UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

NUMERACY VS. INTELLIGENCE: A MODEL OF THE RELATIONSHIP BETWEEN COGNITIVE ABILITIES AND DECISION MAKING

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE

By

JINAN N. ALLAN
Norman, Oklahoma
2018

NUMERACY VS. INTELLIGENCE: A MODEL OF THE RELATIONSHIP BETWEEN COGNITIVE ABILITIES AND DECISION MAKING

A THESIS APPROVED FOR THE DEPARTMENT OF PSYCHOLOGY

BY

Dr. Edward T. Cokely, Chair

Dr. Robert Terry

Dr. Scott Gronlund

Dr. Rocio Garcia-Retamero

Acknowledgements

This thesis would not have been possible without the support and guidance of my advisor, Dr. Edward T. Cokely. I must also thank my co-advisors and committee members, Dr. Rocio Garcia-Retamero, Dr. Adam Feltz, Dr. Robert Terry, and Dr. Scott Gronlund, from whom I have learned so much. Special thanks go to my friends and colleagues at the Decision Analytics Lab, the National Institute for Risk and Resilience, and the Risk Literacy Team. I feel so fortunate to be surrounded by so many wonderful and hardworking people. Finally, a special thanks to my dad and my friends and family for their unwavering support and encouragement.

Table of Contents

List of Tables

List of Figures

Abstract

For nearly 150 years, psychological research and theory has documented a link between general intelligence and decision making performance. This suggests that individual differences in decision making (and other life outcomes – e.g., health, wealth, and happiness) may primary follow from individual differences in heritable and relatively stable general cognitive capacities (e.g., fluid intelligence). However, over the past 40 years there have been great developments in decision making measurement and theory. These developments allow for more precise and comprehensive assessments of essential judgment and decision making tasks, which have traditionally been neglected in general intelligence research. Recent research further indicates that the influence of statistical numeracy (i.e., practical probabilistic math skills) on decision making skill tends to be far greater than that of fluid intelligence or other general cognitive abilities (Cokely et al., 2012; Ghazal, 2014). Here we report results from one of the most comprehensive studies of cognitive abilities and decision making skill, including data from 300 participants who completed a five hour assessment battery. Using confirmatory factor analysis and structural equation modeling, we present a new model of general decision making skill where numeracy mediates the relationship between intelligence and decision making skill. Discussion focuses on implications for a refined factor structure of human cognitive abilities and related applications (e.g., adaptive training, risk communications).

Keywords: Cognitive abilities, decision making skill, numeracy, intelligence

Chapter 1: General Decision Making Skill

For nearly 150 years psychological research and theory has documented a link between general intelligence and better life outcomes (e.g., health, wealth, happiness; Lubinski, 2016). This research suggests that individual differences in decision making primarily follow from individual differences in heritable and relatively stable cognitive capacities (e.g., fluid intelligence). Traditional accounts have emphasized causal mechanisms that assume these differences are a function of an individual's ability to suppress emotions and process large amounts of information, to make optimal decisions in accord with expected utility and multi-attribute functions (Edwards, 1954; Savage, 1954; Tversky & Kahneman, 1974). A large body of previous research has focused on traditional cognitive abilities (i.e., fluid intelligence and working memory) as an explanation for why some individuals are better able to detect errors and override biased judgments (i.e., System 1 and System 2 processing; Tversky & Kahneman, 1974). By this logic, only some individuals are able to engage in superior decision making: those who have the mental capacity to suppress their first intuitive (heuristic) judgment, and override it with a judgment that is more consistent with subjective expected utility (Del Missier et al., 2010; Frederick, 2005). However, people's decision processes rarely follow standard Bayesian probability theory or subjective expected utility calculations; human decision making often *approximates* normative standards of rationality (Cokely & Kelley, 2009; Gigerenzer, 1996). Recent research further suggests that there are individual differences in decision making skill and heuristic use, such that some people are more effective and efficient decision makers than others (Aczel et al., 2015; Cokely et al., 2018; Dhami, Schlottmann, & Waldmann, 2012; Stanovich, West, & Toplak, 2016).

However, a growing body of research suggests this picture may be incomplete. Research on numeracy and inductive reasoning suggests individual differences in decision making should be viewed as an *acquired* skill, and decision making skills may be largely independent of general fluid intelligence. Most broad and integrative analyses of general decision making quality are most strongly linked to acquired math skills. These acquired math skills (e.g., statistical numeracy) may also fully mediate the link between intelligence and decision quality among typical, healthy young adults from industrialized countries. However, until now, there has not been a robust test between statistical numeracy and intelligence on a broad decision making skill battery.

Numeracy and Decision Making

The importance of numeracy (i.e., ones' ability to reason with numbers, especially in real-world contexts) was appreciated in the social sciences long before psychology and economics researchers began to study its empirical importance (Huff, 1954; Paulos, 1988). In the last twenty years, empirical research has grown in this area as a result of the introduction of a simple three-item psychometric numeracy test. This test was created to assess the relationship between numeracy and ones' understanding of "risk reduction expressions about the benefit of screening mammography" (Schwartz, Woloshin, Black, & Welch, 1997, p. 966). Moreover, early research was primarily focused on everyday problem-solving for real world issues (e.g., medical, political, and financial decision-making; Lusardi, 2008; Nelson et al., 2008; Schapira, Walker, & Sedivy, 2009).

Recent research, however, converges on the notion that numeracy often predicts naturalistic, high-stakes, ecologically valid decisions (Cokely, et al., 2012; Lipkus & Peters, 2009; Peters & Levin, 2008; Peters et al., 2006; Petrova et al., 2016; Reyna, Nelson, Han, & Dieckmann, 2009), and often is one of the strongest predictors of other more theoretical and paradigmatic decision making tasks (Ashby, 2017; Ghazal, 2014; Parker, Bruine de Bruin, & Fischhoff, 2015; Szaszi, Palfi, Szollosi, Kieslich, & Aczel, 2018; See also Cokely et al., 2012; 2018; Garcia-Retamero & Cokely, 2013; 2017).

After much development in psychometric testing (Cokely et al., 2012; Fagerlin et al., 2007; Lipkus et al., 2001; Schapira et al., 2009; Weller et al., 2013), the Berlin Numeracy Tests are now "the strongest single predictors of individual differences in general decision making skill, including the ability to evaluate and understand risk (i.e., risk literacy)" (Cokely et al., 2018, p. 478). However, until recently, broad efficient research assessments covering the full range of quantitative numeracy skills have not been widely available. To address this gap we turn to a comprehensive adult numeracy framework derived following a systematic review of 29 existing mathematical and numeracy frameworks and related national education standards (Ginsburg, Manly, & Schmitt, 2006). This framework indicates that the modern core collection of essential components of adult numeracy in industrialized societies typically involves: *operations* including computation, estimation, rates, ratios, proportions, and percentages; *probability* including collection, organization, and display of data, analysis and interpretation of data, chance and probability, and inferential reasoning; *geometry* including measurement units, trigonometric ratios, angles and lines, shapes, perimeter, area, and volume, length, width, height, and radius; and *algebra* including algebraic expressions, symbols, equations, and functions.

Building on this and related statistical literacy frameworks (Gal, 2003; Ginsburg et al., 2006; Kutner et al., 2006), the most recent developments within the Berlin Numeracy Tests family is the Component form of the Berlin Numeracy Test (BNT-C). This work provides simultaneous estimates of: (i) Full-scale adult numeracy, (ii) Adult numeracy subscales (i.e., operations, probability, algebra, geometry), (iii) Statistical numeracy (i.e., a composite of operations and probabilities), and (iv) Conventional numeracy (i.e., a composite of algebra and geometry; see Ghazal, 2014).

Measuring Decision Making Skill

Mapping and measuring individual differences in judgment and decision making (JDM) quality has been one focus of recent research within the JDM field (Baron, 1985; 2008; Bruine de Bruin et al., 2007; Stanovich & West, 2000; Toplak, West, & Stanovich, 2011). Prominently, the Adult Decision Making Competence (ADMC) was one of the first ventures to create a comprehensive measure of decision making skill. This test includes seven components (see Table 2), and theoretically assesses individual ability across four major decision making domains: belief assessment (judging the likelihood of outcomes), value assessment (evaluating outcomes), integration (combining beliefs and values in making decisions), and metacognition (understanding the limits of one's abilities; Bruine de Bruin et al., 2007; but see also Edwards, 1954).

This measure was critical for many developments in the JDM field, including the findings that (i) decision making is a general skill that is related to life outcomes and (ii) similar to intelligence, decision making provides a positive manifold (i.e., tasks and variables are positively correlated; Bruine de Bruin et al., 2007; Del Missier et al., 2013).

Meanwhile, following from the heuristics and biases approach to decision making (Tversky & Kahneman, 1974), Toplak, West, & Stanovich (2011) also created a battery of 15 items used to assess major forms of biases that follow from heuristics. Theoretically, this assessment would be a direct test of System 1 and System 2 processing, and was validated using the Cognitive Reflection Test (Frederick, 2005).

However, both of these tests are largely ones of *paradigmatic* decision making and as such have limits in terms of their practical implications. In order to more effectively measure individual differences in broad decision making skills, research would require other measures that capture *ecological* decision making (i.e., creating stimuli in which environmental properties are preserved; Brunswik, 1955; Dhami, Hertwig, & Hoffrage, 2004). Only within the last decade, has research begun to focus on these more high-stakes, ecologically valid decisions (Banks, O'Dea, & Oldfield, 2010; Cokely, et al., 2012; Garcia-Retamero & Cokely, 2014; Gigerenzer, Gaissmaier, Kurz-Milcke, Schwartz, & Woloshin, 2007; Pleskac & Hertwig, 2014). Leveraging this research, an Ecological Decision Making Battery has been included to broadly assess decision making skill across domains.

Current Study

The present study aims to explain *why* some people are better at making decisions than others by mapping the observed relations between traditional measures of intelligence (fluid and crystallized intelligence), decision making skill, and other measures of acquired skills (e.g., component numeracy). Though I have not ventured to create any new measures of decision making skill in the present study, I have provided an improved assessment of decision making skill, by leveraging a broader set of existing and validated decision tasks. As such, while previous studies have typically only assessed one of the three types of decision tasks (paradigmatic, ecological, or heuristics & biases), the present study aimed to assess all three concurrently. Taken together with standard measures of fluid and crystallized intelligence, as well as robust measures of component numeracy (i.e., statistical numeracy versus conventional numeracy), the present research hopes to begin to answer the question: *Why* are some people better at making decisions than others?

Converging research indicates that the influence of statistical numeracy (e.g., practical probabilistic math skills) tends to be far greater than that of working memory or other general cognitive abilities (e.g., fluid intelligence) (Bruine de Bruin, Parker, & Fischhoff, 2007; Cokely et al., 2018; Cokely et al., 2012; Del Missier et al., 2012). However, the relative contributions of statistical numeracy on intelligence and standard measures of cognitive abilities (i.e., fluid and crystalized intelligence, as measured by tests of verbal and quantitative reasoning) go relatively unmapped. To further map and test the robustness of observed relations, I aimed to conduct one of the most comprehensive integrative studies of cognitive abilities and decision making skill. The present study (i) included better full scale intelligence (fluid and crystallized intelligence), (ii) was tested on a large, diverse, and perhaps more representative sample, and (iii) tested an even broader decision making skill set (i.e., included tests of paradigmatic and ecological decision tasks as well as direct tests of System 1 and System 2 processing – e.g., heuristics and biases). I hypothesize that numeracy will be the best single predictor of decision making skill, and that it will also mediate the relationship between intelligence and decision making skill.

Chapter 2: Methods and Instruments

Participants and Procedure

510 University of Oklahoma students volunteered to participate in a five-hour study that was completed in eight segments, in exchange for class credit. A total of 309 participants completed all segments of the study in accord with study and course requirements and instructions.[1] The first segment was completed in the lab during a one and a half hour face-to-face session. The remaining seven segments were assigned in a specific order and were completed at home via the Qualtrics Survey Software, an online platform. Each of the seven sections were designed to take roughly thirty minutes so as to provide adequate breaks and considerable flexibility for busy college student participants. All participates were advised that once started, the entire segment should be completed in a quiet setting that would allow for un-interrupted concentration for the duration of that segment.

In the lab, participants completed a large battery of numeracy tests, followed by the National Adult Reading Test (NART). At home, participants completed a battery of decision making tasks (e.g., Adult Decision Making Competence), intelligence measures (fluid and crystallized), and individual differences measures. Of the 309

[1] Due to the at-home nature of the study, many participants deviated from task requirements, skipping or quickly advancing through sections at rates which imply the instructions could not have been read or followed (e.g., participants who spent less than 70 seconds on the Cattell Culture Fair Test or Raven's Advanced Progressive Matrices).

participants, 112 (36.2%) were male and 197 (63.7%) were female. 97% were under the age of 25. All ethical standards as outlined by the IRB were followed.

Berlin Numeracy Tests

The Berlin Numeracy Test (BNT) is one of the most efficient predictors of risk literacy and general decision-making skill, especially for educated individuals from industrialized countries (Cokely et al., 2012; 2018). Taking less than three minutes to complete, this adaptive 2-4 item test has demonstrated strong predictive validity across domains, including medical decision making (Petrova et al., 2016; Reyna et al., 2009), financial decision making (Lusardi, 2012; Skagerlund, et al., 2018), and extreme weather decision making (Allan et al., 2017).

When studying a diverse population, the Berlin Numeracy Test is often paired with the three-item Schwartz et al. (1997) numeracy test. The Schwartz test was one of the first published tests of statistical numeracy and it assesses individuals' understanding of proportions and probabilities, especially for those with lower statistical numeracy ability. Taken together these two tests provide a robust assessment where a wide range of skill can be efficiently measured.

Recent developments on the Berlin Numeracy Test has led to the Components form of the BNT (i.e., the BNT-C). This test includes four components (Operations, Probability, Geometry, and Algebra; Ghazal, 2014). Each component consists of five questions to assess highly educated samples (e.g., normed with a large sample of undergraduate students at a predominately engineering university), as well as an additional four questions intended to be used for the general population (e.g., normed with a representative sample of U.S. residents). In the present study each component included all nine items (36 total; see Appendix A). Combining the four and five item tests was hypothesized to provide a more robust assessment of composite numeracy. With higher fidelity measures of composite and component numeracy, results should better map the relationship between decision making skill and component numeracy skills. In addition, a larger and more representative sample (e.g., undergraduate students at a large public university) would provide an opportunity to complete further item and test analyses (i.e., Item Response Theory).

Item Response Theory (IRT) was conducted using the ltm package in R (Rizopoulus, 2015). This two-parameter logistic item analysis optimizes scales on psychometric difficulty and discriminability parameters. The analysis showed the first item in each component had low difficulty and low discriminability (i.e., it was too easy; almost all participants answered them correctly). This type of item has the propensity to pose problems for scale development (e.g., can impact reliability and obfuscate future results). To optimize the composite and component BNT-C scales these items were removed. Eight items remained in each component score, as shown in Figure 1.

Consistent with the previous analyses the four numeracy components are combined into two sub-scales (i.e., statistical numeracy and conventional numeracy). Statistical numeracy is comprised of the Operations and Probability components, whereas Conventional numeracy uses the Geometry and Algebra components (Ghazal, 2014). These two scales should be theoretically related (i.e., they are both math), but also distinct (i.e., probabilistic math involves more inductive reasoning whereas conventional math requires more deductive logic). BNT-C Statistics (M = 7.48, SD =

3.86) and BNT-C Conventional (M = 8.71, SD = 3.48) also had strong psychometric properties, as they both provided acceptable reliability (α = 0.82 and α = 0.77, respectively). Taken together, the composite score of the four components also demonstrated good test information, across the full range of ability (see Figure 1 Test Information Function and Item Information Curves).

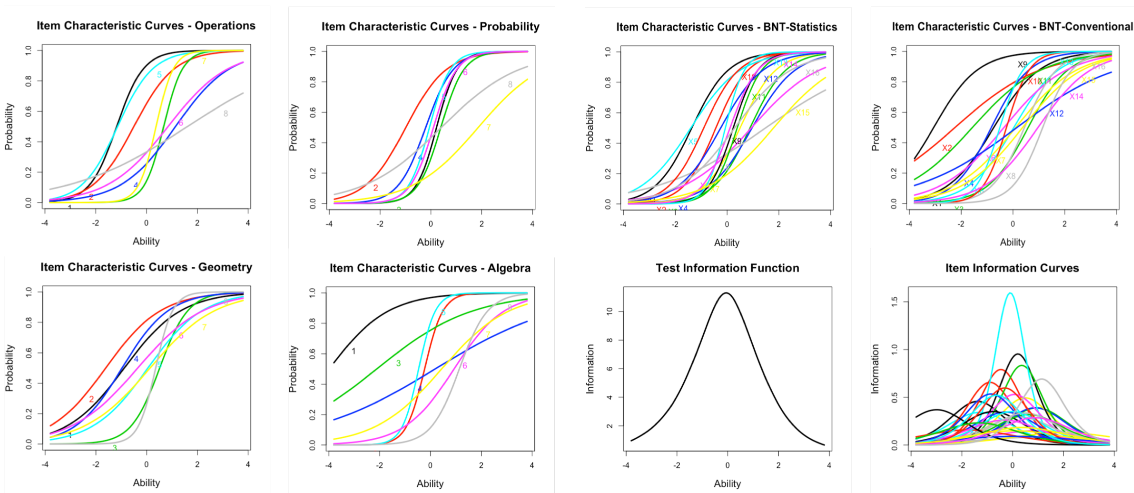


Figure 1. Item Characteristic Curves for BNT-C

Fluid and Crystallized Intelligence

Many previous studies of intelligence and decision making have included a narrow set of fluid intelligence measures (e.g., Raven's Advanced Progressive Matrices), but have not considered a broad range of intelligence factors (e.g., crystallized intelligence, verbal ability, logical inference). As such, this study relied on a number of validated scales to assess intelligence. Following from Carroll's (1993) Model of Cognitive Abilities, the selected tests should be robust measures of the suggested important narrow abilities for (i) fluid intelligence (sequential reasoning, inductive factors, and quantitative reasoning) and (ii) crystallized intelligence (printed language, language comprehension, and vocabulary knowledge).

Measures of Fluid Intelligence

Raven's Advanced Progressive Matrices (RAPM). One of the most common measures of general fluid intelligence, this nonverbal test involves a series of 3x3 matrix reasoning problems. The short-form of this test includes 12 items (Bors & Stokes, 1998; Raven, Court, & Raven, 1988).

Cattell Culture Fair IQ Test, Scale 3 Form A. The Cattell Culture Fair Test was created to estimate intelligence, without the influence of cultural differences. There are four components to this nonverbal test, (i) Series, (ii) Classification, (iii) Matrices, and (iv) Conditions. In part 1, participants have 3 minutes to complete 13 Series problems, where they must select from one of six choices which comes next in the series of four images. In part 2, participants have 4 minutes to complete 14 problems, where they must determine which two figures are in some way different from the other three figures. In part 3, participants again have 3 minutes to complete 13 problems, where they must select which of six options completes the matrix. Here, five are 2x2 matrices, two are complete 3x3 matrices, and the last six are 3x3 matrices that appear to have some portion of the matrix missing. Finally, in part 4 participants have 2.5 minutes to complete 10 problems, in which they must select which of five figure choices duplicates the conditions indicated in the first figure (Cattell, 1973).

Because structural equation modeling with parcels (i.e., scale scores) relies on the strong assumption that the scales are unidimensional, I conducted factor analysis on these four component scores. This analysis indicated that the Cattell Culture Fair best separated into two factors. The first factor (called Cattell Culture Fair A), was made up of indicators 1 and 4 (Series and Conditions), and the second factor (Cattell Culture Fair B), was defined by indicators 2 and 3 (Classification and Matrices).

Measures of Crystallized Intelligence

Wonderlic Personnel Test. This test of cognitive abilities was created to assess aptitude for learning and problem solving in potential employees. This 50-item test is administered in 12 minutes. However, because many participants cannot complete it, a score of 20 indicates average intelligence. Even though some previous research has indicated the Wonderlic is not a measure of fluid or crystallized intelligence (Matthews & Lassiter, 2007), I suggest that this measure is more similar to crystallized intelligence (verbal ability, reading comprehension, etc.) than fluid intelligence (deductive reasoning). The Wonderlic has been validated and used by many employers as a selection assessment (Wonderlic, 2018; but see also Wonderlic & Hovland, 1939).

Employee Aptitude Survey. The EAS is a test of general logic. Participants have five minutes to complete six sets of questions. Each question set provides five facts, and is then followed by five conclusions. Participants are asked to determine if each conclusion is True, False, or Uncertain (Ruch & Ruch, 1963).

National Adult Reading Test. The NART asks participants to read a list of 50 words, which are all irregular with respect to the common rules of pronunciation. The word list increases in difficulty, and previous research demonstrated word reading ability and general intelligence are highly correlated (Nelson & McKenna, 1975). This task was administered in the lab. Each participant sat in a quiet room with a Research Assistant and spoke the words, as they were presented on a screen, one at a time (Nelson & Willison, 1991).

Models of Intelligence

In accord with previous research on intelligence batteries, I hypothesized that Ravens' APM and the Cattell Culture Fair Test would be robust measures of general fluid intelligence, whereas the Wonderlic, EAS, and NART would measure crystallized intelligence. Exploratory Factor Analysis[2] on these six measures of cognitive abilities confirmed the predicted result. Parallel Analysis suggested a two factor solution, and based on this, the EFA produced a model where Ravens' APM and both Cattell Culture Fair indicators made up the first factor (fluid intelligence; g_f), followed by a second factor defined by the Wonderlic, EAS, and NART (crystalized intelligence; g_c) (See Table 1 and Figure 2).

Table 1. Exploratory Factor Analysis on Standard Intelligence Measures

	g_f	g_c
Ravens' APM	**0.72**	-0.06
Cattell Culture Fair B	**0.67**	-
Cattell Culture Fair A	**0.51**	0.12
Wonderlic	-	**0.95**
Employee Aptitude Survey	0.29	**0.37**
National Adult Reading Test	-0.12	**0.44**
Proportional Variance	**22%**	**21%**
Cumulative Variance	**22%**	**43%**

Note. Standardized factor loadings

[2] Confirmatory Factor Analysis returned similar results, with a strong model fit, $\chi^2_8 = 14.82$, $p > 0.05$, $\chi^2/df = 1.85$, CFI = 0.98, TLI = 0.97, RMSEA = 0.05 [0.00, 0.09], SRMR = 0.04.

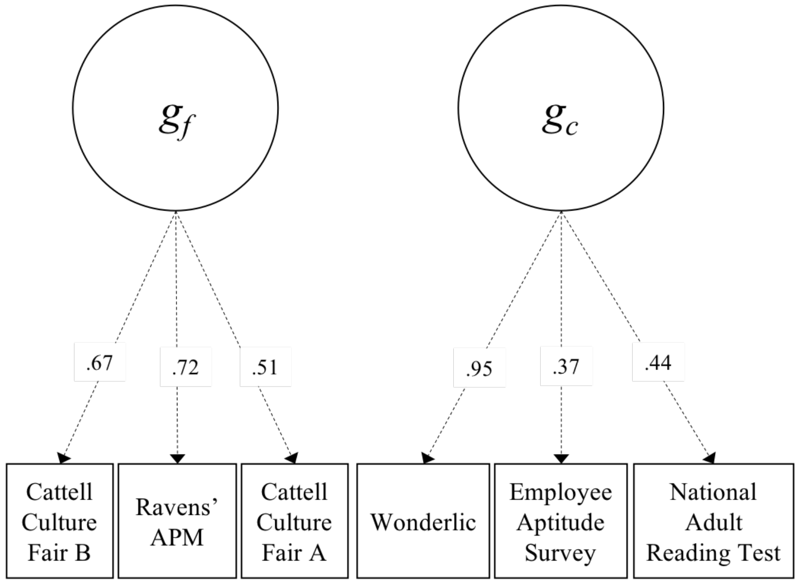


Figure 2. Factor Structure for Standard Intelligence Measures

Consistent with trends across previous analyses (Carroll, 1993), if conventional numeracy (BNT-C Conventional; Figure 3) is included in the EFA with standard measures of intelligence, a two-factor solution remains and conventional numeracy loads with fluid intelligence (g_f). However, if other important components of numeracy (i.e., statistical numeracy) are also properly represented, the EFA suggests a three-factor solution, where numeracy becomes the first factor (Figure 4). Of note, when numeracy is properly represented in the model, conventional numeracy shifts from fluid intelligence and instead primarily loads with numeracy. This finding provides evidence consistent with the notion that numeracy – especially statistical numeracy – was under-represented in previous analyses of intelligence.

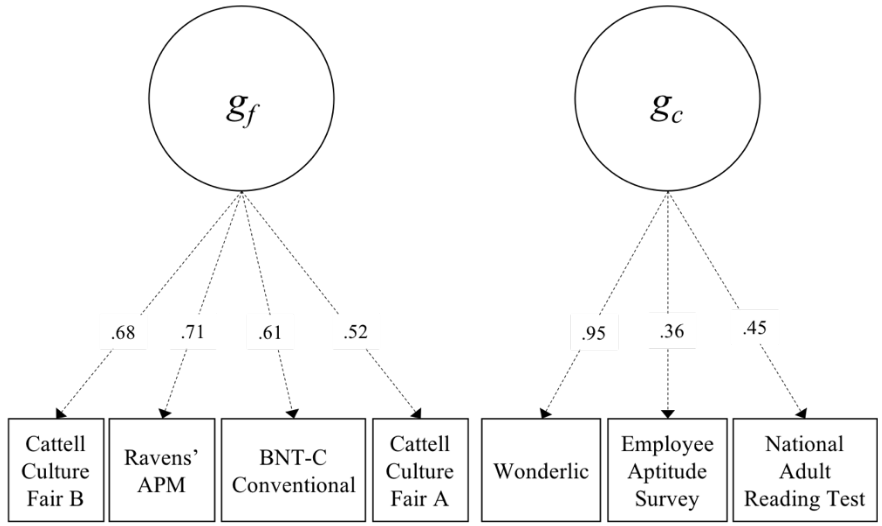


Figure 3. Factor Structure for Intelligence Measures with Conventional Numeracy

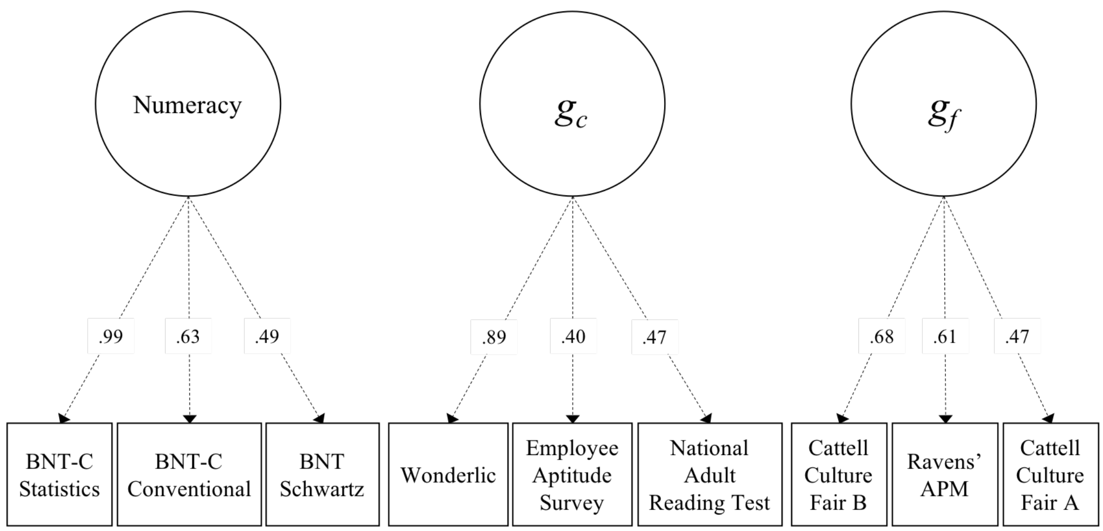


Figure 4. Factor Structure for Intelligence Measures with Full Scale Numeracy

Heuristics and Biases

Following from a tradition of heuristics and biases in the JDM field, a battery of fifteen classic heuristic-and-biases tasks was created (Toplak, West, & Stanovich, 2011). Theoretically, these items reflect aspects of "rational thought, including probabilistic reasoning, hypothetical thought, theory justification, scientific reasoning, and the tendency to think statistically” (Toplak et al., 2011, p. 1277). This battery consists of one causal base-rate problem, two sample-size problems, one problem assessing sensitivity to regression to the mean, two gambler’s fallacy problems, one conjunction problem, one covariation detection problem, one methodological reasoning problem, one Bayesian reasoning problem, a framing problem, one problem assessing denominator neglect, a probability matching assessment, a sunk cost problem, and an outcome bias problem (Appendix B). Toplak et al. (2011) also found a strong correlation between the Cognitive Reflection Test (Frederick, 2005) and the composite score.

General Decision Making Skill

Adult Decision Making Competence (ADMC)

This battery of seven types of decision tasks was created to assess how well individuals make decisions. The original seven decision types include (i) Resistance to framing, (ii) Recognized social norms, (iii) Under/Over confidence, (iv) Applying decision rules, (v) Consistency in risk perception, and (vi) Resistance to sunk cost, and (vii) Path independence (Bruine de Bruin et al., 2007; Appendix C). See Table 2 for sub-scale details and example items.

Ecological Risk Literacy and Paradigmatic Risky Decision Making Tasks

This battery involved seven components of ecological and paradigmatic risky decision tasks. The seven components include (i) Ecological Risk Literacy Medical Decisions, (ii) Ecological Risk Literacy Financial Decisions, (iii) Expected Values, (iv) Choice Consistency in Lotteries, (v) Intertemporal Choice, (vi) Reference Class &

Class-Inclusion Illusions (i.e., Ratio Bias and Denominator Neglect), and (vii) Realistic Risky Decisions (Ghazal, 2014). See Table 3 for sub-scale details and example items.

General Decision Making Skill Assessment (GDMS)

The standardized proportional composite of the seven components of the ADMC and the seven components of the ecological decision battery were combined to make one composite score of general decision making skill. This measure is treated as the dependent variable. In line with previous research, there were strong positive correlations between decision making skill, intelligence and numeracy components $r > .3$, $p < .01$; See Table 4).

Table 2. Components of the Adult Decision Making Competence (ADMC)

A-DMC Component	Description	Example Item
Consistency in Risk Perception	This task asks participants to judge the probability of various events happening in two different time frames.	What is the probability that someone will steal something from you during the next year/in the next 5 years?
Recognizing Social Norms	This test measures how well participants judge social norms. Participants assess 16 undesirable behaviors.	First set: "It is sometimes OK to steal under certain circumstances." Second set: "Out of 100 people your age, how many would say it is sometimes OK to steal under certain circumstances."
Resistance to Sunk Costs	This test measures the ability to ignore prior financial and time investments that are no longer paying off when making decisions.	After a large meal at a restaurant, you order a big dessert with chocolate and ice cream. After a few bites you find you are full and you would rather not eat any more of it. Would you be more likely to eat more or to stop eating it?
Resistance to Framing	This task measures whether value judgments are effected by irrelevant variations in how the problem is presented.	Recent evidence has shown that a pesticide is threatening the lives of 1,200 animals. ... Which option do you recommend: (1) Option A: 600 animals will be lost for sure (2) Option B: 75% chance 400 animals will be lost, and 25% chance that 1,200 animals will be lost. The same item is then presented in a "gain" format (e.g., 600 animals are saved for sure).
Applying Decision Rules	This task evaluates the ability to apply decision rules, by asking participants to choose between DVD players with different ratings and features.	Lisa wants the DVD player with the highest average rating across features. Which one of the presented DVD players would Lisa prefer?
Path Independence	This test presents item pairs posing normatively equivalent choices between gambles. The participants' choice should not be affected by normatively irrelevant changes.	Which do you like best: (1) Flip a coin. If heads, win \$100. If tails, win \$0 (2) Sure Win. Win \$50 for sure. Performance measured by participant's consistency in choices.
Under/Over Confidence	This test measures how well participants can assess their own knowledge. Participants first answer a true/false question, then assess their confidence in that answer.	True or False: Stress makes it easier to form bad habits. How confident are you? 50% (just guessing) to 100% (absolutely sure)

Table 3. Components of the Ecological Decision Battery (Ghazal, 2014)

Component	Description	Example Item
Ecological Risk Literacy		
Medical	Items were included for risky medical decisions that were representative of the natural ecology.	Cervical cancer is very rare. 4 out of 100,000 women are affected by this cancer. The human papillomavirus (HPV) vaccine is federally approved and is being promoted as a method that helps reduce cervical cancer. Given information medical trial results, participants must rate the effectiveness of vaccines.
Financial	Items were included for risky financial decisions that were representative of the natural ecology.	Imagine you take out a $50,000 federal student loan to help pay for college. You are offered four possible repayment plans. Given information about each of these plans, participants must determine which plan is best, based on different criteria.
Prospect Evaluations		
Risky Prospects: Expected Values	Paradigmatic risky prospect evaluations across a wide range of risk and EV ratio ranges.	Which of the two options do you prefer: (1) Lose $50 (2) 50% chance to lose $400
Risky Prospects: Choice Consistency in Lotteries	Given the above expected value problems, to what extent do participants show consistency across gain and loss frames.	Which of the two options do you prefer: (1) Gain $50 (2) 50% chance to win $400
Intertemporal Choice	This test presented a series of theoretically important time-reward preference tasks. Items included had differing time intervals and reward amounts.	Which of the two options do you prefer: (1) $3400 this month (2) $3800 next month
Reference Class & Class-Inclusion Illusions	This task measures ones propensity to struggle with effects like denominator neglect or ratio bias.	With a new drug the risk of death from a heart attack reduced for people with high cholesterol. A study of 900 with high cholesterol showed that 80 of the 800 people who have not taken the drug died after a heart attack, compared with 16 of the 100 people who did take the drug. How beneficial was the drug?
Realistic Risky Decisions	Medical and weather decisions that reflect the natural ecology, including decisions based on advertisements.	Mrs. Jones is told she has a 28 in 1,000 chance of dying from cancer and a 59 in 1,000 chance of dying from a stroke. Mrs. Jones's doctor tells her that a new pill, STROKEX, will lower her chance of dying from stroke by 50%. Another pill, CANCERX will lower her chance of dying from cancer by 50%. Assume she can only take 1 pill. Assuming the 2 pills are equally safe and cost the same, which should she take to minimize her risk of death?

Table 4. Correlations Between Raw Variable Scores

		1	2	3	4	5	6	7	8	9	10	11	12	13
Numeracy														
1	BNT-C Statistics (16 item)	1.0												
2	BNT-C Conventional (16 item)	.72**	1.0											
3	BNT-S (7 item)	.63**	.52**	1.0										
Fluid Intelligence														
4	Raven's APM	.42**	.42**	.37**	1.0									
5	Cattell Culture Fair A	.34**	.37**	.36**	.39**	1.0								
6	Cattell Culture Fair B	.36**	.43**	.34**	.46**	.41**	1.0							
Crystallized Intelligence														
7	Wonderlic	.46**	.38**	.45**	.37**	.41**	.40**	1.0						
8	Employee Aptitude Survey (EAS)	.34**	.29**	.31**	.34**	.30**	.36**	.51**	1.0					
9	National Adult Reading Test (NART)	.13**	0.07	.17**	0.06	.14**	.16**	.35**	.17**	1.0				
Decision Making														
10	Ecological Decision Making	.50**	.47**	.48**	.39**	.33**	.31**	.37**	.36**	.18**	1.0			
11	ADMC	.44**	.33**	.42**	.37**	.33**	.34**	.46**	.32**	.23**	.45**	1.0		
12	Heuristics & Biases	.51**	.41**	.44**	.34**	.32**	.20**	.34**	.23**	.13**	.32**	.39**	1.0	
13	Composite General Decision Making Skill (GDMS 14)	.55**	.45**	.52**	.43**	.38**	.38**	.47**	.39**	.24**	.84**	.86**	.41**	1.0

Chapter 3: Results and Discussion

To test the relationships between decision making skill, numeracy and general cognitive abilities I first start with an assessment of the overall single best predictors of general decision making skill (as measured by the GDMS Assessment). I then derive a best-fitting multiple regression using stepwise regression. Finally, I use latent trait analysis and structural equation modeling for a more robust mediation analysis.

Single Predictors and Stepwise Regression

I hypothesized that taken as single predictors of the general decision making skill assessment, numeracy will out-predict all other intelligence measures. I conducted a series of simple linear regressions, where each numeracy component and standard cognitive ability was an independent variable in a single-predictor regression, and compared R-squared statistics across models (Figure 5). Statistical numeracy out-predicted any single measure of standard cognitive abilities. Further, BNT-Schwartz (a quick, seven item test) accounted for nearly 93% of the total variance explained by BNT-C Statistics (16 items).[3] These results provide evidence that when time and resources are limited, the single best predictor of general decision making skill is statistical numeracy.

I next consider the relationship between these factors more comprehensively. Using stepwise regression to explain the GDMS Assessment, I again find that statistical numeracy (BNT-C Statistics) is the strongest predictor and accounts for 30% of the total variance. The Wonderlic, BNT-Schwartz, Ravens' APM and Employee Aptitude Survey also add incremental unique variance. Of note, the final variance accounted for

[3] Similar trends are present for predicting the Heuristics and Biases Task (Toplak et al., 2011).

is 42%, indicating that statistical numeracy (i.e., BNT-C Statistics and BNT-Schwartz) accounts for 78% of the explained variance (Table 5).

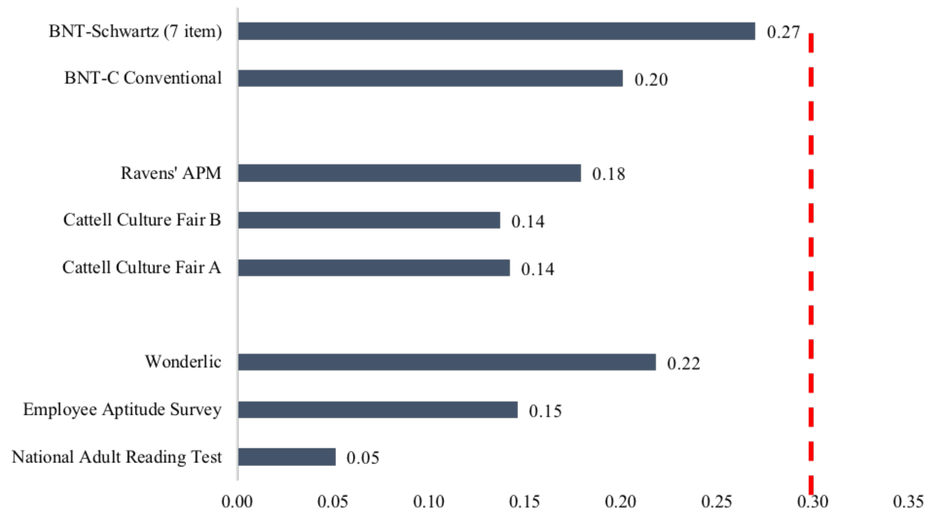


Figure 5. Single Predictors for General Decision Making Skill (GDMS) Assessment
Note. Each bar represents the R^2 from a series of single-predictor simple linear regressions. The red line represents the R^2 for BNT-C Statistics alone.

Table 5. Stepwise Regression Results

Dependent Variable	**Model**	β	R^2	R^2 **Change**	**F Change**
GDMS	**Model 1**	-	0.30	0.30	116.24
	BNT-C Statistics	0.55	-	-	-
	Model 2	-	0.36	0.06	26.55
	BNT-C Statistics	0.42	-	-	-
	Wonderlic	0.28	-	-	-
	Model 3	-	0.39	0.03	13.52
	BNT-C Statistics	0.30	-	-	-
	Wonderlic	0.23	-	-	-
	BNT-Schwartz	0.23	-	-	-
	Model 4	-	0.41	0.02	10.99
	BNT-C Statistics	0.25	-	-	-
	Wonderlic	0.20	-	-	-
	BNT-Schwartz	0.21	-	-	-
	Ravens' APM	0.18	-	-	-
	Model 5	-	0.42	0.01	4.29
	BNT-C Statistics	0.24	-	-	-
	Wonderlic	0.16	-	-	-
	BNT-Schwartz	0.21	-	-	-
	Ravens' APM	0.16	-	-	-
	EAS	0.11	-	-	-

Note. Standardized beta weights.

Structural Equation Modeling

Confirmatory Factor Analysis (CFA) and Structural Equation Modeling (SEM) were next completed using the lavaan package in R (Rosseel et al., 2017). The first measurement model I tested included two factors of intelligence (fluid intelligence and crystallized intelligence, as described in Table 1), and a third numeracy factor. The numeracy factor was defined by the two components of the BNT-C and the BNT-S (See Figure 4). I aimed to have three indicators per factor because nonconvergence or improper solutions are likely to occur when there are only two indicators per factor (Kline, 2011; Marsh & Hau, 1999). Standard convention suggests that a Comparative Fit Index (CFI) and Tucker Lewis Index (TLI) greater than 0.95 and a Root Mean Square Error of Approximation (RMSEA) and Standardized Root Mean Square Residual (SRMR) less than 0.05 suggest good fit. Further, a RMSEA less than 0.08 suggests moderate fit. This first measurement model demonstrated good fit: $\chi^2_{24} = 46.25$, $p = 0.004$, $\chi^2/df = 1.93$, CFI = 0.98, TLI = 0.96, RMSEA = 0.055 [0.03, 0.08], SRMR = 0.04. Though the RMSEA for the model is greater than 0.05 (and the lower bound on the 90% confidence interval does not include 0.00), it is likely this is largely due to measurement error. Further, the other fit indices are above satisfactory conditions, suggesting that this model has moderate to good fit. All factor loadings were strong and significant ($p < 0.001$).

Using this measurement model, I next tested a naïve structural equation model, that predicted the standardized composite General Decision Making Skill (GDMS) dependent variable (Model 1, Figure 6). To test the hypothesis that numeracy may mediate the relationship between intelligence and decision making skill, General

Decision Making Skill (GDMS) was regressed on numeracy, fluid intelligence, and crystallized intelligence (direct effects), numeracy was regressed on fluid intelligence and crystallized intelligence, and GDMS was also regressed on numeracy (indirect effects). With this test, numeracy could mediate the relationship between intelligence and decision making skill. Model 1 had good fit: $\chi^2_{30} = 56.5$, $p = 0.002$, $\chi^2/df = 1.88$, CFI = 0.97, TLI = 0.96, RMSEA = 0.06 [0.03, 0.08], SRMR = 0.04.

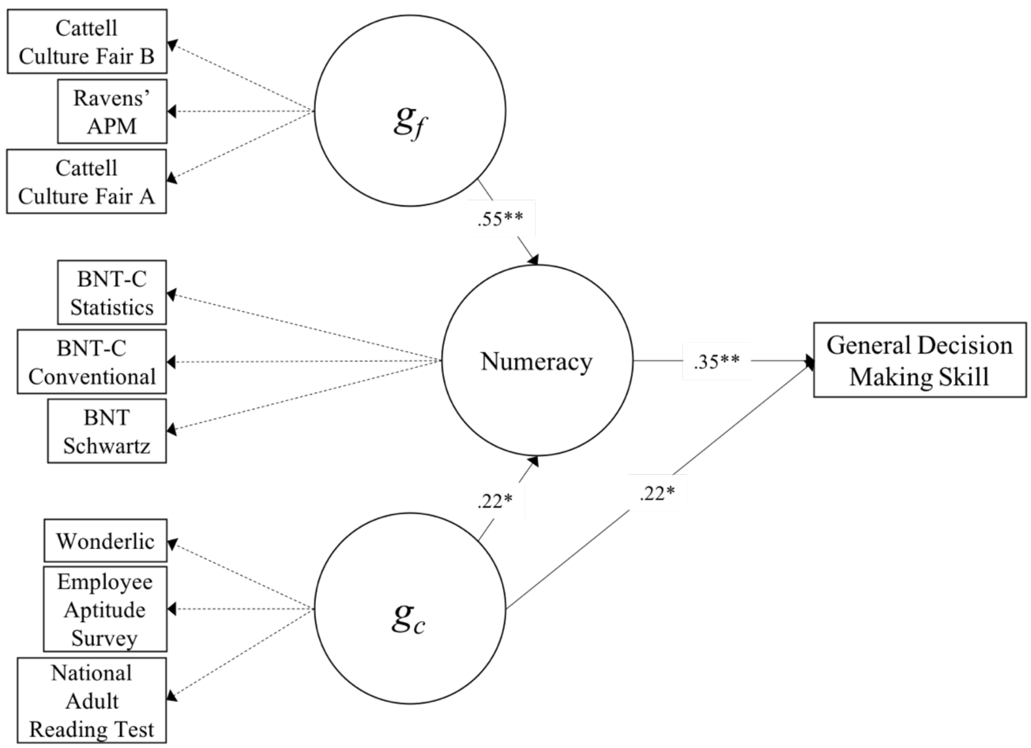


Figure 6. Model 1: Predicting the General Decision Making Skill Assessment
Note. Standardized beta weights. Only significant paths shown.

In Model 1 numeracy fully mediated the relationship between fluid intelligence and decision making, and partially mediated the relationship with crystallized intelligence. However, to aim at the more underlying theoretical constructs and causal factors, I next ventured to test a four factor measurement model, which included the same structure for intelligence and numeracy, but also defined a fourth factor, Decision

Making Skill. This factor was defined by the Ecological Decision Battery, the ADMC composite, and the Heuristics and Biases Test (Model 2, Figure 7). The initial naïve model may have obfuscated the true relationships between these factors (e.g., could be fitting noise), so given a more robust latent trait structure, we should expect more valid results. However, this model also had good fit: $\chi^2_{48} = 83.91$, $p < 0.001$, $\chi^2/df = 1.75$, CFI = 0.97, TLI = 0.95, RMSEA = 0.05 [0.03, 0.07], SRMR = 0.044.

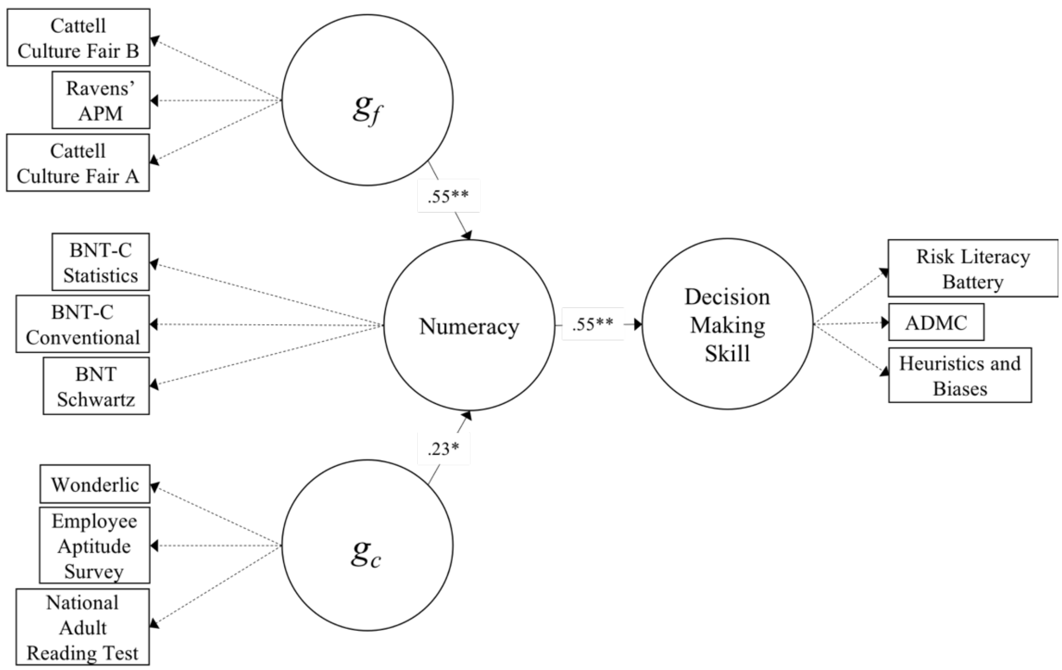


Figure 7. Model 2: Predicting Latent Trait Decision Making Skill
Note. Standardized beta weights. Only significant paths shown.

Model 2 not only demonstrates that numeracy mediates the relationship between fluid and crystallized intelligence and decision making skill, but it also provides evidence for a unifying construct of decision making skill, across the different indicators included. Here, a mediation model was chosen because previous research indicated that numeracy predicts decision making beyond fluid intelligence and memory (as assessed by Ravens' APM; Bruine de Bruin, 2007; Cokely et al., 2018; Ghazal, 2014). Furthermore, when I tested a model with only direct effects (i.e., no mediation), numeracy was the only significant predictor of decision making skill. Though their influence is relatively small, fluid and crystallized intelligence are still important indirect contributors to decision making skill. This model presents a novel finding: numeracy mediates the relationship of both fluid *and* crystallized intelligence. This breakthrough was possible only because all factors were *broadly* assessed.

Chapter 4: General Discussion

For more than 150 years, theory has assumed that intelligence (i.e., innate capacity) causes individual differences in decision making ability (Clynes, 2016; Gould, 1996; Herrnstein & Murray, 1994). However, recent research demonstrates that statistical numeracy is one of the strongest predictors of general decision making skill. Until now previous studies of intelligence had failed to adequately measure decision making skill, such that the relative contributions of statistical numeracy on intelligence and standard measures of cognitive abilities were relatively unmapped.

Given the current analyses, there are a number of notable implications. First, statistical numeracy is the single best predictor of general decision making skill. It is a unique predictor above other measures of fluid and crystallized intelligence, and also mediates the relationship between intelligence and decision making skill. As measured by The Berlin Numeracy Test (a 4 minute, 7 item scale; RiskLiteracy.org), numeracy almost doubles the predictive power of fluid intelligence (e.g., Ravens' APM; see Figure 5). For the first time, the present research provides evidence that crystallized intelligence (e.g., Wonderlic) out-predicts all measures of fluid intelligence. Further, fluid and crystallized intelligence's effect on decision making ability is mediated by numeracy (Figures 6 and 7). Taken together, this may help explain why fluid intelligence measures (i.e., measures of working memory capacity) often fail to independently predict decision making quality.

Restructured Model of Cognitive Abilities

In his seminal work, Carroll (1993) proposed the *Three Strata Theory* by analyzing over 460 data sets from 1927 and 1987. This theory posited that there are eight broad factors of cognitive abilities.[4] Following from Spearman's (1923) hypothesis, these eight broad factors fall under one overarching general ability factor, *g*. Carroll's (1993) model of cognitive abilities found that *g* was largely Fluid Intelligence, Crystallized Intelligence, and General Memory and Learning, with three essential reasoning factors best explaining Fluid Intelligence: (i) sequential reasoning, (ii) inductive factors, and (iii) quantitative reasoning.

Careful inspection, however, revealed that paradigmatic decision making skills as well as full-range math abilities were broadly neglected. First, the category labeled "inductive factors" did not broadly represent inductive logic, considering that all tasks followed deductive logic (i.e., if a person can figure out the rule, they can deduce the answer with complete *certainty*; Cokely et al., 2018). Second, the "quantitative reasoning" tasks focused primarily on conventional mathematical components (e.g., arithmetic, algebra, and geometry), rather than incorporating probabilistic, or statistical reasoning skills which also reflect inductive logic (See Carroll, 1993, p. 213).

[4] Fluid Intelligence (*Gf*), Crystallized Intelligence (*Gc*), Broad Visual Perception (*Gv*), Broad Auditory Perception (*Ga*), General Memory and Learning (*Gy*), Broad Retrieval Ability (*Gr*), Broad Cognitive Speediness (*Gs*), and Reaction Time & Decision Speed (*Gt*).

The reason for this short-coming is likely two-fold. First, while psychometric research on cognitive abilities has been ongoing for over a century, the research on decision making competence has only emerged within the last 40 years (Dhami, Schlottmann, & Waldmann, 2012). Second, while logic can be divided into two major categories (deductive and inductive), fluid intelligence tests have primarily assessed *deductive* reasoning under conditions of *certainty* – ultimately neglecting the importance of inductive logic (dealing with *uncertainty*). Further, the standard tasks used to measure fluid intelligence have relied on working memory and attentional control (Cokely et al., 2018; Cokely, Kelley, & Gilchrist, 2006; Fox & Mitchum, 2013; Kyllonen & Christal, 1990; McCabe, Roediger, McDaniel, Balota, & Hambrick, 2010). Given the importance of long-term representations in memory (i.e., representative understanding), these fluid intelligence tasks are likely not representative of typical human decision-making.

In this final model, I ventured to include all relevant cognitive ability measures: (i) each of the six standard measures of intelligence, (ii) the two components of the BNT-C (Statistics and Conventional)[5], (iii) the BNT-Schwartz (short 7 item numeracy test), (iv) the GDMS (ADMC and Ecological Decision Making composite assessment), and (v) the Heuristics and Biases measure.

There are a number of suggested ways to measure and model general intelligence (Neisser et al., 1996; van der Maas et al., 2006). Using a second order latent variable exploratory factor analysis I present a standard psychometric higher order model of general ability, *g* (Brunner, Nagy, & Wilhelm, 2012). This model

[5] The model structure of cognitive abilities is not changed if the BNT-C Conventional is not included.

demonstrated that among the first order factors, Factor 1 accounted for 22% of the overall variance – but 46% of the explained variance – and was defined by the BNT-C Statistics, BNT-C Conventional, BNT-Schwartz, Heuristics and Biases, and the General Decision Making Skill (GDMS) Composite Assessment. Factor 2 accounted for 13% of the overall variance (27% of explained variance) and was represented by the standard crystallized intelligence measures (i.e., Wonderlic, NART, and EAS). Finally, Factor 3 also accounted for 13% of the variance (27% of the explained variance) and is represented by the standard fluid intelligence measures (Ravens' APM, and Cattell Culture Fair). This higher order factor model had a common general ability *g* factor, and demonstrated good model fit, $\chi^2_{25} = 34.4$, $p < 0.01$, $\chi^2/df = 1.38$, TLI = 0.98, RMSEA = 0.04 [0.0, 0.06], SRMR = 0.03.

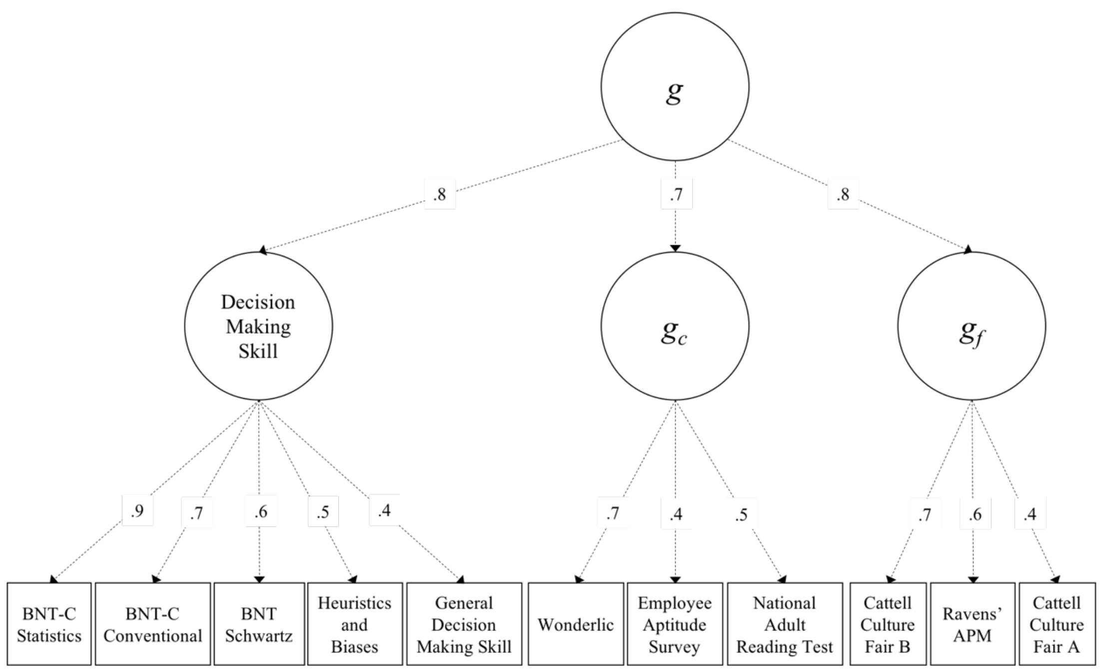


Figure 8. Model of Cognitive Abilities

Table 6. Model of Cognitive Abilities

	DMS	G_c	G_f
BNT-C Statistics	**0.95**		-0.06
BNT-C Conventional	**0.68**	-0.12	0.23
BNT-Schwartz	**0.59**	0.19	
Heuristics & Biases	**0.51**	0.12	
GDMS Assessment	**0.39**	0.30	0.14
Wonderlic	0.07	**0.73**	0.11
National Adult Reading Test		**0.53**	-0.17
Employee Aptitude Survey		**0.41**	0.28
Cattell Culture Fair B		0.08	**0.69**
Ravens' APM	0.16		**0.56**
Cattell Culture Fair A	0.10	0.16	**0.42**
Proportional Variance	**23%**	**13%**	**13%**
Cumulative Variance	**23%**	**36%**	**49%**
Explained Variance	**46%**	**27%**	**27%**

Note. Results from the "psych" package in R (Revelle, 2018).

This model demonstrates that when numeracy and decision making skill are properly represented in a model of general cognitive abilities, a new factor structure emerges whereby (i) numeracy and decision making skill, (ii) crystallized intelligence, and (iii) fluid intelligence are three distinct factors. While models of cognitive abilities may have historically neglected measures of statistical numeracy, inductive logic, and decision making skill, once they are properly included, the structure of cognitive abilities shifts. This restructuring of Carroll's (1993) cognitive ability model demonstrated that when properly represented, there is one strong single factor of statistical numeracy and decision making skill. Given the structure of the factor analysis, this novel finding further suggests that statistical numeracy and decision making skill is an important factor of general cognitive abilities.

Skilled Decision Theory

Skilled Decision Theory provides a theoretical account explaining the influence of numeracy on skilled decision making via a cascade of heuristic deliberation, risk comprehension, and affective calibration (Cokely et al., 2018; Cokely & Kelley, 2009; Ghazal, Cokely, & Garcia-Retamero, 2014). In contrast to theories that emphasize the role of abstract, emotionless decision analysis (e.g., explicitly calculating expected utilities), skilled decision theory emphasizes the role of a personally meaningful and affectively charged understanding of a decision problem, which in turn give rise to strong motivational states, enduring attitudes, specific behavioral intentions, and superior decision making and outcomes (Cokely & Kelley, 2009; Cokely et al., 2012; 2018; see also Ericsson et al., 1993; Gigerenzer et al., 1999; Peters, 2012).

As such, one of the causal mechanisms linking numeracy and decision making is that statistical numeracy tests are themselves representative judgment and decision making tasks. Solving probabilistic math problems involves the same cognitive processes as solving uncertain (situations characterized with *known* probability distributions) or risky (situations characterized by *unknown* probabilities) problems in the real-world; namely, inductive logic. When these skills are combined with personally-meaningful deliberation, they tend to promote a thorough and ecologically representative understanding (e.g., sophisticated mental situation models), circumventing common basic attentional capacity limitations (Cokely et al., 2018). In other words, peoples' expertise in themselves and their ability to use personally relevant and affectively charged heuristics supports superior decision making.

This is often because decision making occurs in high-stakes and *uncertain* situations, and contrary to historical notions of decision-making that highlight rational optimization techniques, recent research indicates that superior decision-making is primarily a function of specialized knowledge and integrated long-term memory representations, that inform adaptive heuristic use consistent with one's goals and values (i.e., *representative understanding*; See Cokely et al., 2018; Ghazal, 2014). Said differently, even though people have limited time and resources (cognitive and otherwise – i.e., bounded rationality), they are able to use adaptive heuristics to make superior, personally rational, decisions (see Skilled Decision theory, Cokely et al., 2018; Gigerenzer & Selten, 2002; Simon, 1956; 1990).

Implications and Future Research

To the extent that the present models hold and are representative of the typical ecology, there are a number of important implications and notions for future research. First, as the strongest single predictor of general decision making skill, numeracy could also be seen as a metric for decision making vulnerability. However, in order to create a vulnerability index, we need to refine current instruments, build norms, and develop adaptive measures of decision making skill. Future research will focus on creating simple yet robust psychometric instruments that facilitate the assessment, prediction and modeling of decision vulnerability among diverse individuals, and across wide-ranging, high-stakes decisions.

While the present research aims to model how these important factors are related, future research should also consider how these skills and abilities can be trained so that individuals are better able to understand information (i.e., build a representative understanding), and be more prepared to make informed decisions. Future research should also look at the possibilities of improving risk literacy via visual aids, graph literacy and numeracy training (Peters et al., 2017; Petrova, Garcia-Retamero, & Cokely, 2015; Schoenfeld, 1992; Ybarra et al., 2017). In tandem to better risk communications and decision aids, individuals will be empowered to make important decisions on their own and in accord with their personal values and goals. Creating and promoting new training interventions can create numerous ethical implications (e.g., tradeoffs of who to train what and when; Peters, 2012; Peters et al., 2017). However, we aim to address these issues and new technologies should be accessible to all via a web-based system that provides individuals an opportunity to take some of these verified scales and receive personalized feedback (e.g., RiskLiteracy.org).

Limitations

Even though there exists some statistical evidence to the contrary, decision making skill is clearly not a unidimensional trait. Therefore, only to the extent that the measures included in this study (e.g., ADMC, Heuristics & Biases, Ecological Decision Battery) actually capture the true nature of decision making skill, can the models represent the underlying causal mechanisms. Converging research indicates these scales are sufficient, and as such, the present study should provide evidence that numeracy likely mediates the relationship between intelligence and decision making skill. However, to the extent that this analysis missed important pieces of decision making skill and/or intelligence, the naïve models (presented in Figure 3) still demonstrate that numeracy is the best stand-alone predictor.

Contrary to some common criticism about the nature of statistical numeracy: statistical numeracy and decision making skill are not one in the same (Cokely et al., 2018). Rather, statistical numeracy is simply a strong and efficient test of decision making skill. Previous research on the efficacy of visual aids and graph literacy training on improving decision making quality (Peters, 2017; Ybarra et al., 2017; Woller-Carter, 2015), provides evidence that even though improving numeracy may lead to improvement in decision making skill, these two constructs are also *separable*. Though statistical numeracy and some decision making tasks may share identical elements (e.g., interpreting risk information in a medical decision problem is similar to solving a conditional probability problem), it is important to note the ecological validity of these decision items – often realistic, high stakes decisions require reckoning with probabilities and uncertainty. To that end, according to Skilled Decision Theory, it is not being numerate that makes someone a better decision maker, but rather that those with more decision making skill are better able to solve numeracy problems. Taken together, this research provides converging evidence for the causal mechanisms between statistical numeracy, visual aids, and decision making skill (see Skilled Decision Theory for a review; Cokely et al., 2018).

Despite the contributions made by the present effort, there are other limitations that should be mentioned. The data provided here is correlational, and the study did not impose any experimental manipulation. Therefore, though I suggest a path analysis, the structure imposed here is fully based on previous research and analyses.

Further, while the numeracy tests were administered in the lab, the decision tasks and intelligence measures were collected at-home in an online setting. Unfortunately, it is possible that given the unstructured at-home environment, students may have been less focused, and may not have demonstrated similar effort – leading to more noise and variance in the at-home measures. That being said, real life decisions are not made in a vacuum, and it is equally possible that the at-home environment was more representative of ecological decision making.

Finally, though demographic data was collected, the present analyses do not include any gender or demographic differences, and there could be many important implications for hiring and human resources. Many hiring decisions are made on the basis of historical views of cognitive abilities (e.g., fluid and crystallized intelligence). However, to the extent that the present analyses are representative of the typical ecology, future decisions should perhaps instead be made on the basis of decision making skill.

Conclusions

Extensive evidence demonstrates that quantitative skills are among the most influential educational variables associated with economic prosperity in industrialized countries (Hunt & Wittmann, 2008; Hanushek & Woessmann, 2010). The results presented here may start to tell us why. Contrary to popular convention, it may not be about making more scientists or better experts, but rather about making people who are more intelligent – in the way that truly matters – in decision making skill.

As a researcher, the results presented here demonstrate that if you have limited time and resources and want to predict general mental ability, you should use statistical numeracy. However, in many ways, the history of intelligence is the history of discrimination. And as a citizen of the world, the present results tell us that general decision making skill does *not* require high levels of basic cognitive capacities (i.e., fluid intelligence). While many researchers have assumed the opposite – that good decision making requires great intelligence – and have thusly worked to determine *who* is intelligent "enough" (Clynes, 2016; Herrnstein & Murray, 1994; Plomin & von Stumm, 2018) as well as who should be *allowed* to make autonomous decisions (Thaler

& Sunstein, 1999), this finding is empowering because the converging evidence demonstrates exactly the opposite. With proper training and access to resources, nearly anyone can have the power to make informed decisions, in accord with their beliefs, values, and goals.

References

Aczel, B., Bago, B., Szollosi, A., Foldes, A., & Lukacs, B. (2015). Measuring individual differences in decision biases: Methodological considerations. *Frontiers in psychology, 6*, 1770.

Allan, J. N., Ripberger, J. T., Ybarra, V. T., & Cokely, E. T. (2017). The Oklahoma Warning Awareness Scale: A Psychometric Analysis of a Brief Self-Report Survey Instrument. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 61, No. 1, pp. 1203-1207). Sage CA: Los Angeles, CA: SAGE Publications.

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

Banks, J., O'Dea, C., & Oldfield, Z. (2010). Cognitive function, numeracy and retirement saving trajectories. *The Economic Journal*, *120*(548).

Baron, J. (1985). *Rationality and intelligence*. Cambridge University Press.

Baron, J. (2008). *Thinking and deciding (4th edn.).* Cambridge University Press.

Bors, D. A., & Stokes, T. L. (1998). Raven's Advanced Progressive Matrices: Norms for first-year university students and the development of a short form. *Educational and Psychological Measurement*, *58*(3), 382-398.

Bruine de Bruin, W., Parker, A. M., & Fischhoff, B. (2007). Individual differences in adult decision-making competence. *Journal of personality and social psychology*, *92*(5), 938.

Brunner, M., Nagy, G., & Wilhelm, O. (2012). A tutorial on hierarchically structured constructs. *Journal of personality*, *80*(4), 796-846.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review, 62*, 193–217.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* Cambridge University Press.

Cattell, R. B. (1973). *Culture-fair intelligence test*. Institute for personality and ability testing.

Clynes, T. (2016). How to raise a genius: lessons from a 45-year study of super-smart children. *Nature News*, *537*(7619), 152.

Cokely, E. T., Feltz, A., Ghazal, S., Allan, J. N., Petrova, D., & Garcia-Retamero, R., (2018). Decision making skill: From intelligence to numeracy and expertise. In K. A. Ericsson, R. R. Hoffman, A. Kozbelt, & A. M. Williams (2nd Eds.), *Cambridge Handbook of Expertise and Expert Performance.* New York, NY: Cambridge University Press.

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

Cokely, E. T., & Kelley, C. M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgment and Decision Making*, *4*(1), 20.

Cokely, E. T., Kelley, C. M., & Gilchrist, A. L. (2006). Sources of individual differences in working memory: Contributions of strategy to capacity. *Psychonomic Bulletin & Review*, *13*(6), 991-997.

Del Missier, F., Mäntylä, T., & Bruine de Bruin, W. (2010). Executive functions in decision making: An individual differences approach. *Thinking & Reasoning, 16*(2), 69-97.

Del Missier, F., Mäntylä, T., & Bruin, W. B. (2012). Decision-making competence, executive functioning, and general cognitive abilities. *Journal of Behavioral Decision Making*, *25*(4), 331-351.

Del Missier, F., Mäntylä, T., Hansson, P., Bruine de Bruin, W., Parker, A. M., & Nilsson, L.-G. (2013). The Multifold Relationship Between Memory and Decision Making: An Individual-differences Study. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 39*(5), 1344–1364.

Dhami, M. K., Hertwig, R., & Hoffrage, U. (2004). The role of representative design in an ecological approach to cognition. *Psychological bulletin*, *130*(6), 959.

Dhami, M. K., Schlottmann, A., & Waldmann, M. R. (Eds.). (2012). *Judgment and decision making as a skill: Learning, development and evolution*. Cambridge University Press.

Edwards, W. (1954). The theory of decision making. *Psychological Bulletin, 51*, 380–417.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological review*, *100*(3), 363.

Fagerlin, A., Zikmund-Fisher, B. J., Ubel, P. A., Jankovic, A., Derry, H. A., & Smith, D. M. (2007). Measuring numeracy without a math test: development of the Subjective Numeracy Scale. *Medical Decision Making*, *27*(5), 672-680.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic perspectives, 19*(4), 25-42.

Fox, M. C., & Mitchum, A. L. (2013). A knowledge- based theory of rising scores on “culture-free” tests. *Journal of Experimental Psychology: General, 142*(3), 979–1000.

Gal, I. (2003). Teaching for statistical literacy and services of statistics agencies. *American Statistician, 57*, 80–84.

Garcia-Retamero, R., & Cokely, E. T. (2013). Communicating health risks with visual aids. *Current Directions in Psychological Science*, *22*(5), 392-399.

Garcia-Retamero, R., & Cokely, E. T. (2014). The influence of skills, message frame, and visual aids on prevention of sexually transmitted diseases. *Journal of Behavioral Decision Making*, *27*(2), 179-189.

Garcia-Retamero, R., & Cokely, E. T. (2017). Designing visual aids that promote risk literacy: A systematic review of health research and evidence-based design heuristics. *Human factors*, *59*(4), 582-627.

Ghazal, S. (2014). *Component numeracy skills and decision making.* PhD Dissertation, Michigan Technological University.

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

Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M., & Woloshin, S. (2007). Helping doctors and patients make sense of health statistics. *Psychological Science in the Public Interest,* 8(2), 53-96.

Gigerenzer, G., & Selten, R. (Eds.). (2002). *Bounded rationality: The adaptive toolbox*. MIT press.

Gigerenzer, G., Todd, P. M., & ABC Research Group, T. (1999). Simple heuristics that make us smart.

Ginsburg, L., Manly, M., & Schmitt, M. J. (2006). The components of numeracy. NCSALL Occasional Paper. Boston, MA: National Center for the Study of Adult Learning and Literacy.

Gould, S. J. (1996). *The mismeasure of man*. WW Norton & Company.

Hanushek, E. A., & Woessmann, L. (2010). The economics of international differences in educational achievement (No. 15949). *National Bureau of Economic Research*.

Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and Class Structure in American Life,* New York.

Huff, D. (1954). *How to Lie with Statistics*. New York: W. W.

Hunt, E., & Wittmann, W. (2008). National intelligence and national prosperity. *Intelligence*, *36*(1), 1-9.

Kline, R. B. (2011). *Principles and Practices of Structural Equation Modeling*. New York: The Guilford Press.

Kutner, M., Greenburg, E., Jin, Y., & Paulsen, C. (2006). The health literacy of America's adults: Results from the 2003 National Assessment of Adult Literacy. NCES 2006-483. Washington, DC: National Center for Education Statistics.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?!. *Intelligence*, *14*(4), 389-433.

Lipkus, I. M., & Peters, E. (2009). Understanding the role of numeracy in health: Proposed theoretical framework and practical insights. *Health Education & Behavior*, *36*(6), 1065-1081.

Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical decision making*, *21*(1), 37-44.

Lubinski, D. (2016). From Terman to today: A century of findings on intellectual precocity. *Review of Educational Research*, *86*(4), 900-944.

Lusardi, A. (2012). *Numeracy, financial literacy, and financial decision-making* (No. w17821). National Bureau of Economic Research.

Marsh, H. W., & Hau, K. T. (1999). Confirmatory factor analysis: Strategies for small sample sizes. *Statistical strategies for small sample research*, *1*, 251-284.

Matthews, T. D., & Lassiter, K. S. (2007). What does the wonderlic personnel test measure?. *Psychological reports*, *100*(3), 707-712.

McCabe, D. P., Roediger III, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: evidence for a common executive attention construct. *Neuropsychology*, *24*(2), 222.

Neisser, U., Boodoo, G., Bouchard Jr, T. J., Boykin, A. W., Brody, N., Ceci, S. J., ... & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American psychologist*, *51*(2), 77.

Nelson, H. E., & McKenna, P. A. T. (1975). The use of current reading ability in the assessment of dementia. *British Journal of Clinical Psychology*, *14*(3), 259-267.

Nelson, H. E., & Willison, J. (1991). *National adult reading test (NART)*. Windsor: Nfer-Nelson.

Nelson, W., Reyna, V. F., Fagerlin, A., Lipkus, I., & Peters, E. (2008). Clinical implications of numeracy: theory and practice. *Annals of behavioral medicine*, *35*(3), 261-274.

Parker, A. M., Bruine de Bruin, W., & Fischhoff, B. (2015). Negative decision outcomes are more common among people with lower decision-making competence: an item-level analysis of the Decision Outcome Inventory (DOI). *Frontiers in psychology*, *6*, 363.

Paulos, J. A. (1988). *Innumeracy: Mathematical illiteracy and its consequences*. Macmillan.

Peters, E. (2012). Beyond comprehension: The role of numeracy in judgments and decisions. *Current Directions in Psychological Science*, *21*(1), 31-35.

Peters, E. (2017). Educating good decisions. *Behavioural Public Policy, 1*(2), 162-176.

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

Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological science*, *17*(5), 407-413.

Peters, E., Shoots-Reinhard, B., Tompkins, M. K., Schley, D., Meilleur, L., Sinayev, A., Tusler, M., Wagner, L., & Crocker, J. (2017). Improving numeracy through values affirmation enhances decision and STEM outcomes. *PloS ONE*, *12*(7).

Petrova, D., Garcia-Retamero, R., Catena, A., Cokely, E., Heredia Carrasco, A., Arrebola Moreno, A., & Ramírez Hernández, J. A. (2016). Numeracy predicts risk of pre-hospital decision delay: A retrospective study of acute coronary syndrome survival. *Annals of Behavioral Medicine, 51*(2), 292-306.

Petrova, D., Garcia-Retamero, R., & Cokely, E. T. (2015). Understanding the harms and benefits of cancer screening: A model of factors that shape informed decision making. *Medical Decision Making*, *35*(7), 847-858.

Pleskac, T. J., & Hertwig, R. (2014). Ecologically rational choice and the structure of the environment. *Journal of Experimental Psychology: General*, 143(5), 2000-2019.

Plomin, R., & von Stumm, S. (2018). The new genetics of intelligence. *Nature Reviews Genetics*.

Raven, J.C., Court, J.H., & Raven, J. (1988). *Manual for Ravens' Progressive Matrices and Vocabulary Scales* (Section 4). London: H. K. Lewis.

Revelle, W. (2018). Package 'psych'. *Retrieved April 6*, 2018.

Reyna, V. F., Nelson, W. L., Han, P. K., & Dieckmann, N. F. (2009). How Numeracy Influences Risk Comprehension and Medical Decision Making. *Psychological Bulletin, 135*(6), 943–973.

Rosseel, Y., Oberski, D., Byrnes, J., Vanbrabant, L., Savalei, V., Merkle, E., ... & Chow, M. (2017). Package 'lavaan'. *Retrieved April 6*, 2018.

Ruch, F. L., & Ruch, W. W. (1963). *Employee aptitude survey: Technical report*. Los Angeles: Psychological Services.

Savage, L. J. (1954). *The foundations of Statistics*.

Schapira, M. M., Walker, C. M., & Sedivy, S. K. (2009). Evaluating existing measures of health numeracy using item response theory. *Patient Educational Counseling*, *75*, 308–314.

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. In D. Grouws (Ed.) *Handbook of research on mathematics teaching and learning*. New York: Macmillan

Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H. G. (1997). The role of numeracy in understanding the benefit of screening mammography. *Annals of internal medicine*, *127*(11), 966-972.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review, 63*, 129–138.

Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology, 41*, 1–20.

Skagerlund, K., Lind, T., Strömbäck, C., Tinghög, G., & Västfjäll, D. (2018). Financial literacy and the role of numeracy–How individuals' attitude and affinity with numbers influence financial literacy. *Journal of Behavioral and Experimental Economics*, *74*, 18-25.

Spearman, C. (1923). *The nature of 'intelligence' and the principles of cognition (2nd ed.).* London: Macmillan.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate?. *Behavioral and brain sciences*, *23*(5), 645-665.

Stanovich, K. E., West, R. F., & Toplak, M. E. (2016). *The rationality quotient: Toward a test of rational thinking*. MIT Press.

Szaszi, B., Palfi, B., Szollosi, A., Kieslich, P. J., & Aczel, B. (2018). Thinking dynamics and individual differences: Mouse-tracking analysis of the denominator neglect task. *Judgment and decision making*, *13*(1), 23.

Thaler, R. H., & Sunstein, C. R. (1999). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT Yales University Press.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & cognition*, *39*(7), 1275.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124-1131.

van der Maas, H. L., Dolan, C. V., Grasman, R. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. (2006). A dynamical model of general intelligence: the positive manifold of intelligence by mutualism. *Psychological review*, *113*(4), 842.

Weller, J. A., Dieckmann, N. F., Tusler, M., Mertz, C. K., Burns, W. J., & Peters, E. (2013). Development and testing of an abbreviated numeracy scale: A Rasch analysis approach. *Journal of Behavioral Decision Making*, *26*(2), 198-212.

Woller-Carter, M. (2015). *Development of the intelligent graphs for everyday risky decisions tutor* (Doctoral dissertation, Michigan Technological University). Michigan Technological University.

Wonderlic. (2018). The Official Wonderlic Test Online. Retrieved April 06, 2018, from https://www.wonderlic.com/

Wonderlic, E. F., & Hovland, C. I. (1939). The Personnel Test: A Restandardized abridgment of the Otis S-A test for business and industrial use. *Journal of Applied Psychology, 23*(6), 685-702.

Ybarra, V. T., Cokely, E. T., Adams, C., Woller-Carter, M., Allan, J. N., Feltz, A., & Garcia-Retamero, R. (2017). Training Graph Literacy: Developing the RiskLiteracy.org Outreach Platform. *Cogsci Proceedings 2017,* 3566-3571.

Appendix A: Berlin Numeracy Tests

Berlin Numeracy Test Components (BNT-C): 4 Sections, 9 Questions Each

Part 1: Operations - 9 Questions
A school is having a field trip and many parents are going on the fieldtrip with the children. What is the child to parent ratio if there are 20 children and 5 parents?
- ❍ 2 children for every one parent
- ❍ 20 children for every 1 parent
- ❍ 1 child to every 5 parents
- ❍ 5 children to every 1 parent
- ❍ 4 children to every 1 parent

The mileage meter of an old motorcycle is malfunctioning and registers only 3 miles for every 4 miles driven. If the meter indicates 54 miles, how many miles has the motorcycle actually driven?
- ❍ 162
- ❍ 108
- ❍ 72
- ❍ 53
- ❍ 36

Kristie has a collection of adventure, comic, and romantic novels. If the ratio of adventure novels to comic novels is 5 to 1 and the ratio of comic to romantic novels is 5 to 3, what is the ratio of adventure to romantic novels?
- ❍ 5:3
- ❍ 10:5
- ❍ 6:3
- ❍ 20:3
- ❍ 25:3

Imagine that goods imported into a country increased by 40% and exports decreased by 30% during a certain year. What was the ratio of imports to exports at the end of the year compared to the beginning of the year?
- ❍ 1/2
- ❍ 3/2
- ❍ 4/3
- ❍ 2/1
- ❍ 1

If a sack of dried dog food feeds 4 dogs or 5 puppies for one week, then 5 sacks of the food will feed 15 puppies and how many dogs in one week?

The odometer of a new automobile functions improperly and registers only 2 miles for every 3 miles driven. If the odometer indicates 48 miles, how many miles has the automobile actually been driven?

- 144
- 72
- 64
- 32
- 24

Helpers are needed to prepare for the fete. Each helper can make either 2 large cakes per hour or 35 small cakes per hour. The kitchen is available for 3 hours and 20 large cakes and 700 small cakes are needed. How many helpers are required?

- 10
- 15
- 20
- 25
- 30

If United States imports increased 20 percent and exports decreased 10 percent during a certain year, the ratio of imports to exports at the end of the year was how many times the ratio at the beginning of the year?

- 12/11
- 4/3
- 11/8
- 3/2
- 2

n and p are integers greater than 1.
5n is the square of a number.
75np is the cube of a number.
The smallest value for n + p is

- 14
- 18
- 20
- 30
- 50

Part 2: Probability - 9 Questions

People often roll dice when playing games. Most dice have 6 sides and each side has a different number on it ranging from 1-6. If you rolled one of the dice, on average what is the probability that it would land on 5?

- ❍ 1 time out of 6 rolls of the dice
- ❍ 5 times out of 6 rolls of the dice
- ❍ 1 time out of 2 rolls of the dice
- ❍ 1 out of 5 rolls of the dice
- ❍ 6 out of 1 roll of the dice

Imagine that you are throwing 2 regular 6-sided dice up in the air. If each side has a different number on it ranging from 1-6, on average what is the probability that both of them land on even numbers?

- ❍ 1 out of 36 rolls of the dice
- ❍ 3 out of 6 rolls of the dice
- ❍ 1 out of 4 rolls of the dice
- ❍ 2 out of 6 rolls of the dice
- ❍ 2 out of 36 rolls of the dice

Imagine that the probability of a child getting sunburned at the beach is 65% while the probability of an adult getting sunburned at the beach is 15%. If there were 300 people who spent a day at the beach, and 60% of the people were children, how many people are likely to get a sunburn?

- ❍ About 195
- ❍ About 150
- ❍ About 135
- ❍ About 80
- ❍ About 64

Suppose you are taking an 8 question multiple choice test and each question has 4 options. Imagine that you don't know anything about the test and so you guess without reading the questions. What's the probability that you would get 100% correct on this test just by chance alone?

- 1/4
- 1/8
- 1/4096
- 1/16384
- 1/65536

Imagine that you are throwing 6 dice up in the air. What is the probability that all of them would land on even numbers?

- 1/432
- 3/216
- 1/64
- 3/6
- 1/21

Imagine you are drawing a picture, and are missing 2 spots you want to color. There are 7 colors to choose from. What's the probability that both spots end up colored orange?

- 1/49
- 2/49
- 1/7
- 2/7
- 6/7

Imagine you are throwing 8 dice up in the air. What's the probability that half will land on an even number, while the other half land on 1?

- 1/10368
- 1/20736
- 1/432
- 1/1728
- 1/6

Phil is holding 4 cards in his hand: 8 of clubs, 5 of hearts, king of hearts, and ace of diamonds. If he places them on a table in random order, what is the probability that the first and last cards will both be hearts?

- 1/2
- 1/3
- 1/4
- 1/6
- 1/8

n is an integer chosen at random from the set {5, 7, 9, 11}
p is chosen at random from the set {2, 6, 10, 14, 18}
What is the probability that n + p = 23?

- 0.1
- 0.2
- 0.25
- 0.3
- 0.4

Part 3: Geometry - 9 Questions

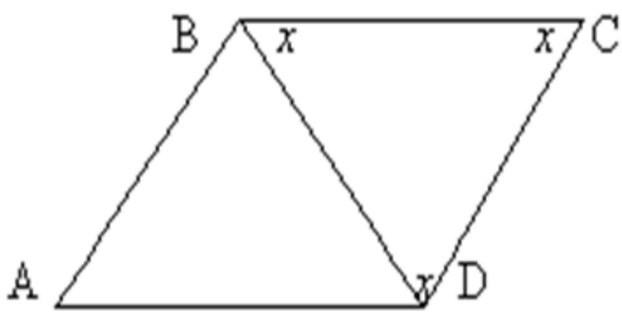
Imagine there is a rectangle that has an area of 20 square meters. If its length is 4 meters, what is its width?

- 10
- 5
- 4
- 3
- None of the above

The slope of a line through points P (1, 4) and Q (-5, X) is 1/3. What is the value of X?

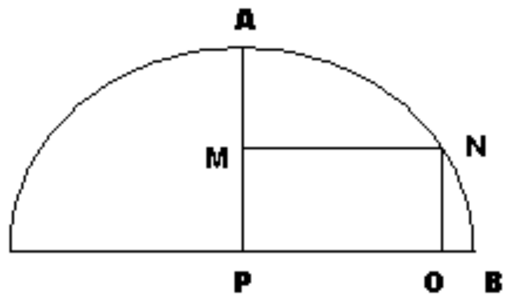
- 1
- 2
- 3
- 4
- 5

The perimeter of the parallelogram ABCD is 16. Each angle in the triangle BCD is equal. What is the length of the side AD?



- $2\sqrt{16}$
- $\sqrt{16}$
- 2
- $2\sqrt{2}$
- 4

APB is a quarter circle. MNOP is a rectangle with sides MN = 4 and MP = 3. What is the length of the arc ANB?



- ❍ 2.5π
- ❍ 5π
- ❍ 10π
- ❍ 7
- ❍ 10

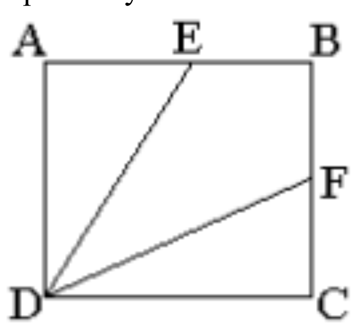
The slope of a line through points P (1, 1) and Q (k, 7) is 3/2. What is the value of k?

- ❍ 4
- ❍ 5
- ❍ 6
- ❍ 7
- ❍ 8

A, B, C, and D are points on a line, with D the midpoint of BC. The lengths of AB, AC and BC are 10, 2, and 12, respectively. What is the length of AD?

- ❍ 2
- ❍ 4
- ❍ 6
- ❍ 10
- ❍ 12

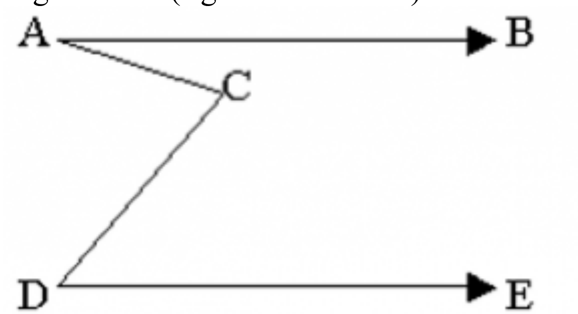
ABCD is a square of side 3, and E and F are the mid points of sides AB and BC respectively. What is the area of the quadrilateral EBFD?



- ❍ 2.25
- ❍ 3
- ❍ 4

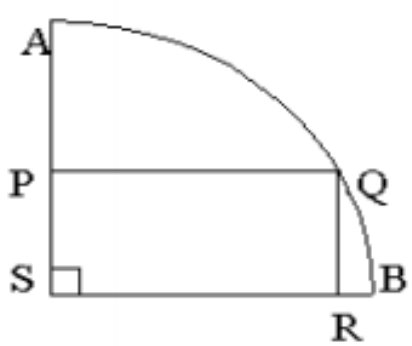
❍ 4.5
❍ 6

AB and DE are parallel. Angle BAC = 30, angle CDE = 50. What is the measure of angle ACD? (figure not to scale)



❍ 100
❍ 90
❍ 80
❍ 70
❍ cannot be determined from the information

ASB is a quarter circle. PQRS is a rectangle with sides PQ = 8 and PS = 6. What is the length of the arc AQB?



❍ 5π
❍ 10π
❍ 25
❍ 14
❍ 28

Part 4: Algebra - 9 Questions

If 40/X = 5, what is the value of X?
- ❍ 5
- ❍ 8
- ❍ 10
- ❍ 15
- ❍ None of the above

Solve for b in this equation: a + b = d
- ❍ b = a - d
- ❍ b = a + d
- ❍ b = ad
- ❍ b = d - a
- ❍ None of the above

What is the value of a - c + b in this equation? $(2x - 1)(x + 3) = ax^2 + bx + c$.
- ❍ 5
- ❍ 8
- ❍ 9
- ❍ 10
- ❍ 11

What is the simplified result of following the three steps below performed in order?
1-Add 2a to 3b
2-Multiply the sum by 4
3-Subtract a + b from the product
- ❍ 7a + 13b
- ❍ 7a + 11b
- ❍ 7a + 7b
- ❍ 8a + 12b
- ❍ a + 2b

If the equation of a line p in the coordinate plane is y = 3x + 2, what is the equation of line q which is a reflection of line p in the x-axis?
- ❍ y = -3x + 2
- ❍ y = -3x - 2
- ❍ y = 3x - 2
- ❍ y = -1/3x - 5
- ❍ y = -1/3x + 5

$(3x + 2)(2x - 5) = ax^2 + kx + n$. What is the value of $a - n + k$?

- 5
- 8
- 9
- 10
- 11

If $f(x) = x^2 – 3$, where x is an integer, which of the following could be a value of f(x)?

I 6
II 0
III -6

- I only
- I and II only
- II and III only
- I and III only
- I, II and III

Six years ago Anita was P times as old as Ben was. If Anita is now 17 years old, how old is Ben now in terms of P?

- 11/P + 6
- P/11 + 6
- 17 - P/6
- 17/P
- 11.5P

If x / y is an integer, which of the following statements must be true?

- both x and y are integers
- x is an integer
- either x or y is negative
- y / x is an integer
- x = ny where n is an integer

Berlin Numeracy Test (BNT-Schwartz; Cokely et al., 2012)

You will now be asked to solve a few problems. Please note that you are allowed to enter numbers that include up to 2 decimal points (for example, 1.11). You are also welcome to use a calculator to help solve these problems.

Imagine that we flip a fair coin 1,000 times. What is your best guess about how many times the coin would come up heads in 1,000 flips?

In the BIG BUCKS LOTTERY, the chance of winning a $10 prize is 1%. What is your best guess about how many people would win a $10 prize if 1,000 people each buy a single ticket to BIG BUCKS?

____

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

______________________ percent

Out of 1,000 people in a small town 500 are members of a choir. Out of these 500 members in a choir 100 are men. Out of the 500 inhabitants that are not in a choir 300 are men. What is the probability that a randomly drawn man is a member of the choir? Please indicate the probability as a percent.

______________________ percent

Imagine we are throwing a five-sided die 50 times. On average, out of these 50 throws how many times would this five-sided die show an odd number (1, 3 or 5)?

____

Imagine we are throwing a loaded die (6 sides). The probability that the die shows a 6 is twice as high as the probability of each of the other numbers. On average, out of these 70 throws how many times would the die show the number 6?

____

In a forest, 20% of the mushrooms are red, 50% are brown, and 30% are white. A red mushroom is poisonous with a probability of 20%. A mushroom that is not red is poisonous with a probability of 5%. What is the probability that a poisonous mushroom in the forest is red? Please indicate the probability as a percent.

______________________ percent

Appendix B: Heuristics and Biases Test

(Toplak et al., 2011)

1. The Caldwells had long ago decided that when it was time to replace their car they would get what they called "one of those solid, safety-conscious, built-to-last Swedish cars"-either a Volvo or a Saab. As luck would have it, their old car gave up the ghost on the last day of the closeout sale for the model year both for the Volvo and for the Saab. The model year was changing for both cars and the dollar had recently dropped substantially against European currencies; therefore, if they waited to buy either a Volvo or a Saab, it would cost them substantially more-about $1200. They quickly got out their Consumer Reports where they found that the consensus of the experts was that both cars were very sound mechanically, although the Volvo was felt to be slightly superior on some dimensions. They also found that the readers of Consumer Reports who owned a Volvo reported having somewhat fewer mechanical problems than owners of Saabs. They were about to go and strike a bargain with the Volvo dealer when Mr. Caldwell remembered that they had two friends who owned a Saab and one who owned a Volvo. Mr. Caldwell called up the friends. Both Saab owners reported having had a few mechanical problems but nothing major. The Volvo owner exploded when asked how he liked his car. "First that fancy fuel injection computer thing went out: $250 bucks. Next I started having trouble with the rear end. Had to replace it. Then the transmission and the clutch. I finally sold it after 3 years for junk."

   Given that the Caldwells are going to buy either a Volvo or a Saab today, in order to save $1200, which do you think they should buy?

   ❍ Volvo

   ❍ Saab

2. A certain town is served by two hospitals. In the larger hospital about 45 babies are born each day, and in the smaller hospital about 15 babies are born each day. As you know, about 50 percent of all babies are boys. However, the exact percentage varies from day to day. Sometimes it may be higher than 50 percent, sometimes lower. For a period of 1 year, each hospital recorded the days on which more than 60

percent of the babies born were boys. Which hospital do you think recorded more such days?

- ❍ The larger hospital
- ❍ The smaller hospital
- ❍ About the same (that is, within 5 percent of each other)

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

- ❍ Playing to 9 points
- ❍ Playing to 15 points
- ❍ Both scoring systems would give player A a better chance of winning

4. After the first 2 weeks of the major league baseball season, newspapers begin to print the top 10 batting averages. Typically, after 2 weeks, the leading batter often has an average of about .450. However, no batter in major league history has ever averaged .450 at the end of the season. Why do you think this is? Circle one:

- ❍ When a batter is known to be hitting for a high average, pitchers bear down more when they pitch to him.
- ❍ Pitchers tend to get better over the course of a season, as they get more in shape. As pitchers improve, they are more likely to strike out batters, so batters' averages go down.
- ❍ A player's high average at the beginning of the season may be just luck. The longer season provides a more realistic test of a batter's skill.
- ❍ A batter who has such a hot streak at the beginning of the season is under a lot of stress to maintain his performance record. Such stress adversely affects his playing.
- ❍ When a batter is known to be hitting for a high average, he stops getting good pitches to hit. Instead, pitchers "play the corners" of the plate because they don't mind walking him.

5. When playing slot machines, people win something about 1 in every 10 times. Julie, however, has just won on her first three plays. What are her chances of winning the next time she plays?
Please give your answer in the format: "____ out of ____."

____

6. Imagine that we are tossing a fair coin (a coin that has a 50/50 chance of coming up heads or tails) and it has just come up heads 5 times in a row.
   For the 6th toss do you think that:
   - ❍ It is more likely that tails will come up than heads.
   - ❍ It is more likely that heads will come up than tails.
   - ❍ Heads and tails are equally probable on the sixth toss.

7. Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.
   Is it more likely that...
   - ❍ Linda is a bank-teller
   - ❍ Linda is a bank-teller and a feminist

8. A doctor had been working on a cure for a mysterious disease. Finally, he created a drug that he thinks will cure people of the disease. Before he can begin to use it regularly, he has to test the drug. He selected 300 people who had the disease and gave them the drug to see what happened. He selected 100 people who had the disease and did not give them the drug in order to see what happened. The table below indicates what the outcome of the experiment was:

   Judge whether this treatment was positively or negatively associated with the cure for this disease by selecting a number from the scale ranging from −10 (strong negative association) to +10 (strong positive association).

- ❍ -10
- ❍ -8
- ❍ -6
- ❍ -4
- ❍ -2
- ❍ 0
- ❍ 2
- ❍ 4
- ❍ 6
- ❍ 8
- ❍ 10

9. The city of Middleopolis has had an unpopular police chief for a year and a half. He is a political appointee who is a crony of the mayor, and he had little previous experience in police administration when he was appointed. The mayor has recently defended the chief in public, announcing that in the time since he took office, crime rates decreased by 12%. Which of the following pieces of evidence would most deflate the mayor's claim that his chief is competent?

   - ❍ The crime rates of the two cities closest to Middleopolis in location and size have decreased by 18% in the same period.
   - ❍ An independent survey of the citizens of Middleopolis shows that 40% more crime is reported by respondents in the survey than is reported in police records
   - ❍ Common sense indicates that there is little a police chief can do to lower crime rates. These are for the most part due to social and economic conditions beyond the control of officials
   - ❍ The police chief has been discovered to have business contacts with people who are known to be involved in organized crime

10. Imagine yourself meeting David Maxwell. Your task is to assess the probability that he is a university professor based on some information that you will be given. This will be done in two steps. At each step you will get some information that you may or may not find useful in making your assessment. After each piece of information you will be asked to assess the probability that David Maxwell is a university professor. In doing so, consider all the information you have received to that point if

you consider it to be relevant. Your probability assessments should be numbers between 0 and 1 that express your degree of belief. 1 means "I am absolutely certain that he is a university professor." .65 means "The chances are 65 out of 100 that he is a university professor," and so forth. You can use any number between 0 and 1, for example, .15, .95, etc.

Step One: You are told that David Maxwell attended a party in which 25 male university professors and 75 male business executives took part, 100 people all together.
Question: What do you think the probability is that David Maxwell is a university professor?

____

Step Two: You are told that David Maxwell is a member of the Bears Club. 70% of the male university professors at the above-mentioned party were members of the Bears Club, and 90% of the male business executives at the party were members of the Bears Club.
Question: What do you think the probability is that David Maxwell is a university professor?

____

11. Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimate of the consequences of the programs are as follows:
If Program A is adopted, 200 people will be saved.
If Program B is adopted, there is 1/3 probability that 600 people will be saved, and 2/3 probability that no people will be saved.
Which of the two programs would you favor?

❍ Program A

❍ Program B

12. Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimate of the consequences of the programs are as follows:
If Program C is adopted 400 people will die.
If Program D is adopted there is 1/3 probability that nobody will die, and 2/3

probability that 600 people will die.
Which of the two programs would you favor?

- ❍ Program C
- ❍ Program D

13. Assume that you are presented with two trays of black and white marbles: a large tray that contains 100 marbles and a small tray that contains 10 marbles. The marbles are spread in a single layer on each tray. You must draw out one marble (without peeking, of course) from either tray. If you draw a black marble, you win $2. Consider a condition in which the small tray contains 1 black marble and 9 white marbles, and the large tray contains 8 black marbles and 92 white marbles. From which tray would you prefer to select a marble in a real situation?

- ❍ Small Tray
- ❍ Large Tray

14. A die with 4 red faces and 2 green faces will be rolled 60 times. Before each roll you will be asked to predict which color (red or green) will show up once the die is rolled. You will be given one dollar for each correct prediction. Assume that you want to make as much money as possible. What strategy would you use in order to make as much money as possible by making the most correct predictions?

- ❍ Strategy A: Go by intuition, switching when there has been too many of one color or the other.
- ❍ Strategy B: Predict the more likely color (red) on most of the rolls but occasionally, after a long run of reds, predict a green.
- ❍ Strategy C: Make predictions according to the frequency of occurrence (4 of 6 for red and 2 of 6 for green). That is, predict twice as many reds as greens.
- ❍ Strategy D: Predict the more likely color (red) on all of the 60 rolls.
- ❍ Strategy E: Predict more red than green, but switching back and forth depending upon "runs" of one color or the other. Which Strategy is best?

15. You are staying in a hotel room on vacation. You paid $6.95 to see a movie on pay TV. After 5 minutes you are bored and the movie seems pretty bad. Would you continue to watch the movie or not?

- ❍ Continue to watch
- ❍ Turn it off

16. You are staying in a hotel room on vacation. You turn on the TV and there is a movie on. After 5 minutes you are bored and the movie seems pretty bad. Would you continue to watch the movie or not?

- ❍ Continue to watch
- ❍ Turn it off

17. A 55 year old man had a heart condition. He had to stop working because of chest pain. He enjoyed his work and did not want to stop. His pain also interfered with other things, such as travel and recreation.
A type of bypass operation would relieve his pain and increase his life expectancy from age 65 to age 70. However, 8% of the people who have this operation die from the operation itself.
His physician decided to go ahead with the operation. The operation succeeded.
Evaluate the physician's decision (on a scale from 1 (incorrect, a very bad decision) to 7 (clearly correct, an excellent decision) to go ahead with the operation.

- ❍ 1 (incorrect, a very bad decision)
- ❍ 2
- ❍ 3
- ❍ 4 (neither correct nor incorrect)
- ❍ 5
- ❍ 6
- ❍ 7 (clearly correct, an excellent decision)

18. A 60-year-old man was having trouble walking due to a hip condition. He had to stop most activity (such as work and enjoyment) as the pain was unbearable with excessive movement.
Arthroplasty (hip replacement) would relieve his pain and increase his life expectancy from 65 to age 75. However, 2% of the people who have this operation die from the operation itself.
His physician decided to go ahead with the operation. Unfortunately, the patient died during the operation.

Evaluate the physician's decision (on a scale from 1 (incorrect, a very bad decision) to 7 (clearly correct, an excellent decision) to go ahead with the operation.

❍ 1 (incorrect, a very bad decision)

❍ 2

❍ 3

❍ 4 (neither correct nor incorrect)

❍ 5

❍ 6

❍ 7 (clearly correct, an excellent decision)

Appendix C: General Decision Making Skill Assessment

Adult Decision Making Competence (Bruine de Bruin et al., 2007)

Please follow the links below for scale and item information.

http://www.sjdm.org/dmidi/Adult_-_Decision_Making_Competence.html

http://www.sjdm.org/dmidi/files/AdultDMCwithoutPathIndep.pdf

Ecological Decision Battery

Ecological Risk Literacy – Medical

Cervical cancer is very rare. 4 out of 100,000 women are affected by this cancer. The human papillomavirus (HPV) vaccine is federally approved and is being promoted as a method that helps reduce cervical cancer. Research studies suggest that the vaccine is 90 percent effective in preventing transmission of certain virus types. This conclusion is based on the results from a large international medical trial of 18,525 women aged 15-25, sponsored by the drug's manufacturer. 23 cases of the HPV virus were detected in the medical trial. Two of these cases were among the 9,258 women receiving the HPV vaccine, and 21 were among the 9,267 controls, who received a hepatitis A vaccine. The mean follow-up time was 14.8 months.

To what extent is the conclusion that "the vaccine was 90% effective" correct?
Please provide an answer in using the following 7 point scale.

- ❍ 1 - Completely confident that conclusion is right
- ❍ 2
- ❍ 3
- ❍ 4
- ❍ 5
- ❍ 6
- ❍ 7 - Completely confident that conclusion is wrong

What is the relative effectiveness of the vaccine? (express your answer as a percentage)

____

What is the absolute effectiveness of the vaccine? (express your answer as a percentage)

____

What are the chances that a woman gets cervical cancer after getting vaccinated? (express your answer as a percentage)

____

What did the women receive in the control group?

____

Ecological Risk Literacy – Financial

Imagine that you take out a $50,000 federal student loan to help pay for college. You are offered four possible repayment plans. The table below provides examples of the monthly repayments for each plan. Note: For the Graduated (10 years) plan, you would start by paying the minimum amount; the payment amount then increases every two years up to the maximum amount.
Look at the table carefully and answer the following questions.

Debt When Loan Enters Repayment	Standard (10 years)		Graduated (10 years)			Extended-Fixed (25 years)		Extended-Graduated (25 years)		
	Payment	Total Paid	Minimum Payment	Maximum Payment	Total Paid	Payment	Total Paid	Minimum Payment	Maximum Payment	Total Paid
$10,000	$115	$13,810	$66	$199	$14,860	-	-	-	-	-
$20,000	$230	$27,619	$133	$398	$29,720	-	-	-	-	-
$30,000	$345	$41,429	$199	$598	$44,580	-	-	-	-	-
$40,000	$460	$55,239	$266	$797	$59,439	$278	$83,289	$227	$397	$90,207
$50,000	$575	$69,048	$332	$996	$74,300	$347	$104,111	$283	$496	$112,762
$60,000	$690	$82,858	$398	$1,195	$89,160	$416	$124,933	$340	$595	$135,314
$70,000	$806	$96,667	$465	$1,393	$104,020	$486	$145,755	$397	$694	$157,865
$80,000	$921	$110,477	$531	$1,593	$118,880	$555	$166,577	$453	$793	$180,427
$90,000	$1,036	$124,287	$597	$1,791	$133,740	$625	$187,399	$510	$892	$202,980
$100,000	$1,151	$138,096	$664	$1,991	$148,600	$694	$208,222	$567	$992	$225,531

What is the total amount of interest payable on the Extended-Fixed (25 years) plan?

____

How confident are you in your previous answer?

- ❍ 0
- ❍ 1
- ❍ 2
- ❍ 3
- ❍ 4
- ❍ 5
- ❍ 6
- ❍ 7
- ❍ 8
- ❍ 9
- ❍ 10

Which option has the minimum interest payment (least expensive overall)?

____

What is the total interest paid in percentage if you have borrowed $50,000 and returned $69,048?

____

Assume someone has borrowed $50,000 for his studies, and he hopes to get a good job after his graduation in 5 years (when he will be able to pay more toward his debt). In this case which option should he choose?

- ❍ Standard
- ❍ Graduated
- ❍ Extended fixed
- ❍ Extended graduated

Prospect Evaluations – Expected Values & Choice Consistency in Lotteries
For the following questions, please indicate which of the two options you prefer:

- ❍ Lose $50
- ❍ 50% chance to lose $400

- ❍ Lose $120
- ❍ 5% chance to lose $1600

- ❍ Lose $200
- ❍ 1% chance to lose $3000

- ❍ Lose $275
- ❍ 20% chance to lose $900

- ❍ Lose $400
- ❍ 70% chance to lose $480

For the following questions, please indicate which of the two options you prefer:

- ❍ Gain $50
- ❍ 50% chance to win $400

- ❍ Gain $120
- ❍ 5% chance to win $1600

- ❍ Gain $200
- ❍ 1% chance to win $3000

- ❍ Gain $275
- ❍ 20% chance to win $900

- ❍ Gain $400
- ❍ 70% chance to win $480

For the following questions, please indicate which of the two options you prefer:

- ❍ $100 for sure
- ❍ 75% chance of $200

- ❍ Lose $100 for sure
- ❍ 60% chance to lose $250

- ❍ $500 for sure
- ❍ 15% chance of $1,000,000

- ❍ Lose $100 for sure
- ❍ 5% chance to lose $7000

For the following questions, please indicate which of the two options you prefer:

- ❍ 25% chance to win $6,000
- ❍ 25% chance to win $4000 and 25% chance to win $2000

- ❍ 33% chance to win $2500 and 67% chance of winning nothing
- ❍ 34% chance of winning $2400 and 66% chance of winning nothing

- ❍ 15% chance to lose $20 and 85% chance to lose nothing
- ❍ 10% chance to lose $25 and a 90% chance of losing nothing

Prospect Evaluations – Intertemporal Choice
For the following questions, please indicate which of the two options you prefer:

- ❍ $3400 this month
- ❍ $3800 next month

- ❍ $100 now
- ❍ $140 next year

- ❍ $100 now
- ❍ $1100 in 10 years

- ❍ $9 now
- ❍ $100 in 10 years

- ❍ $40 immediately
- ❍ $1000 in 10 years

- ❍ $100 now
- ❍ $20 every year for 7 years

- ❍ $400 now
- ❍ $100 every year for 10 years

- ❍ $500 in eight months
- ❍ $1060 in sixteen months

- ❍ $500 now
- ❍ $2400 in 2 years

- ❍ $1000 in six months
- ❍ $2400 in two years

- ❍ $100 now
- ❍ $200 next year

What is the smallest amount in 4 days preferred to $170 in 2 months?
Please express your answer in dollars ($).

____

Reference Class & Class-Inclusion Illusions
With the new drug BENOFRENO, the risk of death from a heart attack reduced for people with high cholesterol. A study of 900 with high cholesterol showed that 80 of the 800 people who have not taken the drug died after a heart attack, compared with 16 of the 100 people who did take the drug.

How beneficial was the Benofreno?

- ❍ 1 - Not beneficial
- ❍ 2
- ❍ 3
- ❍ 4
- ❍ 5
- ❍ 6
- ❍ 7 - Very beneficial

How confident are you about your decision?

- ❍ 1 - Not sure
- ❍ 2
- ❍ 3
- ❍ 4
- ❍ 5
- ❍ 6
- ❍ 7 - Very sure

Suppose you have a chance to win $5 by drawing a red ball from either of the two bags. Bag A contains 1 red ball out of 9, and Bag B contains 10 red balls out of 100. Indicate the bag from which you wish to draw a ball:

- ❍ Bag A
- ❍ Bag B

Imagine that you have finished your studies and you need to find a job. You are looking through the newspaper and you read an advertisement from a company that is looking for people like you. This company offers two types of job positions: Type P and Type Q. Both are of the same category and you like them equally. Therefore, you quickly go to the company to present your application to work in either position. Once there, they tell you that you cannot request both at the same time, you have to opt for one of them: P or Q.
For the Type P job, 2 people are needed and only 10 candidates are admitted (one of them would be you).
For the Type Q job, 10 people are needed and only 100 candidates are admitted (one of them would be you).

What job type would you choose?

- ❍ Type P
- ❍ Type Q
- ❍ No preference

What job type do you believe most people would choose?

- ❍ Type P
- ❍ Type Q
- ❍ No preference

What job type do you believe a completely logical person would choose?

- ❍ Type P
- ❍ Type Q
- ❍ No preference

Cancer causes deaths; below are two situations that present cancer risk statistics. Please rate how risky they appear to you.

Cancer kills 1286 people out of 10,000

- ❍ 1 - No risk at all
- ❍ 2
- ❍ 3
- ❍ 4
- ❍ 5
- ❍ 6
- ❍ 7 - Maximum possible risk

Cancer kills 24.14 people out of 100

- ❍ 1 - No risk at all
- ❍ 2
- ❍ 3
- ❍ 4
- ❍ 5
- ❍ 6
- ❍ 7 - Maximum possible risk

Realistic Risky Decisions

Mrs. Jones is told she has a 28 in 1,000 chance of dying from cancer and a 59 in 1,000 chance of dying from a stroke.
Mrs. Jones's doctor tells her that a new pill, STROKEX, will lower her chance of dying from stroke by 50%. Another pill, CANCERX will lower her chance of dying from cancer by 50%.

Assume she can only take 1 pill. Assuming the 2 pills are equally safe and cost the same, which should she take to minimize her risk of death?

- ❍ STROKEX pill
- ❍ CANCERX pill

How confident are you in your previous answer?

- ❍ 0
- ❍ 1
- ❍ 2
- ❍ 3
- ❍ 4
- ❍ 5
- ❍ 6
- ❍ 7
- ❍ 8
- ❍ 9
- ❍ 10

Mrs. Jones decides to take the CANCERX pill. Now, what is her chance of dying from cancer?

- ❍ 0 in 1,000
- ❍ 7 in 1,000
- ❍ 14 in 1,000
- ❍ 21 in 1,000

How confident are you in your previous answer?

- ❍ 0
- ❍ 1
- ❍ 2
- ❍ 3
- ❍ 4
- ❍ 5
- ❍ 6
- ❍ 7
- ❍ 8
- ❍ 9
- ❍ 10

Imagine the weather forecast indicates "There is a 40% chance of rain tomorrow." Which interpretation is most appropriate:

- ❍ The forecaster thinks it will rain in about 40% of the region tomorrow.
- ❍ The forecaster thinks it will rain about 40% of the time tomorrow.
- ❍ The forecaster thinks it will rain for at least 1 hour on 4 out of 10 days like tomorrow.
- ❍ The forecaster thinks there is a 50% chance it will rain in about 80% of the region tomorrow.
- ❍ The forecaster thinks there is a 40% chance it will rain in at least 40% of the region tomorrow.

How confident are you in your previous answer?

- ❍ 0
- ❍ 1
- ❍ 2
- ❍ 3
- ❍ 4
- ❍ 5
- ❍ 6
- ❍ 7
- ❍ 8
- ❍ 9
- ❍ 10

Imagine that you see the following advertisement for a new toothpaste:

Zendil—50% reduction in occurrence of gum inflammation.

Zendil is a new toothpaste to prevent gum inflammation. Half as many people using Zendil developed gum inflammation as people using a different toothpaste.

If you wanted to determine how much the average person could benefit from using Zendil, which single piece of information below would be most helpful?

- ❍ The risk of gum inflammation for people who do not use Zendil
- ❍ The risk of gum inflammation for people who use a different brand of toothpaste for the same purpose
- ❍ How many people there were in the group who used a different toothpaste
- ❍ How old the people who participated in the study were
- ❍ How much a weekly dose of Zendil costs
- ❍ Whether Zendil has been recommended by a dentists' association for this use

Prostate cancer screening means checking a man's prostate for cancer with the Prostate specific antigen (PSA) test before there are symptoms of the disease. The data below shows results from screening for men (50 years or older), depending on whether they participated in prostate cancer screening for 11 years.

From 1,000 men who participated in screening:

7 men died of prostate cancer.

20 men underwent unnecessary cancer treatment. They were diagnosed with prostate cancer but this cancer would not have been found without screening and would not have threatened their lives. The treatment often included surgery to remove the prostate or radiation therapy, which can cause incontinence or impotence.

From 1,000 men who did NOT participate in screening:

7 men died of prostate cancer.

No men were treated unnecessarily.

For each 1,000 men who participated in prostate cancer screening, how many men were saved (i.e., how many men would have died otherwise)?

Please give your answer in terms of "____ men out of 1,000"

____

If 2,000 men participate in screening for 11 years, how many of these men will undergo unnecessary treatment for prostate cancer?

Please give your answer in terms of "____ men out of 2,000"

____

Screening reduces the risk of dying from prostate cancer by ____%.

- ❍ 0
- ❍ 20
- ❍ 50
- ❍ 80
- ❍ 100