics Condition	0.02 (0.11)	
Gender*Statistics Condition	0.10 (0.11)	
Age S1*Statistics Condition	0.11 (0.16)	
df	(15)	(6)
X^2	41.92, p=.0002	39.19, p<.0001

Note. S1 = Session 1. All predictor variables were standardized. The *B*-values indicate the change in log odds of participating in Session 2 for each standard deviation increase in the predictor. Significant predictors (p<.05) are in bold for emphasis.

Results indicated that participants who were female (*B*= 0.33, p<.05), those who were Statistics students (at the mean level of SNS; *B*= 0.38, p<.05), those who had higher vs. lower baseline ONS scores (*B*= 0.21, p<.05; i.e., they were more objectively numerate at the beginning of the semester) and those who had higher vs. lower baseline vocabulary scores (*B*= 0.21, p<.05; i.e., they had greater verbal ability) were more likely to complete Session 2 (see Table 6 for means). There were no significant Session 2 retention differences based on Session 1 FOS, working memory, or age.

Table 6. Session 1 descriptive data by Session 2 retention status

S2 Complete?	*N*	Session 1 Variable	Mean (SE)	Std Dev	Min	Max
No	162	FOS	19.75	21.11	0	100
		ONS	10.69	2.47	2	15
		SNS	4.48	0.79	1	5.88
		Statistics Condition (1 = Statistics, 0 = REP)	0.16	0.37	0	1
		Working memory	5.62	1.41	2	9
		Vocabulary	18.03	6	1	30
		Gender (1= female, 0= males)	0.53	0.5	0	1
		Age	19.69	3.29	17	50
Yes	587	FOS	18.19	19.11	0	100
		ONS	11.1	2.28	3	16
		SNS	4.46	0.77	1.88	6
		Statistics Condition (1 = Statistics, 0 = REP)	0.34	0.47	0	1
		Working memory	5.38	1.27	2	9
		Vocabulary	19.39	5.43	0	32
		Gender (1= female, 0= males)	0.67	0.47	0	1
		Age	19.87	3.04	17	59

One potential reason for the retention differences based on higher objective numeracy and vocabulary is that these participants may be more responsible, agreeable, or conscientious relative to those participants lower in objective numeracy and vocabulary. Additionally, Statistics participants may have been more likely to complete Session 2 relative to REP participants because REP participants had the option of completing other researchers' experiments to fulfill their research experience requirements for their psychology course. For example, REP participants may have already fulfilled their research experience requirements by the end of the semester or they may not have wanted to wait until the end of the semester to complete their requirements.

The results also indicated a significant interaction between subjective numeracy and Statistics condition in predicting Session 2 participation (B= -0.30, $p<.05$). Specifically, the proportion of Statistics students who completed Session 2 was higher for those Statistics students who were lower compared to higher in subjective numeracy at the beginning of the semester. Based on median splits of baseline SNS, 91% vs. 85% of lower vs. higher SNS Statistics students, respectively, completed Session 2. Among REP participants, 72% vs. 76% of lower vs. higher SNS individuals completed Session 2. Figure 3 illustrates Session 1 SNS means by Statistics condition and Session 2 completion status. No other significant interaction effects emerged. This Session 2 retention difference based on the interaction of Statistics condition and SNS was likely due to the extra-credit incentive for participation among Statistics students. Statistics participants lower in subjective numeracy likely thought they would do poorly in the statistics class (because they had lower confidence in their ability) and they thought they would need the extra credit for their grade; thus, they were more likely to complete Session 2.

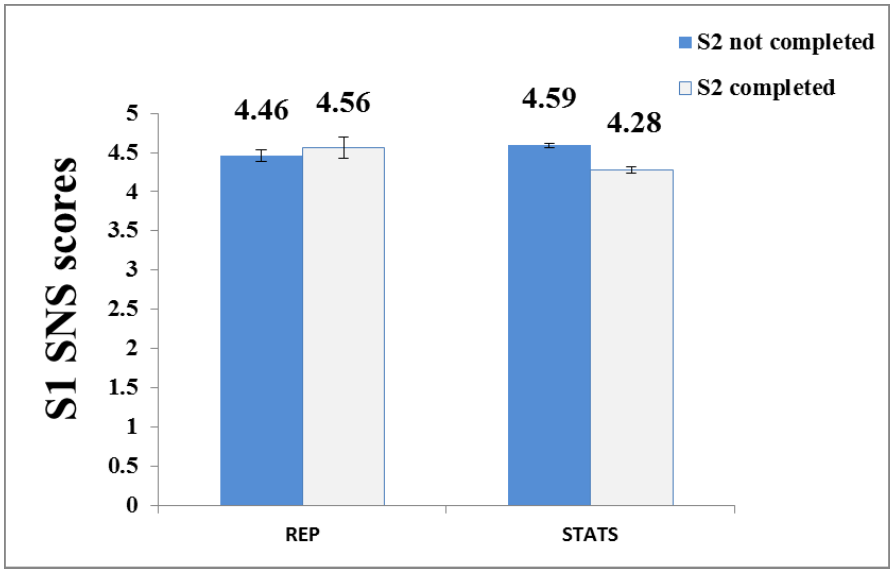

Figure 3. Session 1 SNS scores by Statistics condition and Session 2 completion status

Moving forward, we acknowledge that testing and interpreting the effects of taking a statistics course on changes in numeric competencies and financial outcomes is confounded due to Session 2 retention differences (e.g., the Statistics condition by SNS interaction), but we will provide additional analyses and discussion of this limitation in the coming sections.

To examine relations between changes in numeric competencies and outcomes over time, we computed change scores for each participant by subtracting Session 1 SNS, ONS, and FOS values from Session 2 SNS, ONS, and FOS values. A positive change score on ONS and SNS indicated improvements in the numeric competency from beginning to end of the semester, whereas negative change scores indicated decrements in the numeric competency. Because higher FOS scores indicated worse financial outcomes, a positive FOS change score indicated an increase in bad outcomes experienced and a negative FOS change score indicated a decrease in bad outcomes experienced.

Descriptive data for SNS, ONS, and FOS at Session 1 and 2 are shown in Table 7. On average, participants decreased in subjective numeracy ($M\Delta = -.12$, $t=-4.81$, $p<.0001$), decreased in objective numeracy ($M\Delta = -.29$, $t=-2.90$, $p<.01$), and increased in negative financial outcomes ($M\Delta = 1.93$, $t=2.72$, $p=.01$) from the beginning to the end of the semester (see Table 7).

Table 7. Session 1 and 2 scale descriptive data and change scores (S1ΔS2)

(*N*=587)	Mean (SE)		
Variable	Session 1	Session 2	S1ΔS2
SNS	4.46(.03)	4.35(.04)	-.12(.02)
ONS	11.10(.09)	10.81(.12)	-.29(.10)
FOS	18.19(.79)	20.12(.85)	1.93(.71)
Vocabulary	19.39(.22)	N/A	N/A
Working memory	5.38(.05)	N/A	N/A

Test of Hypothesis 2:

Taking a statistics course (vs. not) would improve objective and/or subjective numeracy from the beginning to the end of the semester.

To test Hypothesis 2, we conducted t-tests for SNS and ONS change scores by Statistics Class condition. Inconsistent with Hypothesis 2, there were no differences in ONS changes over the semester by Statistics condition ($t= -.77$, $p=.44$; see Table 8). Thus, taking a statistics course did not improve participants' objective numeracy from the beginning to the end of the semester on average overall or compared to the REP sample.

Table 8. Session 1 and 2 scale descriptive data and change scores by condition (S1ΔS2)

REP (*N*=388)	Mean (SE)		
Variable	Session 1	Session 2	S1ΔS2
SNS	4.56(.04)	4.41(.04)	-.15(.03)
ONS	11.07(.11)	10.72(.15)	-.34(.12)
FOS	16.92(.94)	18.99(1.2)	2.07(.86)
Vocabulary	19.12(.29)	N/A	N/A
Working memory	5.51(.06)	N/A	N/A
Statistics (*N*=199)	Mean (SE)		
Variable	Session 1	Session 2	S1ΔS2
SNS	4.28(.05)	4.23(.06)	-.06(.04)
ONS	11.16(.16)	10.98(.20)	-.18(.18)
FOS	20.66(1.42)	22.32(1.54)	1.66(1.25)
Vocabulary	19.91(.35)	N/A	N/A
Working memory	5.13(.09)	N/A	N/A

Results also indicated that taking a statistics course (vs. not) marginally predicted SNS changes from the beginning to the end of the semester ($t=1.84$, $p=.07$). Specifically, Statistics students had stable SNS on average ($M_{Statistics}\ \Delta = -.06$, CL -.14 to .03, Std Dev = .62), but non-Statistics students decreased in SNS ($M_{REP}\ \Delta = -.15$, CL -.21 to -.09, Std Dev = .59; see Table 8). In other words, on average, REP students went from a 4.56 to a

4.41 on the SNS scale from the beginning to the end of the semester. These results are also inconsistent with Hypothesis 2 because we did not find that taking a statistics course improved subjective numeracy scores among Statistics students.

Although we did not find improvements in objective or subjective numeracy scores from taking a Statistic course, taking a statistics course appeared to provide a protective effect to decrements in subjective numeracy from the beginning to the end of the semester. On average, REP participants decreased in subjective numeracy from the beginning to the end of the semester and Statistics participants had stable SNS.

As previously reported, Hypothesis 2 test results were confounded by retention differences based on the interaction of Statistics condition and Session 1 subjective numeracy. The proportion of Statistics students who completed Session 2 was higher for those Statistic students who were lower compared to higher in subjective numeracy at the beginning of the semester. As a result, the statistic course's protective effect to decrements in subjective numeracy could be due to differences in attrition from Session 1 to Session 2, such that the remaining Statistics students were closer to floor on the subjective numeracy scale than the non-Statistics students. However, Statistics participants who participated in Session 2 still had ample room to show declines on the subjective numeracy scale from Session 1 to Session 2 ($M_{Statistics}$ S1 SNS = 4.28 on a 6-point scale from 1 to 6). Therefore, the effect is unlikely to be due to floor effects in terms of scale usage.

Based on retention results, we know that Statistics participants who had lower Session 1 objective numeracy and higher Session 1 subjective numeracy were underrepresented in our Session 2 data. These Statistics participants were overconfident in their math abilities. It is possible that these participants would have decreased the most in subjective numeracy by the end of the semester because they would have gotten feedback on their math ability from exam scores. This is another potential explanation for why attrition may account for the self-protective effect of Statistics training on changes in subjective numeracy. We examined if lower grades on the first and second exam predicted attrition. We selected the first and second exam grades because the end of the semester follow-up was conducted prior to the third exam. If lower grades predicted attrition, then the self-protective effect of statistics training on changes in subjective numeracy may be a false effect due to attrition. Results indicated that lower exam scores did not predict attrition (see Table 9 for model estimates). These results provide further evidence that the self-protective effect of statistics training on changes in subjective numeracy was likely not due to retention differences.

Table 9. Logistic regression results for Session 2 retention among Statistic participants based on exam scores (N=225)

Variable	Full Model b(SE)
Intercept	2.86 (2.13)
Exam 1 scores	0.07 (0.11)
Exam 2 scores	-0.08 (0.13)
df	(2)
X^2	0.52, p=.77

Note. There were no significant predictors.

Test of Hypothesis 3:

Changes in objective and/or subjective numeracy would predict changes in financial outcomes from the beginning to the end of the semester

To test Hypothesis 3, we conducted multiple linear regression analysis of FOS change scores from ONS/SNS change scores from Session 1 to 2, Session 1 vocabulary, Session 1 working memory, age, and gender. Non-significant predictors (defined as $p>.05$) were removed one at a time, leaving the SNS change score as the only remaining predictor ($b=-2.76$, $t=-2.33$, $p= .02$; see Table 10 for full and final model estimates).

Table 10. Multiple regression results for change in financial outcomes from Session 1 to Session 2 (FOS S1ΔS2)

Variable	Full Model b(SE)	Final Model b(SE)
Intercept	2.67 (6.36)	**1.60 (0.72)**
SNS S1ΔS2	**-2.67 (1.19)**	**-2.76 (1.18)**
ONS S1ΔS2	-0.32 (0.30)	
Working memory S1	0.02 (0.57)	
Vocabulary S1	-0.08 (0.13)	
Gender (1 Female, 0 Male)	0.49 (1.54)	
Age	0.00 (0.23)	
R^2	.01	.01
df	(6,579)	(1, 587)
F	1.23, p=.29	5.45, p=.02

Note. Significant predictors are in bold for emphasis.

Consistent with Hypothesis 3, decreases in individuals' subjective numeracy (but not objective numeracy) from the beginning to end of the semester were associated with increases in the proportion of negative financial outcomes experienced during that time (see Figure 4). Thus, changes in subjective numeracy appeared to be more important than changes in objective numeracy in predicting financial outcome changes from the beginning to the end of the semester. These results are consistent with our results from Hypothesis 1 that lower-subjective numeracy was associated with worse financial outcomes at baseline.

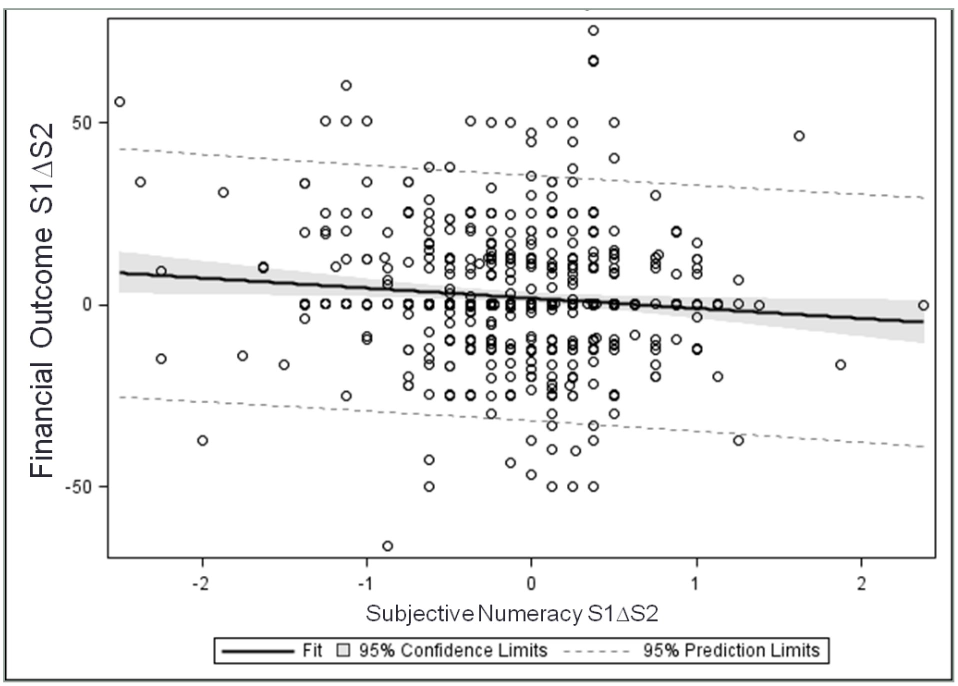

Figure 4. Subjective numeracy changes predicting financial outcome changes (S1∆S2)
Note. Data were jittered for visual purposes only.

We also tested if change on one of the SNS subscales was more important in predicting individuals' financial outcome changes over the semester. We conducted the same multiple regression analysis as described in Hypothesis 3 testing, but calculated separate change scores for the two subscale components. The subjective ability change variable was the only remaining predictor in the model (b=-2.23, t=-2.32, p=.02; see Table 11 for full and final model estimates). Results indicated that decreases in individuals' subjective ability from Session 1 to Session 2 were predictive of increases in the proportion of negative financial outcomes experienced during that time. Consistent with the Hypothesis 1 subscale results, subjective ability changes were more important than changes in preferences for numbers in predicting financial outcomes over time.

Table 11. Multiple regression results for change in financial outcomes from Session 1 to Session 2 (FOS S1ΔS2) using SNS subscales

Variable	Full Model b(SE)	Final Model b(SE)
Intercept	2.51 (6.37)	**1.54 (0.73)**
SNS S1ΔS2 – Ability	**-1.92 (1.00)**	**-2.23 (0.96)**
SNS S1ΔS2 – Number Preference	-0.81 (0.95)	
ONS S1ΔS2	-0.31 (0.30)	
Working memory S1	0.02 (0.57)	
Vocabulary S1	-0.08 (0.13)	
Gender (1 Female, 0 Male)	0.51 (1.54)	
Age	0.00 (0.23)	
R^2	.01	.01
df	(7, 578)	(1, 585)
F	1.13, p=.34	5.37, p=.02

Note. Significant predictors are in bold for emphasis.

Test of Hypothesis 4:

Taking a statistics course (vs. not) would predict improved financial outcomes from the beginning to end of the semester through changes in objective and/or subjective numeracy

We used structural equation modeling to test the indirect effect of Hypothesis 4 while controlling for vocabulary, working memory, age, and gender among the financial outcome change scores (dependent variable). Non-significant predictors (defined as $p>.10$) were removed one at a time, leaving only Statistics condition predicting SNS change scores (Hypothesis 2 result) and SNS change scores predicting FOS change scores (Hypothesis 3 result) from the beginning to end of the semester. We relaxed the criteria for removing variables in this model because of the marginal relationship between Statistics condition and changes in subjective numeracy from the beginning to the end of the semester ($p=.07$). We wanted to examine a signal of an indirect effect. The indirect effect, however was not significant ($t=-1.45$; see Appendix I for Figures of the full and final model estimates).

Inconsistent with Hypothesis 4, we did not find a statistically significant indirect effect of taking a statistics course on changes in financial outcomes through changes in one or both numeric competencies. Additionally, there was no direct effect of taking a statistics course on changes in financial outcomes from beginning to end of the semester.

Session 3: Over One Year Later

To examine retention differences, we conducted a logistic regression using Session 1 FOS, ONS, SNS, Statistics condition, vocabulary, working memory, gender, age, and all two-way interactions with Statistics condition to predict participation in the one-year follow-up. Non-significant predictors (defined as p>.05) were removed one at a time (see Table 12 for full and final model estimates).

Table 12. Logistic regression results for Session 3 retention (N=749)

Variable	Full Model *B*(SE)	Final Model *B*(SE)
Intercept	**-0.75 (0.09)**	**-0.74 (0.08)**
FOS S1	**-0.30 (0.09)**	**-0.29 (0.09)**
Statistics Condition	-0.05 (0.10)	
SNS S1	0.11 (0.09)	
ONS S1	**0.20 (0.09)**	**0.23(0.08)**
Working memory S1	0.02 (0.08)	
Vocabulary S1	0.17 (0.09)	
Gender	**0.35 (0.09)**	**0.31 (0.08)**
Age S1	-0.07 (0.10)	
FOS S1*Statistics Condition	-0.03 (0.09)	
SNS S1*Statistics Condition	-0.11 (0.10)	
ONS S1*Statistics Condition	0.19 (0.10)	
Working memory S1*Statistics Condition	0.03 (0.08)	
Vocabulary S1*Statistics Condition	0.12 (0.09)	
Gender*Statistics Condition	-0.04 (0.10)	
Age S1*Statistics Condition	-0.15 (0.11)	
df	(15)	(3)
X^2	38.11, p=.001	26.61, p<.0001

Note. S1 = Session 1. All predictor variables were standardized. The *B*-values indicate the change in log odds of participating in Session 3 for each standard deviation increase in the predictor. Significant predictors are in bold for emphasis.

Results indicated that participants who were female (*B*=.31, p<.05) and those who had lower vs. higher baseline FOS scores (*B*= -.29, p<.05; i.e., better financial outcomes) and higher vs. lower baseline ONS scores (*B*= .23, p<.05; i.e., they were more objectively numerate) were more likely to complete Session 3 (see Table 13 for means). One potential reason for the retention differences is that these participants may have been more responsible, agreeable, or conscientious relative to those who did not continue. No significant interactions existed with Statistics condition.

Table 13. Session 1 descriptive data by Session 3 retention status

S3 Complete?	*N*	Session 1 Variable	Mean (SE)	Std Dev	Min	Max
No	502	FOS	20.15 (0.91)	20.29	0.0	100.0
		ONS	10.89 (0.11)	2.43	2.0	16.0
		SNS	4.44 (0.04)	0.81	1.0	6.0
		Statistics Condition (1 = Statistics, 0 = REP)	0.30 (0.02)	0.46	0.0	1.0
		Working memory	5.42 (0.06)	1.29	2.0	9.0
		Vocabulary	18.85 (0.25)	5.53	1.0	31.0
		Gender (1= female, 0= males)	0.61 (0.02)	0.49	0.0	10
		Age	19.91(0.15)	3.30	18.0	59.0
Yes	247	FOS	15.22 (1.12)	17.55	0.0	100.0
		ONS	11.26 (0.13)	2.08	4.0	16.0
		SNS	4.53 (0.04)	0.71	2.5	6.0
		Statistics Condition (1 = Statistics, 0 = REP)	0.30 (0.03)	0.46	0.0	1.0
		Working memory	5.46 (0.08)	1.33	2.0	9.0
		Vocabulary	19.59 (0.36)	5.67	0.0	32.0
		Gender (1= female, 0= males)	0.72 (0.03)	0.45	0.0	1.0
		Age	19.67(0.17)	2.65	18	44

We considered that the financial incentive of the one-year follow-up may explain the observed retention differences based on financial outcomes, but the pattern is the opposite of what we would expect. Specifically, we thought that those with worse financial outcomes would be more likely to complete the one-year follow-up because of the financial incentive, but we found that those with better financial outcomes were more likely to complete the follow-up. Results also indicated that females were more likely to complete the one-year follow-up relative to males. There were no retention differences based on Statistics condition, subjective numeracy, working memory, vocabulary, and age. Additionally, no interactions with Statistics condition existed; thus, testing a condition effect among those people who were willing to continue was still of interest.

Descriptive data for SNS, ONS, and FOS at Session 1, Session 2, and Session 3 across all participants who completed all three sessions (N= 247) and by Statistics condition (N=174 and 73, respectively in the REP and Statistics groups) are shown in Table 14. On average, participants marginally increased in negative financial outcomes and objective numeracy from the beginning of the semester to over a year later (M_{ONS} S1ΔS3 = .26, t=1.80, p=.07; M_{FOS} S1ΔS3 = 2.12, t=1.74, p=.08; see Table 14). There were no significant changes in SNS from Session 1 to 3 on average (M_{SNS} S1ΔS3 =-.01, t=-.25, p=.80).

Table 14. Session 1, 2, and 3 subjective numeracy, objective numeracy, financial outcomes, and change scores (S1ΔS2, S1ΔS3)

All Participants (N=247)	Mean (SE)				
Variable	Session 1	Session 2	Session 3	S1ΔS2	S1ΔS3
SNS	4.53(0.04)	4.44(0.05)	4.53(0.05)	-0.10(0.04)	-0.01(0.04)
ONS	11.26(0.13)	11.09(0.17)	11.51(0.15)	-0.17(.16)	0.26(0.14)
FOS	15.22(1.12)	15.86(1.2)	17.34(1.23)	0.64(.92)	2.12(1.21)
REP (N=174)	Mean (SE)				
Variable	Session 1	Session 2	Session 3	S1ΔS2	S1ΔS3
SNS	4.60(0.05)	4.42(0.07)	4.57(0.06)	-0.18(0.04)	-0.03(0.04)
ONS	11.05(0.17)	10.83(0.22)	11.22(0.19)	-0.22(0.19)	0.17(0.18)
FOS	14.49(1.32)	16.51(1.51)	18.26(1.53)	2.03(1.1)	3.77(1.42)
Statistics (N=73)	Mean (SE)				
Variable	Session 1	Session 2	Session 3	S1ΔS2	S1ΔS3
SNS	4.38(0.09)	4.48(0.09)	4.41(0.09)	0.10(0.06)	0.03(0.08)
ONS	11.75(0.19)	11.73(0.26)	12.22(0.22)	-0.03(0.28)	0.47(0.20)
FOS	16.97(2.08)	14.31(1.89)	15.13(2.0)	-2.66(1.61)	-1.83(2.29)

Test of Hypothesis 5:

Taking a statistics course (vs. not) would be associated with better financial outcomes through changes in objective numeracy or subjective numeracy over a year later.

We used structural equation modeling to test the indirect effect of Hypothesis 5 while controlling for vocabulary, working memory, age, and gender among the financial outcome change scores (dependent variable). Non-significant predictors (defined as $p>.05$) were removed one at a time, leaving only Statistics condition predicting FOS change scores from the beginning of the semester to over one year later (*B*(SE)= -.13

(.06); see Appendix J for Figures of the full and final model estimates). Inconsistent with Hypothesis 5, we did not find an indirect effect of taking a statistics course on changes in financial outcomes through changes in objective or subjective numeracy from the beginning of the semester to over a year later. There were no significant differences in ONS changes (t= -1.10) or SNS changes (t= -.65) by Statistics condition from Session 1 to Session 3. Additionally, we did not find a relationship between SNS changes and FOS changes in the model.

On average, statistics students remained stable (with a slight improvement) in negative financial outcomes from the beginning of the semester to over a year later (M_{FOS} Δ(SE) = -1.83(2.29)), but REP students increased in negative financial outcomes over that time (M_{FOS} Δ(SE) = 3.77(1.42); see Figure 5). Among those people willing to complete Session 3, taking a statistics course provided a protective effect to detrimental changes in financial outcomes from the beginning of the semester to over one year later. These results are somewhat consistent with Hypothesis 5. We did not find that taking a statistics course was predictive of significant improvements in financial outcomes, but Statistics participants (on average) did not experience the increase in negative financial outcomes that REP participants experienced. Thus, taking a statistics course appeared to have more long-term benefits as it provided a protective effect to detrimental changes in financial outcomes for those college students.

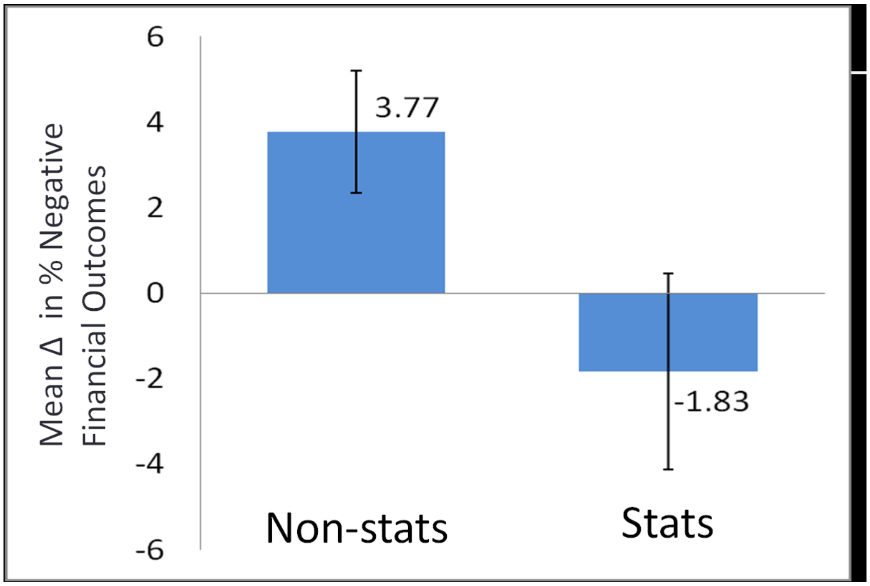

Figure 5. Mean financial outcomes change from Session 1 to 3 by Statistics condition

Note. Error bars indicate standard error of the means.

General Discussion

The present research had several aims. The first aim was to examine the possible dissociable roles of objective and subjective numeracy in predicting the experience of negative financial outcomes. Consistent with Hypothesis 1, we found that participants with lower subjective numeracy reported experiencing worse financial outcomes (controlling for intelligence proxies) at baseline. Additionally, we found that participants' subjective ability (the SNS-ability subscale) was more predictive than preferences (SNS-preference subscale) for numbers in predicting financial outcomes at baseline. This result suggests that greater confidence or beliefs in one's abilities is more important than preferences for numbers to experiencing better financial outcomes. Greater confidence or beliefs in one's abilities may result in better financial outcomes by driving individuals to play a more active role in financial decisions and behaviors, although future research is needed to determine the behavioral consequences of greater subjective numeric ability.

Surprisingly, we did not find that objective numeracy was predictive of financial outcomes at baseline (inconsistent with Hypothesis 1). In fact, we did not find objective numeracy to be associated with our financial outcome measure based on a simple correlation. This result may be due to the limited number and types of financial items in our measure. For example, objective numeracy may play a larger role in financial domains not captured by our items (e.g., mortgages, Gerardi, Goette, & Meier, 2013; investments, Banks & Oldfield, 2007; payday loans, Sinayev & Peters, 2015; child support, etc.). These types of items were not included because they were less relevant to a student population.

We also attempted to improve individuals' numeric competencies and financial outcomes through statistics training. On average, across the full sample, we found that participants decreased in SNS and ONS from the beginning to the end of the semester. It is unclear what may have caused the negative changes in subjective and objective numeracy over the semester across participants, but one potential contributor may be that some students were tired and fatigued by the end of the semester. Inconsistent with Hypothesis 2, we did not find that taking a statistics course improved participants' objective numeracy from the beginning to the end of the semester compared to those who did not take the course. A potential explanation for this failure to improve objective numeracy is that one undergraduate statistics course may not have been sufficient training, or training of the right type, to improve individuals' performance on the objective numeracy scale in particular. The course material may not have matched the skills needed to answer the objective numeracy items. Additionally, learning probabilities is a component of the undergraduate statistics course, but it is not the only focus (e.g., learning how to display data using computer software to create graphs, box plots, stem-and-leaf displays, etc.). Further courses or training in statistics and probabilities may be necessary to improve individuals' objective numeracy.

Alternatively, more basic training in the mental manipulation of numerical quantities without the use of numerical symbols may be useful for improving the understanding of probabilities and numerical concepts. Recent research found that repeated training on a non-symbolic approximate arithmetic task improved symbolic arithmetic performance (Park & Brannon, 2013). In these studies, the researchers trained participants on approximate addition and subtraction with arrays of dots and found that this training improved performance on symbolic math problems. A similar non-symbolic training of proportions with arrays of dots may help increase understanding and use of proportions in math problems.

Also inconsistent with Hypothesis 2, we did not find that taking a statistics course (vs. not) *improved* subjective numeracy over the semester. Nonetheless, we did find several benefits of statistics training. Taking a statistics course (vs. not) appeared to provide a protective effect to detrimental changes in subjective numeracy over the semester. On average, REP participants decreased in subjective numeracy from the beginning to the end of the semester and Statistics participants had stable SNS. We are unable to identify the specific factors that led to the decline in subjective numeracy for

REP participants, but we would expect it to be due to factors related to numeric vs. non-numeric domains, such as working on difficult tasks involving numbers.

Similar to Hypothesis 1 result, we found that changes in subjective numeracy from the beginning to the end of the semester were predictive of changes in financial outcomes over that time. More specifically, we found that declines in subjective numeracy were associated with increases in the proportion of negative financial outcomes experienced over a semester. Changes in objective numeracy were not predictive of changes in financial outcomes. In supplemental analyses, we found that changes in subjective numeric ability over the semester were more important than changes in preferences for numbers in these predictions. This result was consistent with the exploratory analysis results from Hypothesis 1. These results again support the idea that perceptions of one's ability were more important than a preference for numbers in predicting financial outcomes.

Inconsistent with Hypothesis 4, we did not find a significant indirect effect of taking a statistics course (vs. not) on changes in financial outcomes through changes in one or both numeric competencies over the course of a semester. Taking a statistics course (vs. not) was marginally associated with changes in subjective numeracy ($p=.07$), and changes in subjective numeracy were associated with changes in financial outcomes, but the indirect effect did not reach statistical significance. However, the Statistics condition's protective effect on negative changes in subjective numeracy was small. The average change on the SNS scale (range = 1 to 6) was -.15 for REP participants. We would likely need a larger sample size to obtain a statistically significant indirect effect, and given the effect size, this may not be an important enough determinant of financial outcomes to pursue. Additionally, there was no direct effect of taking a statistics course on changes in financial outcomes from beginning to end of the semester. The time frame of one semester may have been too small to see an impact of taking a statistics course on financial outcomes.

The final aim of the present research was to examine if taking a statistics course would be associated with better financial outcomes through changes in ONS or SNS over a year later, controlling for intelligence proxies and demographics (Hypothesis 5). Inconsistent with Hypothesis 5, we did not find an indirect effect of taking a statistics course on changes in financial outcomes through changes in objective or subjective numeracy from the beginning of the semester to over a year later. Among those willing to participate, there were no significant differences in ONS or SNS changes by Statistics condition from the beginning of the semester to over one year later. Perhaps subjective numeracy may be more stable and less susceptible to change over longer periods of time. Additionally, we did not find a relationship between SNS changes and FOS changes in the model. However, there may not have been enough variance in SNS changes from the beginning of the semester to one year later to detect a relationship between SNS changes and FOS changes over that time. On average, there was no change in subjective numeracy across participants from the beginning of the semester to over a year later (M_{SNS} S1ΔS3 = -.009, Std Dev= .59, p=.80). It is possible that we did not have enough variance in SNS changes or a large enough sample to detect a relationship between SNS changes and FOS changes over this longer period of time.

However, we did find that taking a statistics course appeared to provide a protective effect to detrimental changes in financial outcomes from the beginning of the semester to over one year later. On average, statistics students remained stable in negative financial outcomes from the beginning of the semester to over a year later, but REP students increased in negative financial outcomes over that time. These results are somewhat consistent with Hypothesis 5. More specifically, we did not find that taking a statistics course was predictive of significant improvements of financial outcomes, but statistics course participants (on average) did not experience the increase in negative financial outcomes that REP participants experienced. One potential reason why we observed an increase in negative financial outcomes is that our participants may have been taking on more responsibilities in their life as time passed (e.g., more difficult classes, more leadership in extra-curricular activities) and REP participants, in particular, did not keep up with their financial responsibilities. Thus, taking a statistics course appeared to have more long-term benefits by providing a protective effect to detrimental changes in financial outcomes for those college students. The timeframe of a semester may have been too short to show an impact of statistics training of financial outcomes within the term, but the statistics training may have affected decision-making by the end of the course that had an impact on financial outcomes down the road.

Across time frames (both over a semester and over a year long-period) we tended to see participants (by Statistics condition or across all participants) get worse or remain stable in subjective numeracy and financial outcomes. We did not, however, find any average improvements across all participants (or by statistics condition) on these measures. One potential explanation for this result may be due to floor and ceiling effects. Participants were already scoring high on the subjective numeracy measure (M= 4.5 at baseline on a 6 point scale) and low on the financial outcome measure (M= 18% at baseline on a 100% scale, but measured with 10 binary items). There was less opportunity to detect improvements on subjective numeracy or financial outcomes with these measures when, at baseline, participants scored near the top of the subjective numeracy scale and they did not experience many negative financial outcomes. There was much more opportunity to detect decrements in these measures. In future studies, adjustments should be made to the items to allow greater opportunity to detect improvements in these measures (e.g., increase the SNS scale range, add more financial items to FOS).

Limitations

There are several limitations to our findings. As previously mentioned, there were a limited number of financial items (some of which were knowledge based) that were tailored to undergraduate students and were self-reported outcomes (vs. actual outcomes). Thus, our findings may not generalize to adults who have more (and different) financial responsibilities. Additionally, subjective numeracy explained only a small amount of variance in the financial outcomes (e.g., most of the models' $R^2 = .01$). The effect size is small, but it is not trivial (Cohen, 1992). In particular, for a person scoring lowest on SNS (score of 1), the predicted percentage of negative financial outcomes experienced was 27%, compared to only 15% for a person scoring highest on SNS (score of 6). This difference could have a big impact considering the variety and severity of financial outcomes a person can experience in life.

An additional limitation to the present findings is that we do not know how our results vary by income. We did not have data on income, but we suspect that our college student sample was less variable relative to the general population and more dependent (but variable in that dependence) on parental income. Lastly, but perhaps most noteworthy, we cannot make any causal claims with this research because it was a quasi-experimental design with non-random group assignment. We had baseline differences in subjective numeracy and financial outcomes between our statistics training and control group. We controlled for these baseline differences in our analyses, but the groups may have differed on other important dimensions that were not measured in this study and, thus, we were unable to account for these differences. In addition, we had overall sample differences based on retention from Session 1 to Session 2. There were significant retention differences from Session 1 to Session 2 related to an interaction of baseline subjective numeracy and Statistics condition. Therefore, we are unable to definitively rule out the possibility that attrition may explain the self-protective effect of Statistics condition on detrimental changes in subjective numeracy. However, additional analyses did not support this attrition explanation. We also had overall sample differences based on retention from Session 1 to Session 3. Retention differences from Session 1 to Session 3, however, were unrelated to our Statistics condition, allowing us to continue our tests. Retention differences do limit the generalizability of our results to a student population.

Despite these limitations, the present research contributes to the numeracy literature by providing additional evidence that objective and subjective numeracy are related, but separable constructs with dissociable roles.

Future Directions

Future research is needed to determine the key underlying psychological factors driving these effects and the behavioral consequences of greater subjective numeracy. In addition, very little is known about the malleability of subjective numeracy. For example, how susceptible is SNS to change over short and long periods of time? How long do these changes last?

One approach to changing subjective numeracy is to change individuals' perceptions of their numeric ability by giving them a difficult or easy math task. Additionally, we could further manipulate individuals' perceptions of their abilities by providing them with fake performance scores of the average student for the task (e.g., high average scores in a difficult test condition and low average scores in a easy test condition). These manipulations should alter individuals' subjective ability through the experienced difficulty (or ease) of performing the math task and through social comparison.

Another approach to changing individuals' subjective numeracy is to change their liking of numeric information. Intrinsic motivation research has shown that when people freely alternate between two disliked activities, the decision to work on one of the activities is encoded as a decision not to avoid the task (Fujita & MacGregor, 2011; Higgins, Trope, & Kwon, 1999). Switching between the two disliked tasks thus leads individuals to conclude that the activities are not so bad. Allowing participants to freely choose between two numeric tasks may increase preferences for numeric activities or information among low-SNS individuals. However, in theory, this manipulation would have the opposite effect on high-SNS individuals. Deciding not to approach a likeable numeric activity with each task switch should lead individuals to conclude that the numeric activities are not so enjoyable.

The studies would require pre-and post-test assessments of subjective and objective numeracy, as well as the inclusion of non-numeric comparison tasks. Additional assessments of subjective numeracy over variable time-periods would assess how long these changes, if any, in subjective numeracy last.

Conclusions

The present research highlights the independent importance of subjective numeracy in financial outcomes, a variable that has not been considered in the financial literature. These findings should be taken into consideration when developing strategies to improve financial outcomes. Governments, business, and organizations worldwide have spent billions of dollars creating educational programs to try and improve financial behaviors, but these programs have shown almost no benefit (Fernandes, Lynch, & Netemeyer, 2014). Educational programs that also include interventions for improving subjective numeric ability may be more effective.

References

Anderson, B. L., Obrecht, N. A., Chapman, G. B., Driscoll, D. A., & Schulkin, J. (2011). Physicians' communication of Down syndrome screening test results: The influence of physician numeracy. *Genetics in Medicine*, *13,* 744-749.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191-215. doi:10.1037/0033-295X.84.2.191

Banks, J., O'Dea, C., & Oldfield, Z. (2010). Cognitive function, numeracy and retirement saving trajectories. *The Economic Journal*, *120*, 381-410.

Banks, J., & Oldfield, Z. (2007). Understanding pensions: Cognitive function, numerical ability, and retirement saving. *Fiscal Studies*, *28*, 143-170.

Betz, N. (2013). Assessment of Self-efficacy. In B. A. Bracken, J. F. Carlson, J. C. Hansen, N. R. Kuncel, S. P. Reise, & M. C. Rodriguez (Eds.), *APA Handbook of Testing and Assessment in Psychology* (pp.379-391). Washington, DC: APA.

Bruine de Bruin, W., Parker, A. M., & Fischhoff, B. (2007). Individual differences in adult decision-making competence. *Journal of Personality and Social Psychology*, *92*, 938-956.

Cohen, J. (1992). A power primer: Quantitative methods in psychology. *Psychological Bulletin*, *112*, 155-159.

Couper, M. P., & Singer, E. (2009). The role of numeracy in informed consent for surveys. *Journal of Empirical Research on Human Research Ethics*, *4*, 17-26.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297- 334.

Dolan, J., Cherkasky, O. A., Li, Q., Chin, N., & Veazie, P. J. (2014). Should health numeracy be assessed objectively or subjectively? *Medical Decision Making*. Advance online publication. doi: 10.1177/0272989X15584332

Eisinga, R., Grotenhuis, M. T., & Pelzer, B. (2013). The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? *International Journal of Public Health, 58*, 637-642.

Ekstrom, R. B., French, J. W., Harman, H., & Derman, D. (1976). *Kit of factor-referenced cognitive tests* (rev. ed.). Princeton, NJ: Educational Testing Service.

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*, 672-680.

Fernandes, D., Lynch, J. G., & Netemeyer, R. G. (2014). Financial literacy, financial education, and downstream financial behaviors. *Management Science*, *60*, 1861-1883.

Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology*, *18*, 253-292.

Fujita, K., & MacGregor, K. E. (2012). Basic goal distinctions. In H. Aarts & A. J. Elliot (Eds.), *Goal-oriented behavior* (pp. 85-114). New York: Psychology Press/Taylor & Francis.

Gerardi, K., Goette, L., & Meier, S. (2013). Numerical ability predicts mortgage default.

Proceedings of the National Academy of Sciences, *110*, 11267-11271.

Gibbs, G. (1981). *Teaching students to learn.* Milton Keynes, England: Open University Press.

Glasgow, R. E., Strycker, L. A., King, D. K., & Toobert, D. J. (2014). Understanding who benefits at each step in an internet-based diabetes self-management program application of a recursive partitioning approach. *Medical Decision Making*, *34*, 180-191.

Hadar, L., Sood, S., & Fox, C. R. (2013). Subjective knowledge in consumer financial decision. *Journal of Marketing Research*, *50*, 303-316.

Hanoch, Y., Miron-Shatz, T., & Himmelstein, M. (2010). Genetic testing and risk interpretation: How do women understand lifetime risk results? *Judgment and Decision Making*, *5*, 116-123.

Hess, R., Visschers, V. H. M., & Siegrist, M. (2011). Risk communication with pictographs: The role of numeracy and graph processing. *Judgment and Decision Making*, *6*, 263-274.

Higgins, E. T., Trope, Y., & Kwon, J. (1999). Augmentation and undermining from combining activities: The role of choice in activity engagement theory. *Journal of Experimental Social Psychology*, *35*, 285–307.

Hong, T. (2006). The Internet and tobacco cessation: The roles of Internet self-efficacy and search task on the information-seeking process. *Journal of Computer-Mediated Communication*, *11*, 536-556.

Keller, C., Siegrist, M., & Visschers, V. (2009). Effect of risk ladder format on risk perception in high- and low-numerate individuals. *Risk Analysis*, *29*, 1255-1264.

Liberali, J. M., Reyna, V. F., Furlan, S., Stein, L. M., & Pardo, S. T. (2012). Individual differences in numeracy and cognitive reflection, with implications for biases and fallacies in probability judgment. *Journal of Behavioral Decision Making*, *25*, 361-381.

Lipkus, I. M., Peters, E., Kimmick, G., Liotcheva, V., & Marcom, P. (2010). Breast cancer patients' treatment expectations after exposure to the decision aid program Adjuvant Online: The influence of numeracy. *Medical Decision Making*, *30*, 464-473.

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

Lusardi, A. (2012). Numeracy, financial literacy, and financial decision-making. *Numeracy*, 5. Retrieved from http://dx.doi.org/10.5038/1936-4660.5.1.2

Madrian, B. C., & Shea, D. F. (2001). The power of suggestion: Inertia in 401(k) participation and savings behavior. *The Quarterly Journal of Economic*, *116*, 1149-1187.

Main, A. (1980). *Encouraging effective learning.* Edinburgh, Scotland: Scottish Academic Press.

Miron-Shatz, T., Hanoch, Y., Doniger, G. M., Omer, Z. B., & Ozanne, E. M. (2014). Subjective but not objective numeracy influences willingness to pay for BRCA1/2 genetic testing. *Judgment and Decision Making*, *9*, 152-158.

Nelson, W. L., Moser, R. P., & Han, P. J. (2013). Exploring objective and subjective numeracy at a population level: Findings from the 2007 Health Information National Trends Survey (HINTS). *Journal of Health Communication*, *18*, 192-205.

OECD (2013). *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills*. Paris: OECD Publishing. Retrieved from http://dx.doi.org/10.1787/9789264204256-en

Park, J., & Brannon, E. M. (2013). Training the approximate number system improves math proficiency. *Psychological Science*, *24*, 2013-2019.

Peters, E. (2012). Beyond comprehension: The role of numeracy in judgments and decisions. *Current Directions in Psychological Science*, *21*, 31-35.

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

Peters, E., Dieckmann, N. F., Västfjäll, D., Mertz, C. K., Slovic, P., & Hibbard, J. H. (2009). Bringing meaning to numbers: The impact of evaluative categories on decisions. *Journal of Experimental Psychology: Applied*, *15*, 213–227.

Peters, E., Tompkins, M. K., Meilleur, L., & Bjalkebring, P. (in review). Numeric competencies in health: Integrative review and future research opportunities.

Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., Dickert, S. (2006). Numeracy and decision making. *Psychological Science*, *17*, 407-413.

Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review*, *96*, 341-357.

Schley, D. R., & Fujita, K. (2014). Seeing the math in the story on how abstraction promotes performance on mathematical word problems. *Social Psychological and Personality Science*, *5*, 953-961.

Sinayev, A., & Peters, E. (2015). Cognitive reflection vs. calculation in decision making. *Frontiers in Psychology*, *6*, 1-16.

Smith, J. P., McArdle, J. J., & Willis, R. (2010). Financial decision making and cognition in a family context. *The Economic Journal*, *120*, 363-380.

Soll, J. B., Keeney, R. L., & Larrick, R. P. (2013). Consumer misunderstanding of credit card use, payments, and debt: Causes and solutions. *Journal of Public Policy & Marketing*, *32*, 66-81.

Tait, A. R., Voepel-Lewis, T., Zikmund-Fisher, B. J., & Fagerlin, A. (2010). The effect of format on parents' understanding of the risks and benefits of clinical research: a comparison between text, tables, and graphics. *Journal of Health Communication*, *15*, 487-501.

U.S. Courts. (2014). Bankruptcy Filings: Table F. Statistics & Reports. Retrieved from http://www.uscourts.gov/statistics/table/f/bankruptcy-filings/2014/12/31

U.S. Federal Reserve. (2015). Consumer Credit: G.19. Economic Research & Data. Retrieved from http://www.federalreserve.gov/releases/g19/current/default.htm

Västfjäll, D., Peters, E., & Starmer, C. (in preparation). Numeracy, incidental affect and the construction of prices.

Wechsler, D. (1997). *Wechsler adult intelligence scale* (3rd ed.). New York: Psychological Corporation.

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*, 198-212.

Zikmund-Fisher, B., Fagerlin, A., Roberts, T. R., Derry, H. A., Ubel, P. A. (2008). Alternate methods of framing information about medication side effects: Incremental risk versus total risk of occurrence. *Journal of Health Communication*, *13*, 107-124.

Zikmund-Fisher, B., Smith, D., Ubel, P., & Fagerlin, A. A. (2007). Validation of the subjective numeracy scale: Effects of low numeracy on comprehension of risk communication and utility elicitations. *Medical Decision Making, 27*, 663-671.

Appendix A: Subjective Numeracy Items

1. How good are you at working with fractions?

Not at all good Extremely good

1 2 3 4 5 6

2. How good are you at working with percentages?

Not at all good Extremely good

1 2 3 4 5 6

3. How good are you at calculating a 15% tip?

Not at all good Extremely good

1 2 3 4 5 6

4. How good are you at figuring out how much a shirt will cost if it's 25% off?

Not at all good Extremely good

1 2 3 4 5 6

5. When reading the newspaper, how helpful do you find tables and graphs that are parts of a story?

Not at all helpful Extremely helpful

1 2 3 4 5 6

6. When people tell you the chance of something happening, do you prefer that they use **words** ("it rarely happens") or **numbers** ("there's a 1% chance")?

Always prefer words Always prefer numbers

1 2 3 4 5 6

7. When you hear a weather forecast, do you prefer predictions using **percentages** (e.g., "there will be a 20% chance of rain today") or predictions using only **words** (e.g., "there is a small chance of rain today")?

Always prefer words Always prefer percentages

1 2 3 4 5 6

8. How **often** do you find numerical information to be useful?

Never Very often

1 2 3 4 5 6

Appendix B: Objective Numeracy Items

1. Imagine that we roll a fair, six-sided die 1,000 times. Out of 1,000 rolls, how many times do you think the die would come up as an even number?

Answer: Half the time, 50%, any number between 490-510, 1:2

2. In the BIG BUCKS LOTTERY, the chances of winning a $10.00 prize are 1%. What is your best guess about how many people would win a $10.00 prize if 1,000 people each buy a single ticket from BIG BUCKS?

Answer: ___10__people

3. In the ACME PUBLISHING SWEEPSTAKES, the chance of winning a car is 1 in 1,000. What percent of tickets of ACME PUBLISHING SWEEPSTAKES win a car?

Answer: ___.1__%

4. Which of the following numbers represents the biggest risk of getting a disease?

Answer: ___ 1 in 100 ___ 1 in 1000 _X_ 1 in 10

5. Which of the following numbers represents the biggest risk of getting a disease?

Answer: ___ 1% _X_ 10% ___5%

6. If Person A's risk of getting a disease is 1% in ten years, and Person B's risk is double that of A's, what is B's risk?

Answer: _2_% in_10_ years or 1% in 5 years

7. If the chance of getting a disease is 10%, how many people would be expected to get the disease:

A: Out of 100
Answer: ____10__people
B: Out of 1000
Answer: __100___people

8. If Person A's chance of getting a disease is 1 in 100 in ten years, and person B's risk is double that of A, what is B's risk?

Answer: 2 in 100 in 10 years or 1 in 50 in 10 years or 1 in 100 in 5 years.

9. If the chance of getting a disease is 20 out of 100, this would be the same as having a __20__% chance of getting the disease.

10.The chance of getting a viral infection is .0005. Out of 10,000 people, about how many of them are expected to get infected?

Answer: __5__people

11. Which of the following numbers represents the biggest risk of getting a disease?
__X__ 1 chance in 12
_____ 1 chance in 37

12. Suppose your friend just had a mammogram. The doctor knows from previous studies that, of 100 women like her, 10 have tumors and 90 do not. Of the 10 who do have tumors, the mammogram correctly finds 9 with tumors and incorrectly says that 1 does not have a tumor. Of the 90 women without tumors, the mammogram correctly finds 81 without tumors and incorrectly says that 9 have tumors. The table below summarizes this information. Imagine that your friend tests positive (as if she had a tumor), what is the likelihood that she actually has a tumor?

	Tested Positive	Tested Negative	Totals
Actually has a tumor	9	1	10
Does not have a tumor	9	81	90
Totals	18	82	100

Answer: __9___ out of __18___

13. Imagine that you are taking a class and your chances of being asked a question in class are 1% during the first week of class and double each week thereafter (i.e., you would have a 2% chance in Week 2, a 4% chance in Week 3, an 8% chance in Week 4). What is the probability that you will be asked a question in class during Week 7?

Answer: __64__%

14. Suppose that 1 out of every 10,000 doctors in a certain region is infected with the SARS virus; in the same region, 20 out of every 100 people in a particular at-risk population also are infected with the virus. A test for the virus gives a positive result in 99% of those who are infected and in 1% of those who are not infected. A randomly selected doctor and a randomly selected person in the at-risk population in this region both test positive for the disease. Who is more likely to actually have the disease?

_____ They both tested positive for SARS and therefore are equally likely to have the disease
_____ They both tested positive for SARS, and the doctor is more likely to have the disease
__X__ They both tested positive for SARS and the person in the at-risk population is more likely to have the disease.

15. A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost?

Answer: __5___cents

16. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?

Answer: __5__minutes

17. In a lake, there is a patch of lilypads. 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?

Answer: __47___days

Appendix C: Financial Outcome Items

Financial Items	% who made the decision	% experienced bad outcome given decision
Do you have a savings account or emergency fund? (no=1)	applies to all	13.6 (101/745)
• Do you know how much money you have in your savings or emergency fund? (no=1)	86.4 (644/745)	10.9 (70/644)
Do you have a credit card? (neutral)	55.2 (411/745)	
• Do you know your credit card balance? (no=1)		22.1 (91/411)
• Do you know your credit card limit? (no=1)		17.3 (71/410)
• Do you know your credit card interest rate? (no=1)		58.1 (238/410)
• **In the past year**, have you had more than $5,000 in credit card debt? (yes=1)		4.9 (20/411)
In the past year, have you had an overdraft or bounced a check from your account? (yes=1)	applies to all	14.4 (106/736)
In the past year, have you used up your funds (e.g., checking, food points) early (for example, before your next paycheck or regular allowance from family)?(yes=1)	applies to all	23.2 (172/741)
In the past year, have you been responsible for paying utilities (e.g., electricity, gas, cable, or water)? (neutral)	38.9 (291/748)	
• **In the past year**, have you ever had a utility (e.g., electricity, gas, cable, or water) shut off due to late or no payment? (yes=1)		8.3 (24/290)
Have you ever been responsible for rent or mortgage payments? (neutral)	36.6 (273/745)	
• Have you ever been late on a rent or mortgage payment?(yes=1)		17.2 (47/273)

Note. Item frequencies presented are across all Session 1 participants (*N*=759). Questions labeled as neutral did not contribute to the Financial Outcomes Score.

Appendix D: Working Memory Items

AB
CD
ZH
MP
DN
JRQ
FCS
TGK
IXVT
LBEI
HNYD
GMAZT
UFKEV
TQSOJ
UFKEV
BRCYPZ
XIDHWA
KFMOQU
XIDHWA
KFMOQU
VSGNTEJ
AOMRLKB
WFYCEHD
AOMRLKB
WFYCEHD
ZURINPQX
TFGCKAMD
OELNWBHS
TFGCKAMD
OELNWBHS
HTIDSRABK
MXFOPNJLD
AGCEUWTQV
MXFOPNJLD
AGCEUWTQV

Appendix E: Vocabulary Items

This is a test of your knowledge of word meanings. Look at the sample below. One of the five numbered words has the same meaning or nearly the same meaning as the word above the numbered words. Mark your answer by putting an X through the number in front of the word you select.

jovial
1 - refreshing
2 - scare
3 - thickset
4 - wise
~~5 -~~ jolly

The answer to the sample item is number 5; therefore, an X has been put through number 5.

Your score will be the number of items you answer correctly minus a fraction of the number marked incorrectly. Therefore, it will not be to your advantage to guess unless you are able to eliminate one or more of the answer choices as wrong.

There are two parts to this test. Each part has one page. When you have finished with the first page please continue to the second page of the test.

Vocabulary Test
Please do not spend more than a few minutes on this page.
If you are unsure of the answer, remember that an incorrect answer will deduct more points than a question left blank.

1. handicraft
   1 - cunning
   2 - fast boat
   3 - utility
   4 - manual skill
   5 - guild

2. resistant
   1 - confusing
   2 - conjunctive
   3 - systematic
   4 - assisting
   5 - opposing

3. ejection
   1 - restoration
   2 - expulsion
   3 - reformation
   4 - bisection
   5 - exposition

4. yawl
   1 - tropical storm
   2 - foghorn
   3 - carouse
   4 - sailboat
   5 - turn

5. listless
   1 - aggressive
   2 - adaptable
   3 - indifferent
   4 - sorrowful
   5 - ugly

7. unobservant
   1 - analytic
   2 - conclusive
   3 - heedless
   4 - uninformed
   5 - timid

8. perambulator
   1 - coffeepot
   2 - drunkard
   3 - baby carriage
   4 - liar
   5 - camel

9. masticate
   1 - chew
   2 - massage
   3 - manufacture
   4 - create
   5 - pollute

10. poignancy
   1 - peignoir
   2 - gloominess
   3 - keenness
   4 - gluttony
   5 - barony

11. salaam
   1 - salivation
   2 - salmon
   3 - salutation
   4 - ransom
   5 - brigand

13. inclement
   1 - balmy
   2 - happy
   3 - righteous
   4 - severe
   5 - apprehensive

14. access
   1 - abundance
   2 - evaluation
   3 - approach
   4 - extremes
   5 - foes

15. bland
   1 - disagreeable
   2 - pale
   3 - soothing
   4 - empty
   5 - musical

16. collusion
   1 - nerve
   2 - rest
   3 - prayer
   4 - conspiracy
   5 - disguise

17. degrade
   1 - lower in rank
   2 - bend downward
   3 - disagree
   4 - sort
   5 - uplift

6. acceptable
1 - affected
2 - suitable
3 - attractive
4 - genial
5 - noteworthy

12. compatible
1 - abridged
2 - congenial
3 - compelling
4 - related
5 - combined

18. evolve
1 - develop gradually
2 - spin
3 - end suddenly
4 - implicate
5 - include

19. dreg
1 - pulled
2 - worthless leftover
3 - wooden pin
4 - wheel spoke
5 - liquid

20. crescendo
1 - repeat
2 - treble clef
3 - decrease in time
4 - eighth note
5 - increase in loudness

21. trilogy
1 - set of four
2 - a pair
3 - vibrations
4 - interjections
5 - set of three

22. budget
1 - civil government
2 - capital punishment
3 - calendar
4 - bulletin
5 - financial plan

23. gritty
1 - frigid
2 - windy
3 - adhesive

25. morbid
1 - moral
2 - attractive
3 - gruesome
4 - caustic
5 - mysterious

26. malignant
1 - deliberate
2 - superior
3 - delirious
4 - malicious
5 - fragrant

27. hauteur
1 - discordancy
2 - arrogance
3 - languor
4 - ignorance
5 - utility

28. nihilism
1 - psychology
2 - optimism
3 - anarchism
4 - biology
5 - chauvinism

29. insipid
1 - benign
2 - changeable
3 - poisonous

31. complacent
1 - Friendly
2 - Smug
3 - Jealous
4 - Angry
5 - Uncivil

32. archaeology
1 - obsolete language
2 - study of ancient cultures
3 - architectural structure
4 - Lineage
5 - study of rock formations

33. canvass
1 - Crack
2 - Flower
3 - Elect
4 - white bird
5 - Examine

34. correlate
1 - ceremony of crowning
2 - relate closely
3 - distant relative
4 - overweight
5 - group of soldiers

Edifice
1 - small insect
2 - Heir
3 - Front

4 - granular
5 - unwieldy

24. alignment
1 - formation
2 - accusation
3 - emblem
4 - brightness
5 - buoyant

4 - colorless
5 - tasteless

30. droll
1 - serious
2 - argument
3 - dwarf
4 - brogue
5 - laughable

4 - large building
5 - Learning

36. flabby
1 - lacking firmness
2 - Giddy
3 - Talkative
4 - noisy and boastful
5 - Affluent

Appendix F: Correlations of Individual Difference Scales in Session 1

(N=749)	1	2	3	4	5
1. SNS S1	1.00				
2. ONS S1	0.38 <.0001	1.00			
3. Vocabulary S1	0.07 0.04	0.17 <.0001	1.00		
4. Working Memory S1	0.08 0.03	0.12 0.00	0.05 0.19	1.00	
5. FOS S1	-0.10 0.01	-0.02 0.58	0.00 0.99	0.00 0.97	1.00

Note. The first data value indicates the correlation and the second data value indicates the p-value in this matrix.

Appendix G: Correlations of Individual Difference Scales in Session 2

(N=587)	1	2	3	4	5	6	7	8
1. SNS S1	1.00							
2. ONS S1	0.39 <.0001	1.00						
3. Vocabulary S1	0.06 0.16	0.14 0.00	1.00					
4. Working Memory S1	0.09 0.03	0.09 0.03	0.02 0.67	1.00				
5. FOS S1	-0.13 0.00	-0.02 0.70	0.01 0.77	-0.01 0.85	1.00			
6. SNS S2	0.74 <.0001	0.43 <.0001	0.03 0.50	0.11 0.01	-0.05 0.20	1.00		
7. ONS S2	0.31 <.0001	0.59 <.0001	0.20 <.0001	0.01 0.79	-0.08 0.05	0.40 <.0001	1.00	
8. FOS S2	-0.16 <.0001	-0.08 0.05	-0.01 0.76	-0.01 0.89	0.63 <.0001	-0.14 0.00	-0.17 <.0001	1.00

Note. The first data value indicates the correlation and the second data value indicates the p-value in this matrix.

Appendix H: Correlations of Individual Difference Scales in Session 3

(N=247)	1	2	3	4	5	6	7	8	9	10	11
1. SNS S1	1.00										
2. ONS S1	0.37 <.0001	1.00									
3. Vocabulary S1	-0.03 0.63	0.10 0.11	1.00								
4. Working Memory S1	0.13 0.05	0.06 0.34	0.00 0.99	1.00							
5. FOS S1	-0.07 0.30	0.05 0.47	0.03 0.69	0.04 0.52	1.00						
6. SNS S2	0.75 <.0001	0.41 <.0001	-0.06 0.37	0.14 0.03	-0.01 0.86	1.00					
7. ONS S2	0.30 <.0001	0.49 <.0001	0.19 0.00	-0.05 0.45	0.00 0.94	0.34 <.0001	1.00				
8. FOS S2	-0.07 0.29	-0.01 0.90	-0.04 0.53	0.05 0.42	0.69 <.0001	-0.13 0.04	-0.08 0.20	1.00			
9. SNS S3	0.69 <.0001	0.33 <.0001	-0.02 0.80	0.09 0.16	-0.09 0.17	0.72 <.0001	0.27 <.0001	-0.07 0.27	1.00		
10. ONS S3	0.34 <.0001	0.52 <.0001	0.19 0.00	0.02 0.81	-0.04 0.56	0.41 <.0001	0.54 <.0001	-0.06 0.37	0.34 <.0001	1.00	
11. FOS S3	0.01 0.89	0.04 0.49	0.01 0.88	-0.06 0.33	0.47 <.0001	-0.04 0.50	0.00 0.95	0.52 <.0001	-0.07 0.26	-0.15 0.02	1.00

Note. The first data value indicates the correlation and the second data value indicates the p-value in this matrix

Appendix I: SEM Results for Indirect Effect of Statistics on Change in Financial Outcomes through Change in Numeric Competencies from Session 1 to Session 2

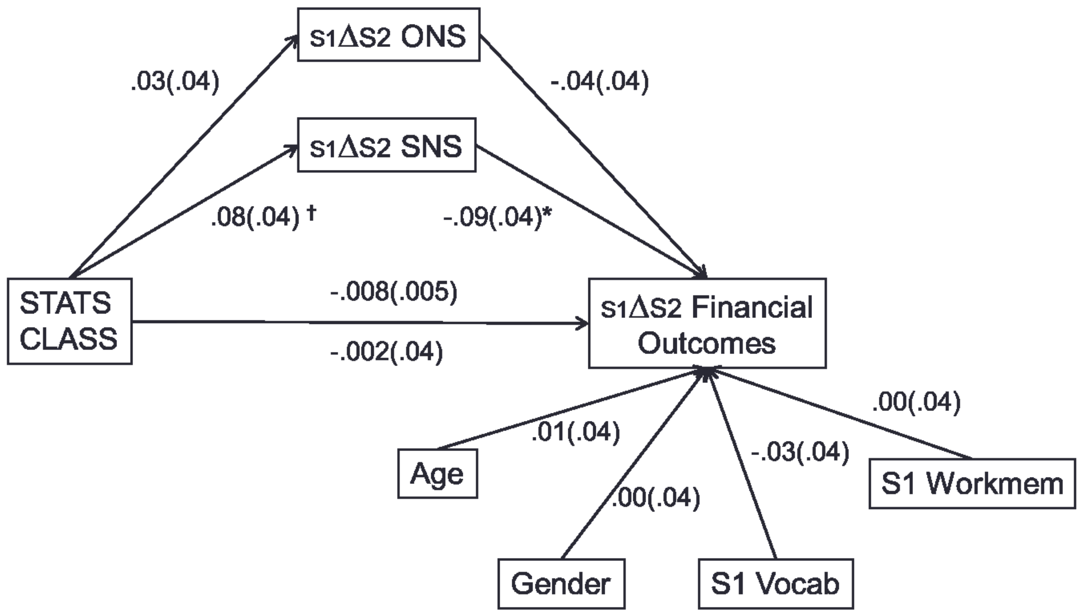

Figure 6. Hypothesis 4 full model, *B*(SE)

Note. $^{\dagger}p< .10$. $^{*}p < .05$. $^{**}p < .01$; $X^2 = 20.83$, df(9),p=.01, Relative $X^2 = 2.31$, RMSEA= .05, CFI= .89, GFI = .99.

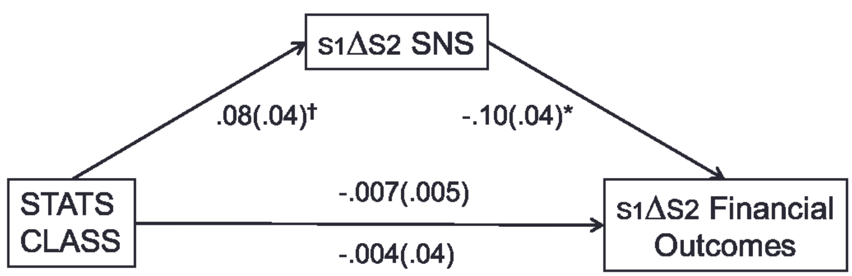

Figure 7. Hypothesis 4 final model, *B*(SE)

Note. $^{\dagger}p < .10$. $^{*}p < .05$. $^{**}p < .01$.

Appendix J: SEM Results for Indirect Effect of Statistics on Change in Financial Outcomes through Change in Numeric Competencies from Session 1 to Session 3

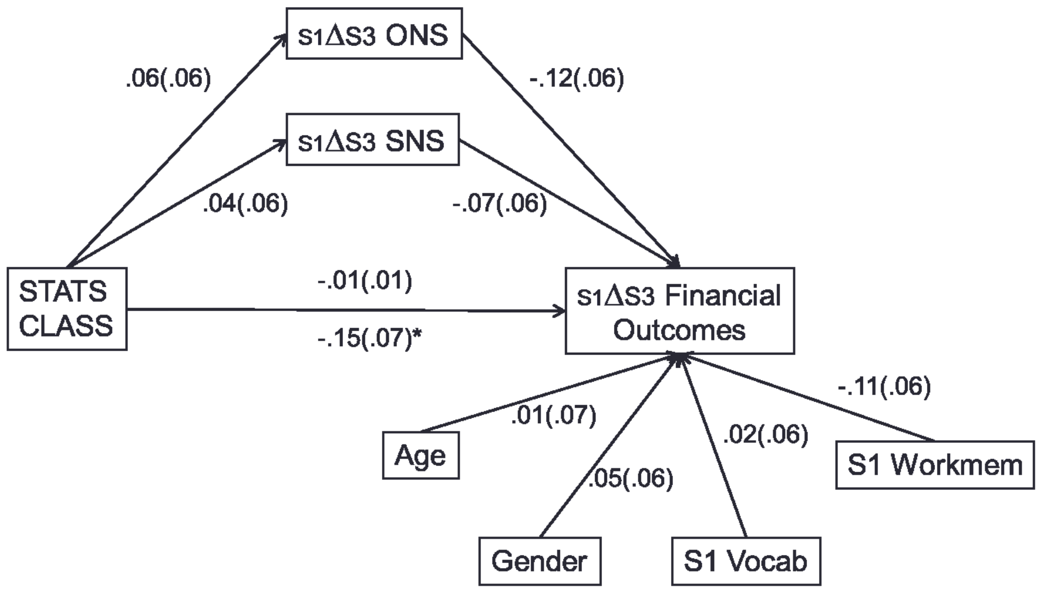

Figure 8. Hypothesis 5 full model, *B*(SE)

Note. $^{\dagger}p < .10$. $^{*}p < .05$. $^{**}p < .01$; $X^2 = 8.98$, df(9), p=.44, Relative $X^2 = 1.00$, RMSEA= .00, CFI =1.00 , GFI = .99.

Figure 9. Hypothesis 5 final model, *B*(SE)

Note. $^{\dagger}p < .10$. $^{*}p < .05$. $^{**}p < .01$.