Journal of Behavioral Decision Making, *J. Behav. Dec. Making*, **30**: 541–554 (2017)
Published online 2 August 2016 in Wiley Online Library (wileyonlinelibrary.com) **DOI:** 10.1002/bdm.1973

Real-World Correlates of Performance on Heuristics and Biases Tasks in a Community Sample

MAGGIE E. TOPLAK,[1]* RICHARD F. WEST[2] AND KEITH E. STANOVICH[3]
[1]*York University, Canada*
[2]*James Madison University, United States*
[3]*University of Toronto, Canada*

ABSTRACT

In the current study, we sought to examine whether performance on several heuristics and biases tasks and thinking dispositions was associated with real-life correlates in a community sample of adults. We examined performance on five heuristics and biases tasks (ratio bias, belief bias in syllogistic reasoning, cognitive reflection, probabilistic and statistical reasoning, and rational temporal discounting), three thinking dispositions (actively open-minded thinking, future orientation, and avoidance of superstitious thinking), and a questionnaire assessing real-world correlates in several domains (substance use, driving behavior, financial behavior, gambling behavior, electronic media use, and secure computing). Our heuristics and biases tasks and thinking disposition measures were modestly associated with several real-world outcomes, including the domains of secure computing, financial behaviors, and the total scores. That is, better performance on the heuristics and biases measures was associated with fewer negative outcomes. We found that the associations were generally higher in males than in females. Heuristics and biases performance and thinking dispositions were unique predictors of real-world outcomes after statistically controlling for educational attainment and sex differences. Copyright © 2016 John Wiley & Sons, Ltd.

KEY WORDS Heuristics and biases; real-world outcomes; thinking dispositions; sex differences

A substantial research literature—one comprising literally hundreds of empirical studies conducted over several decades—has firmly established that people's responses sometimes deviate from the performance considered normative on many reasoning tasks (Baron, 2008, 2014; Evans, 2014; Kahneman, 2011; Kahneman & Tversky, 2000; Koehler & Harvey, 2004; Manktelow, 2012; Thaler, 2015). Much of this research derives from the so-called heuristics and biases tradition inaugurated by Kahneman and Tversky in the early 1970s (Kahneman & Tversky, 1972, 1973; Tversky & Kahneman, 1974). The term biases refers to the systematic errors that people make in choosing actions and in estimating probabilities, and the term heuristic refers to *why* people often make these errors—because they use mental shortcuts (heuristics) to solve many problems. This literature has had enormous influence, including the Nobel Prize to Kahneman in 2002. Nevertheless, this literature is often criticized for its emphasis on laboratory tasks and some have questioned whether the tasks relate at all to real-world behavior. Thus, in the current study, we sought to examine whether performance on several popular heuristics and biases tasks—as well several thinking dispositions—were associated with real-life outcomes in a community sample of adults.

There is a growing but still relatively sparse literature on the relationship between performance on heuristics and biases tasks and real-world behaviors. For example, pathological gamblers have been shown to perform less well on probabilistic reasoning than non-gamblers (Toplak, Liu, Macpherson, Toneatto, & Stanovich, 2007). Also, hindsight bias and overconfidence effects have been associated with less optimal medical decisions (Arkes, 2013). The most well-studied measure of decision making that has been linked to outcomes has been the test of Decision-Making Competence (DMC; Bruine de Bruin, Parker, & Fischhoff, 2007, 2012; Carnevale, Inbar, & Lerner, 2011; Parker & Fischhoff, 2005; Parker & Weller, 2015; Weller, Ceschi, & Randolph, 2015; Weller, Levin, Rose, & Bossard, 2012; Weller, Moholy, Bossard, & Levin, 2015).

The DMC is a battery of decision making measures that includes tasks assessing: consistency in risk perception, recognizing social norms, resistance to sunk costs, resistance to framing, applying decision rules, path independence, and overconfidence. Parker and Fischhoff (2005) reported a significant relationship between performance on the overall DMC composite and several risk behaviors, such as externalizing behavior and substance use in a sample of 110 18–19 year old males. In particular, higher engagement in risk behavior was significantly associated with the component measures of risk perception, decision rules, and overconfidence. Similarly, on a version of the DMC developed for preadolescents, individuals who performed better on the Youth DMC were more likely to report completing difficult tasks and receiving top grades in class, and less likely to miss homework assignments and to be called to the principal's office for bad behavior (Weller, Levin, Rose, & Bossard, 2012).

Bruine de Bruin, Parker, and Fischhoff (2007) conducted a much larger study and reported significant associations between the DMC and a measure called the Decision Outcomes Inventory (DOI) in a larger sample ($N = 360$) of adults ranging in age from 18 to 88 years. The DOI is a self-report

*Correspondence to: Maggie E. Toplak, 126 BSB, Department of Psychology, York University, 4700 Keele St. Toronto, Ontario, Canada M3J 1P3, E-mail mtoplak@yorku.ca
Preparation of this manuscript was supported by grants from the Social Sciences and Humanities Research Council of Canada to Maggie E. Toplak and from the John Templeton Foundation to Keith Stanovich and Richard West. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.

measure of several negative decision outcomes that sample a wide variety of domains and items that vary in severity (from throwing out food to having a mortgage or loan foreclosed). They found that performance on the DMC predicted unique and significant variance on the DOI. Other studies have also found that better critical thinking performance (as measured by the Halpern Critical Thinking Assessment) was associated with an adapted version of the DOI in college students and in a community sample (Butler, 2012; Butler et al., 2012). In addition to decision-making performance, cognitive styles or dispositions relevant to decision making have also been found to predict real-world outcomes (Bruine de Bruin et al., 2007; Parker & Fischhoff, 2005).

In the present study, we examined a variety of risk behaviors such as substance use, gambling, and problematic driving behaviors. We also examined several indicators of financial behavior, including some previously examined (such as, check bounced on overdrawn debit account; Bruine de Bruin et al., 2007) and new items (such as signing up for recurring payments that were later regretted). In addition to these more commonly studied risk behaviors, we focused on some new areas such as technology use and misuse (Subrahmanyam & Smahel, 2011). In particular, we examined the frequency of technology use of several forms of media, such as number of texts sent per day and number of hours spent on online social media. In addition, we surveyed the use of secure computing behaviors, including: opening attachments sent from an unfamiliar email address, backing up computer work, using privacy settings on social networking sites, and using computer virus protection. We examined whether better performance on several important heuristics and biases tasks is associated with less risk behaviors (substance use and gambling), more prudent financial choices, less technology overuse, and more secure computing behaviors.

This study examined performance on five classic paradigms from the heuristics and biases literature: ratio bias, belief bias in syllogistic reasoning, cognitive reflection, probabilistic and statistical reasoning, and temporal discounting. Successful performance on several of the tasks we chose for investigation requires avoiding or resisting shallow information processing (Kahneman, 2011; Stanovich, 2009, 2011; Stanovich et al., 2016). For example, optimal performance on the well-known Cognitive Reflection Test (CRT; Frederick, 2005) requires recognizing and overriding an incorrect intuitive response in favor of a computed analytic response. Performance on the CRT has also been shown to be highly related to numeracy (Sinayev & Peters, 2015). We used a version of a reflection versus intuition measure inspired by Frederick's original task (see Toplak, West, & Stanovich, 2014a). Similarly, the ratio bias task and belief bias task that we employ in this study are also characterized by a strong lure to select the most visually salient option (in the ratio bias task) and to endorse a believable conclusion that conflicts with logic in our syllogistic reasoning task. Temporal discounting is a theoretically complex construct with several contextual factors that may impact choices (Ainslie, 2001; Basile & Toplak, 2015; Frederick, 2005; McClure, Laibson, Loewenstein, & Cohen, 2004). However, at least partially implicated is the inability to suppress the attraction of the immediacy of the smaller reward and simulate the real worth of the larger delayed reward. Finally, shallow information processing also characterizes performance on several of the problems on our measure of probabilistic and statistical reasoning; however, these problems are more knowledge dependent than the other tasks we employed (see Stanovich et al., 2016). Our study also included three thinking dispositions very relevant to decision making: actively open-minded thinking, future orientation, and superstitious thinking. These thinking dispositions have been shown to be associated with better performance on heuristics and biases tasks (Toplak et al., 2011; West et al., 2008). We expected that the endorsement of actively open-minded thinking and future orientation and the avoidance of superstitious thinking would be associated with better performance on our heuristics and biases tasks.

Because of demographic differences in the risk behaviors that we assessed, we examined our data for differences between males and females. A substantial literature has reported that males engage in significantly more risk behaviors than females (Arnau-Sabates et al., 2013; Figner & Weber, 2011; Harris, Jenkins, & Glaser, 2006; Pawlowski et al., 2008; Thomson & Carlson, 2015; Wong, Zane, Saw, & Chan, 2013). For example, based on a meta-analysis of 150 studies, Byrnes, Miller, and Schafer (1999) reported that males engage in more risk-taking behaviors than females. These behaviors included substance use, gambling, and risky driving behaviors. Similarly, males tend to make riskier financial investments than females (Rolison, Hanoch, Wood, & Liu, 2013). Males also seem to experience more problematic technology use. Specifically, college males reported more problematic internet use (such as, excessive use; Odacı, 2013) and more pedestrian injuries related to mobile phone use than females (Nasar & Troyer, 2013). Differences between males and females on heuristics and biases tasks are less thoroughly investigated. However, it has been reported that males tend to outperform females on the Cognitive Reflection Test (Frederick, 2005; Toplak, West, & Stanovich, 2014a). In contrast, females have been reported to show a slight advantage on delay of gratification tasks and temporal discounting tasks (Silverman, 2003a; 2003b). Less consistent differences have been observed on the Iowa Gambling Task (d'Acremont and Van der Linden, 2006; van den Bos, Homberg, & de Visser, 2013), but some studies have shown that males tend to outperform females (Overman & Pierce, 2013; Overman et al., 2004; Weller et al., 2010).

In summary, the main purpose of this study was to examine the association between performance on heuristics and biases tasks and on thinking dispositions relevant to decision making with self-reported real world outcomes in a community sample of adults. Specifically, we examined a reflection versus intuition measure (cognitive reflection), ratio bias, belief bias in syllogistic reasoning, probabilistic and statistical reasoning, and temporal discounting measures. Thinking dispositions included actively open-minded thinking, future oriented thinking, and superstitious thinking. Our outcome categories included electronic media use, secure computing, substance use, driving behavior, financial behavior, and

DOI: 10.1002/bdm

gambling behavior. These associations were also examined separately for males and females. Finally, we also examined associations with age and educational level attained, as some associations have been reported in the literature (see Bruine de Bruin et al., 2007). We expected that performance on our heuristics and biases tasks would be associated with less electronic media use, more secure computing behaviors, less substance use, and less risk taking behaviors in the domains of driving, finance, and gambling. Given the higher frequency of certain risk behaviors in males, we examined whether the association between heuristics and biases performance and real-world outcomes may be higher in males than in females.

METHOD

Participants and procedure

A total of 232 participants (M age = 43.1 years, SD = 8.3; age range 21 – 64 years of age; 93 males and 139 females) took part in the study. The participants were recruited from a middle class community. We approached parents of participants in our developmental study (Toplak, West, & Stanovich, 2014b), who then recruited other members of the same community by word of mouth. Participants were given paper and pencil surveys and returned them in the mail; they were paid $20 for participating in this study. In order to preserve anonymity of the data, participants were asked to mail in their questionnaires separately from their consent and contact information (to obtain reimbursement).

The current sample was well educated: five participants did not complete high school, 14 participants completed only high school, 20 participants had at least one year of college or university, 125 participants had completed college or university, 66 participants had a graduate or professional degree, and two participants did not report their educational attainment. Education was coded numerically 1 to 5 by these five levels of education. The tasks were presented in the following order: demographics, ratio bias, belief bias syllogisms, the reflection versus intuition measure, probabilistic and statistical reasoning, rational temporal discounting, thinking dispositions, and an outcomes questionnaire.

Tasks and variables

Thinking dispositions

Participants completed a self-report questionnaire in which they were asked to rate their agreement with each question using the following six-point scale: Strongly Disagree (1), Disagree Moderately (2), Disagree Slightly (3), Agree Slightly (4), Agree Moderately (5), and Strongly Agree (6). Three different scales were intermixed in the questionnaire so that the target scales of interest would be less transparent to participants.

Actively open-minded thinking (AOT). We used a 12-item version of the original 41-item measure (Stanovich & West, 1997, 2007). Examples of items are "People should always take into consideration evidence that goes against their beliefs" and "Certain beliefs are just too important to abandon no matter how good a case can be made against them" (reverse scored). The score on the scale was obtained by summing the responses of the items. A higher score indicates greater tendency toward open-minded thinking.

Future oriented thinking. This 14-item thinking disposition scale (see Stanovich et al., 2016) assessed the extent to which participants think about and plan for their futures. It was based on other scales such as the Consideration of Future Consequences Scale (Strathman et al., 1994). Sample items include: "I believe in living for today" (reverse scored) and "I tend to think a lot about what my life will be like in the future". The score on the scale was obtained by summing the responses to the items. A higher score indicates more orientation toward the future.

Avoidance of superstitious thinking (ST). The Superstitious Thinking Scale is a measure of knowledge that may inhibit rational thinking. This 12-item scale was composed of items to assess paranormal beliefs and superstitious thinking, based on questionnaires used in previous studies (see Stanovich et al., 2016). Examples of items include: "Astrology can be useful in making personality judgments," "Some numbers are unlucky," and "I do not believe in any superstitions" (reverse scored). The score on the scale was obtained by summing the items. The total score was reflected so that a higher score indicated avoidance of superstitious thinking.

Heuristics and biases tasks

On each of these tasks, a higher score indicated better performance.

Ratio bias. The eight problems on this task were modeled on Kirkpatrick and Epstein (1992; see also Denes-Raj & Epstein, 1994). An example of a trial read as follows:

Assume that you are presented with two trays of black and white marbles (pictured below). The large tray contains 100 marbles. The small tray contains 10 marbles. The marbles inside each tray will be randomly mixed up, and you must draw out a single marble from one of the trays without looking. **If you draw a black marble you win** $**5**. The small tray contains 1 black marble and 9 white marbles, and the large tray contains 8 black marbles and 92 white marbles. [A drawing of two trays with their corresponding numbers of marbles arranged neatly in 10-marbles-rows was pictured.] Which tray would you prefer to select a marble from in a real situation?"

The following scale was used to indicate preferences: (i) Strongly prefer the small tray; (ii) Moderately prefer the small tray; (iii) Slightly prefer the small tray; (iv) Slightly prefer the large tray; (v) Moderately prefer the large tray; (vi) Strongly prefer the large tray. In the remaining seven trials, the ratio of black:white numbers were as follows: 1:19 versus 4:96, 3:12 versus 18:82, 2:8 versus 19:81, 1:7 versus 12:88, 1:4 versus 19:81, 1:14 versus 4:96, and 1:4 versus 17:83. For all of these eight items, the correct response was to select the small tray. There was one additional foil item, where the ratio was reversed so that the large tray was the

DOI: 10.1002/bdm

better option (1:9 versus 25:75); this foil item was not included in the total score. The items were presented together in the battery. We used a continuous scoring scheme for this variable, so that more points were given for strongly preferring the correct response (strongly preferring the small tray was given 6 points, whereas strongly preferring the large tray was given 1 point).

Belief bias in syllogistic reasoning. Eight syllogistic reasoning problems, largely drawn from Markovits and Nantel (1989), were completed by the participants. Each problem was worded such that the validity judgment was in conflict with the believability of the conclusion. There were two types of these inconsistent syllogisms. One type of inconsistent syllogism had a believable conclusion but an invalid format (e.g., "Premises: All living things need water; Roses need water; Conclusion: Roses are living things"—which is invalid). The other type had an unbelievable conclusions in a logically valid format (e.g., "Premises: All things that are smoked are good for the health; Cigarettes are smoked; Conclusion: Cigarettes are good for the health"—which is valid). Therefore, the believability of the content was inconsistent with the logical format of the syllogism in both types. Problems of this type have typically been thought to mirror the critical thinking skill of being able to put aside one's prior knowledge and reason from new premises. After each item, the participants indicated their responses by selecting one of the two alternatives: (i) conclusion follows logically from premises, or (ii) conclusion does not follow logically from premises. The eight syllogisms were presented together in the battery.

There are two biases that contribute to incorrect selections on this task: a believability bias and a validity bias. All of the syllogisms were inconsistent, which means that the believability of the conclusion conflicted with the logical structure. Thus, if the participant responds based on believability of the conclusion, that participant will obtain a score of 0. Also, four of the syllogisms were valid and four of the syllogisms were invalid. Thus, if a participant assesses each syllogism based on a bias toward validity (instead of believability), that participant will receive a score of four correct simply by marking each of the eight syllogisms as valid. We did not want to penalize one bias over another (believability versus validity) in our scoring. Therefore, scores less than 4 were set to 4. This equates the two biases in that individuals showing a complete believability bias will score the same (4) as those displaying a complete validity bias. See Stanovich, West, and Toplak (2016) for a further discussion of these two different types of biases on the belief bias in syllogistic reasoning task.

Reflection versus intuition task. Based on the original three items from Frederick's (2005) Cognitive Reflection Test (CRT) and analogues of four problems used in Toplak et al. (2014a), there were 11 questions used for this task. The defining characteristic of these problems is that a quick, intuitive answer springs to mind, but that this answer is incorrect. The key to deriving the correct solution is to suppress and/or evaluate the first solution that springs to mind (Frederick, 2005). A commonly used sample item is the following: A bat and a ball cost $1.10 in total. The bat costs $1 more than the ball. How much does the ball cost? Most participants give the first response that comes to mind, which is 10 cents, without thinking further and realizing that this cannot be right. The bat would then have to cost $1.10 and the total cost would then be $1.20 rather than the required $1.10. The correct answer is 5 cents. The majority of the responses on each problem were either the intuitive or analytic response. Correct responses ranged from 34% ($n = 78$) to 82% ($n = 189$). The intuitive response as a percentage of the incorrect responses was, in increasing magnitude: 34%, 57%, 60%, 65%, 69%, 72%, 83%, 88%, 90%, 90%, and 95%. We summed correct performance on these 11 items to derive a total score.

Probabilistic and statistical reasoning. A total of 11 problems adapted from previously published studies from the heuristics and biases literature (Fong et al., 1986; Gal & Baron, 1996; Jepson et al., 1983; Kahneman and Tversky, 1982; Lehman et al., 1988; Sloman, et al., 2003; Stanovich et al., 2016; Toplak et al., 2011, 2014a; Tversky & Kahneman, 1974, 1983; West & Stanovich, 2003) were included to assess probabilistic and statistical reasoning. Specifically, we included: one probability matching item (Toplak et al., 2011), four gambler's fallacy items (Toplak et al., 2011; West et al., 2008), two causal baserate items (West et al., 2008), one quantitative baserate neglect items (Stanovich and West, 1998b), one sample size problem (West et al., 2008), one regression to the mean problem (West et al., 2008), and one conjunction problem (Stanovich & West, 1998c). A commonly used sample item of the Gambler's Fallacy is as follows: 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:

a. It is more likely that tails will come up than heads; b. It is more likely that heads will come up than tails; c. Heads and tails are equally probable on the sixth toss. Answer c is the correct response and was scored as 1, while the other two alternatives were scored as 0. Each item was scored as correct or incorrect. Correct scores indicated responses that: recognized maximizing as a more effective strategy on probability matching; independence of trials on the Gambler's Fallacy; preference for large sample information over salient personal testimony on the causal baserate items; taking into account the baserate of an event on the quantitative baserate neglect items; recognizing that larger sample sizes are more likely to approximate the population value; recognition of regression effects; and indicating that the components are less likely than the conjunction of the components on the conjunction problem. We summed performance on these 11 items to derive a total score.

Rational temporal discounting. Two temporal discounting tasks were used in this study (see Stanovich et al., 2016). The first task was composed of 25 items presented in a staircase format. For each item, the participant was asked to choose between an amount "now" versus "$100 in

DOI: 10.1002/bdm

3 months". The amount for the "now" choice started at $1, ascending in increments until the "now" option reached "$99.50 now" (amounts were: $1, $2.50, $5, $7.50, $10, $15, $20, $25, $30, $35, $40, $45, $50, $60, $65, $70, $75, $80, $85, $90, $92.50, $95, $97.50, $99, and $99.50). For example, one item asked participants to choose between "$90 now" or "$100 in 3 months". In this example, a willingness to wait was worth an extra $10, which is the equivalent of an 11% gain in three months and on a simple interest basis would have resulted in a value increase of about 44% if earned annually. Making the delayed choice for the items that had amounts of "$1 now" through to "$90 now" would represent an annual interest rate between 44% and 39 600%. The delay choices for these 20 items were scored as correct. The other 5 items, which had an annual rate of return below 44%, were not scored.

The second temporal discounting task also had 25 items presented in a staircase format. For each item, the participant was asked to choose between an amount "now" versus "$2000 in 1 year". The amount for the "now" choice started at $1990, descending in increments until the "now" option reached "$20 now" (amounts were: $1990, $1980, $1950, $1900, $1850, $1800, $1700, $1600, $1500, $1400, $1300, $1200, $1000, $900, $800, $700, $600, $500, $400, $300, $200, $150, $100, $50, and $20). For example, one item asked participants to choose between "$1600 now" or "$2000 in 1 year". In this example, a willingness to wait was worth an extra $400, which is the equivalent of 25% rate of interest if earned annually. Making the delayed choice for the items that had amounts of "$1600 now" through to "$20 now" would represent an annual interest rate between 25% to 9900%. The delayed choices for these 18 items were scored as correct. The remaining seven items, which had an annual rate of return below 25%, were not scored.

We were interested in differentiating prudent discounting from other choices in this task. To do this, we selected cut-offs that provided at least a 44% and 25% rate of return based on annual interest rates. Given that it has been argued that the rate of discounting should correspond with the current interest rate (Frederick et al., 2002; Senecal et al., 2012), we chose these higher cut-off scores to demarcate those items where it is clearly an irrational choice to choose the immediate option (Basile & Toplak, 2015; Stanovich, West, & Toplak, 2016). The scores on each of these temporal discounting tasks were standardized and a mean of these two scores was derived.

Heuristics and biases composite. The raw scores on each of the five heuristics and biases tasks were standardized into *z*-scores, and summed to create a composite score. A higher score indicated better performance on these tasks.

Adult outcome measure

Adult outcome questionnaire. We created a questionnaire that included six different domains of outcomes, four related to engagement in risk behaviors (substance use, driving behavior, financial behavior, and gambling behavior) and two related to risk inherent in technology use (use of electronic media and secure computing).

Four questions were used to assess substance use (frequency of the following: cigarettes smoked per day, alcoholic drinks per week, alcoholic drinks on one occasion when out, been hung over from drinking in the last year). Nine questions were used to assess driving behavior (frequency of the following while driving: talk on a hand's free phone, talking on a handheld phone, viewing or sending texts or emails, suspension of driver's license, caused an accident, speeding tickets, driven without a driver's license, driven an uninsured car, charged with driving while intoxicated). Eight questions were used to assess financial behavior (frequency of the following: check bounced or overdrawn debit account, used line of credit for household expenses, taken out a very short-term pay day loan, signed up for recurring payments that later regretted, missed a mortgage or car payment; participants were also asked if they spend more than their available disposable income each month, whether they pay their credit card bill in full each month, and how much of a credit card balance they carry on average each month). Four questions about gambling behavior were included (frequency of the following was asked: gambling per week, gambled more than intended, gambled with borrowed money; participants were also asked what the largest amount of money is that they have ever gambled in one day).

Five questions were asked about media usage (number of hours for each: social networking sites, video gaming, virtual world games, downloading music and videos, online gambling). Four questions were used to assess secure computing (opening attachments from an unfamiliar address, back up work, use privacy settings on social networking sites, and use of virus protection).

For high risk behaviors (such as check bouncing), a frequency of one or more was credited with one point for risk outcome. For other items (such as number of alcoholic drinks per week), a cut-off score was used, and scores above the cut-off were credited with one point for risk outcome. There were 34 items in total. An Outcome Total Score (all 34 items) was derived in addition to separate domain scores (electronic media usage, secure computing, substance use, driving, finance, gambling). In addition, we identified several high risk behaviors that were endorsed with low frequency (endorsed by under 16% of the sample for each item). We identified 10 of these items to create an Outcomes Low Frequency Total score. The actual items and scoring of these items are presented in the Appendix. A higher score on each of these domains or total scores indicated higher outcome risk.

RESULTS

Table 1 displays the mean performance, standard deviations, observed range, potential range, and Cronbach's alpha for all of the measures that were used in this study. The mean scores on the outcome measures indicate that our sample was relatively low risk, endorsing low levels on the outcome variables.

DOI: 10.1002/bdm

Table 1. Descriptive statistics for tasks in the study ($N = 232$)

Task	Number of items	Mean	SD	Observed range	Potential range	Cronbach's alpha
Heuristics and biases tasks						
Ratio bias	8	34.93	10.85	8 to 48	8 to 64	0.91
Belief bias in syllogistic reasoning	8	5.11	1.31	4 to 8	4 to 8	0.62
Reflection versus intuition	11	5.77	3.02	0 to 11	0 to 11	0.80
Probabilistic and statistical reasoning	11	6.66	1.87	2 to 11	0 to 11	0.53
Rational temporal discounting—test 1[a]	25	17.55	3.19	10 to 20	0 to 20	0.91
Rational temporal discounting – test 2[a]	25	15.99	2.70	10 to 18	0 to 18	0.93
Heuristics and biases composite *z*-score	5	0.02	3.52	−7.52 to 7.75	N/A	0.77
Thinking dispositions						
Actively open-minded thinking	12	52.63	7.52	34 to 72	12 to 72	0.71
Superstitious thinking	12	25.55	8.89	12 to 48	12 to 72	0.81
Future orientation	14	47.51	7.68	23 to 70	14 to 84	0.82
Real-world correlates/outcomes						
Electronic media use	5	0.68	0.82	0 to 4	0 to 4	0.31
Secure computing	4	0.65	0.84	0 to 3	0 to 3	0.36
Substance use	4	0.46	0.86	0 to 4	0 to 4	0.60
Driving behavior	9	2.36	1.57	0 to 7	0 to 7	0.49
Financial behavior	8	2.75	1.43	1 to 7	0 to 7	0.35
Gambling	4	0.35	0.72	0 to 4	0 to 4	0.55
Total score	34	7.24	3.15	0 to 18	0 to 34	0.56
Low frequency score	10	0.99	1.18	0 to 6	0 to 10	0.47

[a]Note: The two temporal discounting tasks were averaged to derive a single score that was used for the Heuristics and Biases Composite *z*-score.

Table 2 presents the zero-order correlations among the heuristics and biases measures and thinking disposition measures in this study. Male was coded as 1 and female was coded as 2; a positive correlation would indicate that females scored higher than males. All of the heuristics and biases measures were significantly positively intercorrelated. The 10 correlations ranged from. 26 to 60 and were all significant at the $p < .001$ level of significance. Associations between heuristics and biases performance and actively open-minded thinking were significant and in the expected direction, ranging from. 28 to 49, all $p < .001$. Better performance on the heuristics and biases tasks was associated with more actively open-minded thinking. Future orientation was positively associated with heuristics and biases performance and actively open-minded thinking, all ranging from .19 to .32, $p < .01$ to $p < .001$. Finally, avoidance of superstitious thinking was positively associated with all of our measures with correlations ranging from .16 to 45, $p < .01$ to $p < .001$. Avoidance of superstitious thinking was associated with better performance on the heuristics and biases tasks, higher actively open-minded thinking, and more future orientation.

We also examined associations between these measures and some demographic characteristics: age, male or female status, and educational level. We found that age was largely unrelated to the heuristics and biases composite and actively open-minded thinking, $r = -.09$ and $-.06$, *ns*, respectively. Age was significantly associated with avoidance of superstitious thinking and future orientation, $r = -.12$ and $-.15$, $p < .05$ and $p < .01$, respectively. Older participants endorsed more superstitious thinking and less future orientation. Being male or female was significantly associated with the heuristics and biases composite, actively open-minded thinking, superstitious thinking, and future orientation, $r = -.28$, $p < .001$; $r = -.15$, $p < .05$; $r = .20$, $p < .01$; $r = -.16$, $p < .01$, respectively. Males outperformed females on the heuristics and biases composite, and males endorsed more actively open-minded thinking, more avoidance of superstitious thinking, and more future orientation than females. Finally, educational attainment was also significantly correlated with the heuristics and biases composite, actively open-minded thinking, avoidance of superstitious thinking, and future orientation, $r = .23$, $p < .001$, $r = .15$, $p < .05$, $r = .23$, $p < .001$; $r = .21$, $p < .001$, respectively. Higher educational attainment was associated with better heuristics and biases performance, more actively open-minded thinking, avoidance of superstitious thinking, and more future orientation.

Table 3 displays the association between the heuristics and biases composite score and the thinking dispositions with the real-life outcome measures. In the full sample (first column of each variable), the correlations were generally in the expected direction, but several of them did not reach significance. The heuristics and biases composite, actively open-minded thinking, avoidance of superstitious thinking and future oriented thinking were negatively associated with our outcomes. Overall, better performance on our heuristics and biases thinking measures and scales was associated with better real-life outcomes. However, many of these correlations were quite low in absolute magnitude. In the remaining parts of the table we examined whether aggregating the male and female subsamples might be covering up stronger relationships. Males outperformed females on the heuristics and biases thinking composite (see Table 2), yet previous studies that we reviewed in the introduction have indicated that males show more risk behaviors. The direction of this association would tend to attenuate the expected direction of the correlation (higher risk behaviors being associated with lower heuristics and biases scores) in the aggregate sample.

As in previous research, the males in our sample displayed more total risk behaviors ($M = 11.8$; $SD = 4.0$) than the females

DOI: 10.1002/bdm

Table 2. Correlations between demographic variables, heuristics and biases tasks, and thinking disposition measures ($N = 232$)

Task	1	2	3	4	5	6	7	8	9	10	11	12
Demographic variables												
1. Age	—											
2. Male/Female[a]	−.10	—										
3. Educational level[b]	−.19**	.09	—									
Heuristics and biases measures												
4. Ratio bias	−.13*	−.23***	.17*	—								
5. Belief bias syllogistic reasoning	−.02	−.21***	.22***	.39***	—							
6. Reflection versus intuition	−.06	−.23***	.18**	.34***	.58***	—						
7. Probabilistic and statistical reasoning	−.04	−.25***	.18**	.33***	.47***	.60***	—					
8. Rational temporal discounting	−.09	−.06	.08	.26***	.35***	.29***	.32***	—				
9. Heuristics and biases composite *z*-score	−.09	−.28***	.23***	(.65)***	(.78)***	(.79)***	(.76)***	(.60)***	—			
Thinking dispositions												
10. Actively open-minded thinking	−.06	−.15**	.15*	.31***	.39***	.38***	.39***	.28***	.49***	—		
11. Avoidance of superstitious thinking	−.12*	−.20**	.23***	.25***	.28***	.34***	.32***	.16**	.38***	.45***	—	
12. Future orientation scale	−.16*	−.16**	.21***	.19**	.27***	.24***	.26***	.20**	.32***	.22***	.22***	—

*$p < .05$.
**$p < .01$.
***$p < .001$, one-tailed.
Correlations in parentheses reflect part-whole relationships.
[a]Male was coded as 1 and female was coded as 2; a positive correlation would indicate that females scored higher than males.
[b]$N = 230$ for analyses with educational level.

DOI: 10.1002/bdm

Table 3. Correlations between heuristics and biases composite score, thinking disposition measures, and real-world outcomes ($N = 232$)

	Heuristics and biases composite			Actively open-minded thinking			Avoidance of superstitious thinking			Future oriented thinking		
	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female
Electronic media use	−.02	−.03	.02	−.06	−.05	−.05	−.08	.09	−.18*	.09	.07	.11
Secure computing	−.20***	−.29**	−.13	−.14*	−.11	−.17*	−.17**	−.13	−.19*	−.19**	−.24*	−.16*
Substance use	.02	−.11	.02	−.01	−.15	.08	−.02	−.08	−.08	.02	−.04	<.01
Driving	.05	−.10	.05	−.04	−.16	−.02	−.02	−.09	−.07	.06	−.04	.06
Finance	−.16**	−.23*	−.15*	−.18**	−.26**	−.14	−.21***	−.30**	−.17*	−.13*	−.24*	−.08
Gambling	−.08	−.19*	−.09	−.15*	−.22*	−.15*	−.17**	−.15	−.30***	−.08	−.16	−.09
Outcomes total score	−.12*	−.32***	−.10	−.19**	−.35***	−.16*	−.21***	−.25**	−.30***	−.07	−.22*	−.04
Outcomes low frequency score	−.12*	−.30**	−.04	−.08	−.16	−.05	−.17**	−.23*	−.21**	−.10	−.27**	−.03

Note: Total sample $N = 232$; male sample $n = 93$; female sample $n = 139$.
*$p < .05$.
**$p < .01$.
***$p < .001$, one-tailed.

DOI: 10.1002/bdm

(10.1, $SD = 3.6$, $t(230) = 3.45$, $p < .001$, Cohen's $d = 0.45$). Likewise, the males in our sample displayed more low-frequency risk behaviors ($M = 1.23$; $SD = 1.40$) than the females (0.82, $SD = 0.98$, $t(230) = 2.61$, $p < .01$, Cohen's $d = 0.34$).

Thus, we proceeded to examine the correlations between our experimental tasks and outcome measures separately in our males (second column) and in females only (third column) for each task in Table 3. The very modest associations obtained between the heuristics and biases tasks and outcome measures in the full sample displayed a somewhat different pattern when separated into males and females. In particular, the males displayed a significant association between the heuristics and biases composite score and secure computing, financial behavior, gambling behavior, outcome total score, and outcome low frequency score. In our female sample, only the association with financial behavior reached significance. Comparing the correlation coefficients between males and females on the heuristics and biases composite score, the strength of the association was significantly higher in males than in females on both the total outcome and low frequency outcome scores, $z = -1.7$, $p = .04$, one-tailed; $z = -1.98$, $p = .02$, one-tailed.

A similar pattern of findings was apparent between actively open-minded thinking and future oriented thinking and our outcome measures. In the case of avoidance of superstitious thinking, the associations were higher in females than in the males on some of the outcome scores. However, when comparing correlation coefficients between males and females, none of these comparisons reached significance.

The regression analyses in Table 4 explore the power of educational level, male/female status, and the heuristics and biases tasks and thinking dispositions to predict variance in the total outcome score. In the first two columns of the first regression, educational level, male/female status, and heuristics and biases were entered hierarchically. When entered first, educational level explained 8.4% ($p < .001$) of the variance. When entered next, male/female status explained an additional 5.8% ($p < .001$) of the variance when entered second. Finally, when the heuristics and biases score was entered, it explained a further 1.7% of the variance ($p < .05$). Of the three predictors, male/female status actually explained the most unique variance (see column 3) and had the largest standardized beta weight in the final regression equation (column 4).

The second regression analysis in Table 4 explored the power of educational level, male/female status, and the thinking dispositions composite to predict variance in the total outcome score. In the first two columns of the this regression, educational level, male/female status, and the thinking dispositions were entered hierarchically. As before, when entered first, educational level explained 8.4% ($p < .001$) of the variance. When entered next, male/female status explained an additional 5.8% ($p < .001$) of the variance when entered second. Finally, when the thinking disposition composite score were was entered, it explained a further 4.6% of the variance ($p < .05$). Of the three predictors, male/female status actually explained the most unique variance (see column 3) and had the largest standardized beta weight in the final regression equation (column 4).

DISCUSSION

This study found that the five examined heuristics and biases tasks were significantly intercorrelated, which is consistent with previous findings (Frederick, 2005; Toplak et al. 2011; 2014a, 2014b; Stanovich & West, 1998a; West et al., 2008). However, this association has not always been found across different decision-making domains (Bruine de Bruin et al., 2007; Strough et al., 2015). Similarly, the heuristics and biases composite score was significantly correlated with thinking dispositions related to actively open-minded thinking and future oriented thinking, as well as avoidance of superstitious thinking, which is consistent with other previous studies that have demonstrated and an association between

Table 4. Regression results ($N = 230$)

	R^2 change	F to enter	Unique variance explained	Standardized betas in final equation
Criterion variable = total outcome score				
1. Educational level	.084	20.94***	.049***	−.231***
2. Male/female status[a]	.058	15.31***	.073***	−.284***
3. Heuristics and biases composite score	.017	4.64*	.017*	−.142*
Overall regression:				
$F = 14.26$***				
Multiple $R^2 = .16$				
Criterion variable = total outcome score				
1. Educational level	.084	20.94***	.036**	−.201**
2. Male/female status[a]	.058	15.31***	.084***	−301***
3. Thinking dispositions composite: AOT; future oriented thinking; avoidance of superstitious thinking	.046	12.74***	.046***	−.231***
Overall regression:				
$F = 17.41$***				
Multiple $R^2 = .19$				

*$p < .05$.
**$p < .01$.
***$p < .001$.
[a]Male was coded as 1 and female was coded as 2.

DOI: 10.1002/bdm

heuristics and biases performance and decision making styles (Bruine de Bruin et al., 2007; Parker & Fischhoff, 2005; Toplak et al. 2011; 2014a; West et al., 2008).

Age was not consistently correlated with our heuristics and biases tasks and thinking dispositions. Bruine de Bruin et al. (2007) reported associations between age and some of their decision-making competence subtests, including Resistance to Framing and Applying Decision Rules. The current study examined other heuristics and biases tasks and did not have as wide an age range as the Bruine de Bruin et al. (2007) study. The negative correlation between age and the future orientation scale may reflect the somewhat older age of our participants (middle-aged, as opposed to the typical undergrad samples of college students), as it has been suggested that with aging, people are more likely to perceive time as finite and tend to invest fewer resources to gather information and expand their horizons (Carstensen, 2006). The impact of aging on decision making, including deliberative, affective, and experience-based skills, will be an important topic for continued research (Strough, Parker, & Bruine de Bruin, 2015). Males tended to outperform females on the heuristics and biases tasks used in this study, which may not be surprising given the literature that has reported such findings before (Frederick, 2005; Toplak et al., 2014a). Educational level was also significantly associated with the heuristics and biases composite, actively open-minded thinking, future orientation and with avoidance of superstitious thinking, which is consistent with other research that has shown an association between educational attainment and decision making competence (Bruine de Bruin et al., 2007).

In our community sample of adults, a modest association between heuristics and biases performance and real-world outcomes was found. When these associations were examined separately in our males and females, the relationships between the heuristics and biases measures and outcomes were amplified in the male sample. These correlations obtained even though, in terms of the real-world outcome variables, our sample would most accurately be characterized as a low-risk sample. Out of a total of 34 different outcome behaviors, our sample endorsed a mean of 7.24 (SD = 3.15), with an observed range of 0 to 18. We did, however, find that heuristics and biases performance and thinking dispositions predicted outcomes after statistically controlling for educational attainment and male/female status.

This study examined outcomes in both specific domains (substance use, driving behavior, financial behavior, gambling behavior, electronic media use, and secure computing) and overall scores (an outcome total score and an outcome low frequency score). In the total sample, a significant association was obtained between the heuristics and biases composite score and the specific domains of secure computing and finance, but not with the other specific domains (electronic media use, substance use, driving, and gambling). A significant association was obtained between the heuristics and biases composite scores and the outcome total and outcome low frequency scores. Our results suggest that there may be utility in examining specific domains of life outcomes, as has been done in the risk perception literature (Weber, Blais, & Betz, 2002).

Behaviors related to secure computing made up a novel domain explored in this study, composed of four questions asking whether participants: opened attachments from unfamiliar email addresses, backed up their work, used privacy settings on social networking sites, and used virus protection. The domain of secure computing was our strongest association between the heuristics and biases composite score and outcomes. There was slightly more variability in the endorsement of these behaviors in our sample, relative to substance use and gambling.

Secure computing is a particularly interesting domain for several reasons. Secure computing reflects behavior in a relatively new type of activity in our modern society, as the internet has been widely available for less than 20 years. The items we used to assess this domain seem to reflect "good habits" with the use of our modern technological devices, as opposed to engagement in highly risky behaviors, such as substance use and gambling. The items used to assess secure computing resemble passive risk taking behaviors, namely behaviors that may reduce the likelihood of a particular outcome (Keinan & Bereby-Meyer, 2012). One could argue that these behaviors related to secure computing are an increasingly important domain of function in our modern society. The distinction between benign versus hostile environments is particularly relevant here. A benign environment has been defined as containing useful cues that are easily identified by our heuristic, implicit processes, and such environments do not have individuals who will adjust their behavior to alter these cues (Stanovich, 2009, 2010). In contrast, a hostile environment provides meager useful cues for heuristic processes, and if cues are present, other individuals may alter these cues to their own advantage, making these cues unreliable for the agent.

The internet is an example of a more hostile than benign environment. For example, useful cues to prevent computer viruses and protect secure information are not always obvious to the user, and computer viruses are constantly being adjusted to "trick" users to think that they are legitimate sources that require attention by opening attachments. There are already numerous examples of very bright individuals in high-level positions who have demonstrated cognitive failures on the internet that are related to secure computing, such as unintended postings of hurtful comments about friends and colleagues on the internet (Crane, 2012; Das & Sahoo, 2011; Solove, 2007). Being conscious of who can access your personal information and photos on a social networking website may prevent some thoughtless comment from becoming "viral." Similarly, keeping updated virus protection on your computer may protect against potential and permanent loss of information and files. In fairly well-educated samples, such as in the current study, the domain of secure computing may be more diagnostic than many other types of outcomes.

Although we found that secure computing was associated with our heuristics and biases measures, electronic media use was not associated with the heuristics and biases composite score. The domain of electronic media use was assessed as frequency of use, on the basis of high levels of use ("overuse"). Frequency of use may not have been diagnostic in this sample, perhaps also because there was little evidence of

DOI: 10.1002/bdm

overuse. Overall, the domain of technology use and indicators of maladaptive uses are worthy of further exploration.

An association between the heuristics and biases tasks and financial behaviors was found, which is consistent with other studies that have shown a link between financial literacy and decision-making (Thoma et al., 2015). Although the lack of associations with the other specific domains, including substance use, driving behavior, and gambling was not expected, perhaps this is not surprising given the very low risk nature of our community sample. Overall, however, our findings generally support previous studies which have shown that higher risk behaviors are associated with lower decision-making performance (Bruine de Bruin et al., 2007, 2012; Parker & Fischhoff, 2005; Weller et al., 2012; 2015a).

Our outcome measure differed from other measures that have been reported in the literature, such as the DOI (Bruine de Bruin et al., 2007). The DOI was developed based on pilot research that sampled a wide variety of domains and behaviors that also varied in severity. The calculation of the DOI score is based on weighting decision outcomes to take into account the fact that some of the more severe outcomes assessed were very infrequent (such as spending a night in jail). While we did not assess whether participants had had the opportunity to engage in some of these experiences, we did use several of the high risk behaviors from the DOI in the current study, including substance use and financially risky behaviors. We attempted to elaborate some of the behaviors and domains captured on the DOI, including outcomes related to technology use. We also sought to identify additional specific domains of outcomes as has been done in the risk perception literature (Weber et al., 2002). Given the modest reliabilities of our outcome measures, it will be important to examine more items for each domain. Future directions could also involve an item-level analysis of our outcomes questionnaire (as done by Parker, Bruine de Bruin, & Fischhoff, 2015), in order to develop a more refined, shorter measure. However, the frequency of these behaviors may have important demographic associations, such as sex differences and educational level, which need to considered in developing such a measure.

Except for the superstitious thinking measure, associations between the heuristics and biases thinking composite score and real-life outcomes were stronger in the male than in the female subsample. These findings illustrate that investigations of heuristics and biases performance and life outcomes may benefit from more attention to sex differences. Previous research has indicated that males tend to engage in more risky behaviors (Byrnes et al., 1999) and also tend to perform better on some heuristics and biases tasks (Frederick, 2005; Toplak et al., 2014a) compared to females, which is consistent with the pattern of correlations in Tables 2 and 3. It will be useful to consider other variables that may amplify or attenuate the relationship between decision-making and life outcomes.

There were some limitations in the present study. We were unable to obtain a measure of cognitive ability, which would have allowed us to examine the separable contributions of cognitive abilities and heuristics and biases performance in the real-world outcomes. Our regression analyses used education as a proxy for cognitive ability. Finally, the real-world outcomes measure was based on self-report. While a more objective measure of actually experienced outcomes would have been more desirable, measuring outcomes in this way is conventional in this literature (Bruine de Bruin et al., 2007; Butler, 2012; Butler et al., 2012). It will also be important to use diverse samples with more variability in educational levels and longitudinal designs to consider causal associations.

In summary, we found that tendencies to resist miserly information processing on heuristics and biases tasks were associated real-life outcomes, especially the domains of secure computing and financial behaviors. Similarly, thinking dispositions (toward actively open-minded thinking and future orientation and to avoid superstitious thinking) were also associated with these outcomes, even after statistically controlling for educational level and male/female status. These associations may become of growing practical importance as modern society becomes increasingly hostile to people who cannot avoid miserly processing.

REFERENCES

Ainslie, G. (2001). *Breakdown of will*. Cambridge: Cambridge University Press.

Arkes, H. R. (2013). The consequences of the hindsight bias in medical decision making. *Current Directions in Psychological Science*, *22*(5), 356–360. 10.1177/0963721413489988.

Arnau-Sabates, L., Garcia, M. J., Munoz, M. M., & Capdevila, J. M. (2013). The relationship between awareness of road safety measure and accident involvement in pre-drivers: The basis of a road safety programme. *Journal of Risk Research*, *16*(5), 635–650. 10.1080/13669877.2012.761272.

Baron, J. (2008). *Thinking and deciding* (Fourth ed.). Cambridge, MA: Cambridge University Press.

Baron, J. (2014). Heuristics and biases. In E. Zamir, & D. Teichman (Eds.), *The Oxford handbook of behavioral economics and the law* (pp. 3–27). Oxford: Oxford University Press.

Basile, A. G., & Toplak, M. E. (2015). Four converging measures of temporal discounting and their relationships with intelligence, executive functions, thinking dispositions, and behavioral outcomes. *Frontiers in Psychology*, *6*, 13. 10.3389/fpsyg.2015.00728.

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

Bruine de Bruin, W., Parker, A. M., & Fischhoff, B. (2012). Explaining adult age differences in decision-making competence. *Journal of Behavioral Decision Making*, *25*(4), 352–360. 10.1002/bdm.712.

Butler, H. A. (2012). Halpern Critical Thinking Assessment predicts real-world outcomes of critical thinking. *Applied Cognitive Psychology*, *26*(5), 721–729. 10.1002/acp.2851.

Butler, H. A., Dwyer, C. P., Hogan, M. J., Franco, A., Rivas, S. F., Saiz, C., et al. (2012). The Halpern Critical Thinking Assessment and real-world outcomes: Cross-national applications. *Thinking Skills and Creativity*, *7*(2), 112–121. 10.1016/j.tsc.2012.04.001.

Byrnes, J. P., Miller, D. C., & Schafer, W. D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin*, *125*(3), 367. 10.1037/0033-2909.125.3.367.

Carnevale, J. J., Inbar, Y., & Lerner, J. S. (2011). Individual differences in need for cognition and decision-making competence among leaders. *Personality and Individual Differences*, *51*(3), 274–278. 10.1016/j.paid.2010.07.002.

DOI: 10.1002/bdm

Carstensen, L. L. (2006). The influence of a sense of time on human development. *Science*, *312*(5782), 1913–1915. 10.1126/science.1127488.

Crane, C. (2012). Social networking v. The employment-at-will doctrine: A potential defense for employees fired for facebooking, terminated for twittering, booted for blogging, and sacked for social networking. *Washington University Law Review*, *89*(3), 639–672.

d'Acremont, M., & Van der Linden, M. (2006). Gender differences in two decision-making tasks in a community sample of adolescents. *International Journal of Behavioral Development*, *30*(4), 352–358. 10.1177/0165025406066740.

Das, B., & Sahoo, J. S. (2011). Social networking sites—A critical analysis of its impact on personal and social life. *International Journal of Business and Social Science*, *2*(14), 222–228.

Denes-Raj, V., & Epstein, S. (1994). Conflict between intuitive and rational processing: When people behave against their better judgment. *Journal of Personality and Social Psychology*, *66*, 819–829. 10.1037/0022-3514.66.5.819.

Evans, J. S. B. T. (2014). *Reasoning, rationality and dual processes*. London: Psychology Press.

Figner, B., & Weber, E. U. (2011). Who takes risks when and why? Determinants of risk taking. *Current Directions in Psychological Science*, *20*(4), 211–216. 10.1177/0963721411415790.

Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology*, *18*, 253–292. 10.1016/0010-0285(86)90001-0.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*, 25–42.

Frederick, S., Loewenstein, G. F., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, *40*, 351–401.

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

Harris, C. R., Jenkins, M., & Glaser, D. (2006). Gender differences in risk assessment: Why do women take fewer risks than men. *Judgment and Decision Making*, *1*(1), 48–63.

Jepson, C., Krantz, D., & Nisbett, R. (1983). Inductive reasoning: Competence or skill? *Behavioral and Brain Sciences*, *6*, 494–501. 10.1017/S0140525X00017180.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus & Giroux.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, *3*, 430–454.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*, 237–251.

Kahneman, D., & Tversky, A. (1982). On the study of statistical intuitions. *Cognition*, *11*, 123–141.

Kahneman, D., & Tversky, A. (Eds) (2000). *Choices, values, and frames*. Cambridge: Cambridge University Press.

Keinan, R., & Bereby-Meyer, Y. (2012). "Leaving it to chance" – Passive risk taking in everyday life. *Judgment and Decision Making*, *7*(6), 705–715.

Kirkpatrick, L., & Epstein, S. (1992). Cognitive-experiential self-theory and subjective probability: Evidence for two conceptual systems. *Journal of Personality and Social Psychology*, *63*, 534–544. 10.1037/0022.

Koehler, D. J., & Harvey, N. (Eds) (2004). *Blackwell handbook of judgment and decision making*. Oxford, England: Blackwell.

Lehman, D. R., Lempert, R. O., & Nisbett, R. E. (1988). The effects of graduate training on reasoning: Formal discipline and thinking about everyday-life events. *American Psychologist*, *43*(6), 431–442. 10.1037/0003-066X.43.6.431.

Manktelow, K. (2012). *Thinking and reasoning: An introduction to the psychology of reason, judgment and decision making*. Hove, UK: Psychology Press.

Markovits, H., & Nantel, G. (1989). The belief-bias effect in the production and evaluation of logical conclusions. *Memory & Cognition*, *17*, 11–17.

McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science*, *306*, 503–507. 10.1126/science.1100907.

Nasar, J. L., & Troyer, D. (2013). Pedestrian injuries due to mobile phone use in public places. *Accident Analysis & Prevention*, *57*, 91–95. 10.1016/j.aap.2013.03.021.

Odacı, H. (2013). Risk-taking behavior and academic self-efficacy as variables accounting for problematic internet use in adolescent university students. *Children and Youth Services Review*, *35*(1), 183–187. 10.1016/j.childyouth.2012.09.011.

Overman, W. H., & Pierce, A. (2013). Iowa Gambling Task with non-clinical participants: effects of using real + virtual cards and additional trials. *Frontiers in Psychology*, *4*, 935. 10.3389/fpsyg.2013.00935.

Overman, W. H., Frassrand, K., Ansel, S., Trawalter, S., Bies, B., & Redmond, A. (2004). Performance on the IOWA card task by adolescents and adults. *Neuropsychologia*, *42*(13), 1838–1851.

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 of Psychology,Cognition*, *6*Article 363, 1–7. 10.3389/fpsyg.2015.00363.

Parker, A. M., & Fischhoff, B. (2005). Decision-making competence: External validation through an individual differences approach. *Journal of Behavioral Decision Making*, *18*, 1–27. 10.1002/bdm.481.

Parker, A. M., & Weller, J. A. (2015). Greater decision-making competence is associated with greater expected-value sensitivity, but not overall risk taking: an examination of concurrent validity. *Frontiers in Psychology*, *6*. 10.3389/fpsyg.2015.00717.

Pawlowski, B., Atwal, R., & Dunbar, R. I. M. (2008). Sex differences in everyday risk-taking behavior in humans. *Evolutionary Psychology*, *6*(1), 29–42. 10.1177/147470490800600104.

Rolison, J. J., Hanoch, Y., Wood, S., & Liu, P. (2013). Risk-taking differences across the adult life span: A question of age and domain. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*gbt081. 10.1093/geronb/gbt081.

Senecal, N., Wang, T., Thompson, E., & Kable, J. W. (2012). Normative arguments from experts and peers reduce delay discounting. *Judgment and decision making*, *7*, 568–589.

Silverman, I. W. (2003a). Gender differences in delay of gratification: A meta-analysis. *Sex Roles*, *49*(9-10), 451–463. 10.1023/A:1025872421115.

Silverman, I. W. (2003b). Gender differences in resistance to temptation: Theories and evidence. *Developmental Review*, *23*(2), 219–259. 10.1016/S0273-2297(03)00012-1.

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

Sloman, S. A., Over, D., Slovak, L., & Stibel, J. M. (2003). Frequency illusions and other fallacies. *Organizational Behavior and Human Decision Processes*, *91*, 296–309. 10.1016/S0749-5978(03)00021-9.

Solove, D. J. (2007). *The future of reputation: Gossip, rumor, and privacy on the internet*. New Haven: Yale University Press.

Stanovich, K. E. (2009). *What intelligence tests miss: The psychology of rational thought*. New Haven, CT: Yale University Press.

Stanovich, K. E. (2010). *Decision making and rationality in the modern world*. New York: Oxford University Press.

Stanovich, K. E. (2011). *Rationality and the reflective mind*. New York: Oxford University Press.

Stanovich, K. E., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology*, *89*, 342–357. 10.1037/0022-0663.89.2.342.

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

Stanovich, K. E., & West, R. F. (1998b). Who uses base rates and P (D/~H)? An analysis of individual differences. *Memory & Cognition*, *26*, 161–179. 10.3758/BF03211379.

DOI: 10.1002/bdm

Stanovich, K. E., & West, R. F. (1998c). Individual differences in framing and conjunction effects. *Thinking and Reasoning*, *4*, 289–317. 10.1080/135467898394094.

Stanovich, K. E., & West, R. F. (2007). Natural myside bias is independent of cognitive ability. *Thinking & Reasoning*, *13*, 225–247. 10.1080/13546780600780796.

Stanovich, K. E., West, R. F., & Toplak, M. E. (2016). *The Rationality Quotient (RQ): Toward a test of rationality thinking*. Cambridge, MA: MIT Press.

Strathman, A., Gleicher, F., Boninger, D. S., & Scott Edwards, C. (1994). The consideration of future consequences: Weighing immediate and distant outcomes of behavior. *Journal of Personality and Social Psychology*, *66*, 742–752. 10.1037/0022-3514.66.4.742.

Strough, J., Parker, A. M., & Bruine de Bruin, W. (2015). Understanding life-span developmental changes in decision-making competence. In T. Hess, J. Strough, & C. Loeckenhoff (Eds.), *Aging and decision making: Empirical and applied perspectives* (pp. 235–257).

Subrahmanyam, K., & Smahel, D. (2011). *Digital youth: The role of media in development*. New York: Springer.

Thaler, R. H. (2015). Misbehaving: The making of behavioral economics: WW Norton & Company.

Thoma, V., White, E., Panigrahi, A., Strowger, V., & Anderson, I. (2015). Good thinking or gut feeling? Cognitive reflection and intuition in traders, bankers and financial non-experts. *PloS One*, *10*(4). 10.1371/journal.pone.0123202.

Thomson, C. J., & Carlson, S. R. (2015). Increased patterns of risky behaviours among helmet wearers in skiing and snowboarding. *Accident Analysis & Prevention*, *75*, 179–183. 10.1016/j.aap.2014.11.024.

Toplak, M., Liu, E., Macpherson, R., Toneatto, T., & Stanovich, K. E. (2007). The reasoning skills and thinking dispositions of problem gamblers: A dual-process taxonomy. *Journal of Behavioral Decision Making*, *20*, 103–124. 10.1002/bdm.544.

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*, 1275–1289. 10.3758/s13421-011-0104-1.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014a). Assessing miserly processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning*, *20*, 147–168. 10.1080/13546783.2013.844729.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014b). Rational thinking and cognitive sophistication: Development, cognitive abilities, and thinking dispositions. *Developmental Psychology*, *50*, 1037–1048. 10.1037/a0034910.

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

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, *90*, 293–315.

van den Bos, R., Homberg, J., & de Visser, L. (2013). A critical review of sex differences in decision-making tasks: Focus on the Iowa Gambling Task. *Behavioural Brain Research*, *238*, 95–108. 10.1016/j.bbr.2012.10.002.

Weber, E. U., Blais, A., & Betz, N. E. (2002). A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors. *Journal of Behavioral Decision Making*, *15*, 263–290. 10.1002/bdm.414.

Weller, J. A., Ceschi, A., & Randolph, C. (2015a). Decision-making competence predicts domain-specific risk attitudes. *Frontiers in Psychology*, *6*. 10.3389/fpsyg.2015.00540.

Weller, J. A., Levin, I. P., & Bechara, A. (2010). Do individual differences in Iowa Gambling Task performance predict adaptive decision making for risky gains and losses? *Journal of Clinical and Experimental Neuropsychology*, *32*, 141–150.

Weller, J. A., Levin, I. P., Rose, J. P., & Bossard, E. (2012). Assessment of decision-making competence in preadolescence. *Journal of Behavioral Decision Making*, *25*(4), 414–426. 10.1002/bdm.744.

Weller, J. A., Moholy, M., Bossard, E., & Levin, I. P. (2015b). Preadolescent decision-making competence predicts interpersonal strengths and difficulties: A 2-year prospective study. *Journal of Behavioral Decision Making*, *28*(1), 76–88.

West, R. F., & Stanovich, K. E. (2003). Is probability matching smart? Associations between probabilistic choices and cognitive ability. *Memory & Cognition*, *31*, 243–251.

West, R. F., Toplak, M. E., & Stanovich, K. E. (2008). Heuristics and biases as measures of critical thinking: Associations with cognitive ability and thinking dispositions. *Journal of Educational Psychology*, *100*, 930–941. 10.1037/a0012842.

Wong, G., Zane, N., Saw, A., & Chan, A. K. (2013). Examining gender differences for gambling engagement and gambling problems among emerging adults. *Journal of Gambling Studies*, *29*(2), 171–189. 10.1007/s10899-012-9305-1.

APPENDIX. REAL WORLD OUTCOMES QUESTIONNAIRE

Scoring: 0 = good outcome, 1 = bad outcome

Electronic Media Use

1. Which of the following types of electronic media do you use?

Electronic media use includes any materials for playing videogames and/or accessing the Internet for fun activities. (*1 or more hours scored as 1*, *less than 1 hour scored as 0*)

a. Number of hours a day on social networking sites, such as Facebook and Twitter. _____ hours

b. Number of hours a day on video games.

_____ hours

c. Number of hours a day on virtual worlds.

_____ hours

d. Number of hours a day downloading music and videos.

_____ hours

e. Number of hours a day on online gambling.

_____ hours

Secure Computing

1. When using **computers and/or electronic media**, do you: (Circle Yes or No) (*Yes scored as 1 for a*, *others No scored as 1*)

a. Open attachments in emails that are sent by someone you don't know?
Yes No

b. Backup important work and documents on your computer?
Yes No

c. Do you use the privacy settings on social networking sites, such as Facebook and Twitter, to protect your personal information and privacy?
Yes No

d Have virus protection on your home computers?
Yes No

Substance Use

DOI: 10.1002/bdm

1. How many cigarettes do you smoke a day? Options: 0, 1–5, 6–10, 11–15, or More than 15 Cigarettes (*1 or more scored as 1*, *otherwise 0*)
2. How many alcoholic drinks do you have a week? *Estimate over 7 days*. Options: 0, 1–7, 7–14, 15–21, or More than 21 drinks. (*7 or more scored as 1*)
3. How many alcoholic drinks do you usually have on one occasion when you are out? Options: 0, 1–3, 4–6, 7–10, or More than 10 drinks. (*4 or more scored as 1*)
4. How many times have you been hung over from drinking too much in the past year? Options: 0, 1–3, 4–6, or More than 6 times. (*4 or more scored as 1*)

Driving Behavior

1. How many times a week do you talk on a hand's free cell phone while driving? Options: 0, 1–2, 3–4, 5–7, or More than 7 times. (*1 or more scored as 1*)
2. How many times a week do you talk on a handheld cell phone while driving? Options: 0, 1–2, 3–4, 5–7, or More than 7 times (*1 or more scored as 1*)
3. How many times a week do you send or look at text or email messages while driving? _____ times (*1 or more scored as 1*)
4. Have you ever had your driver's license suspended? Options: Never, Once, 2–5 times, 6–10 times, or More than 10 times. (*1 or more scored as 1*)
5. Have you ever caused a car accident? Options: Never, Once, 2–5 times, 6–10 times, or More than 10 times. (*1 or more scored as 1*)
6. Have you ever received a speeding ticket? Options: Never, Once, 2–5 times, 6–10 times, or More than 10 times. (*1 or more scored as 1*)
7. Have you ever driven without a valid driver's license? Options: Never, Once, 2–5 times, 6–10 times, or More than 10 times. (*1 or more scored as 1*)
8. Have you ever driven a car that is not insured? Options: Never, Once, 2–5 times, 6–10 times, or More than 10 times. (*1 or more scored as 1*)
9. Have you ever been charged with a DWI (Driving While Intoxicated)? Options: Never, Once, 2–5 times, 6–10 times, or More than 10 times. (*1 or more scored as 1*)

Financial Behavior

1. Have you ever had a check bounce or overdrawn your debit account? Options: Never, Once, 2–5 times, 6–10 times, or More than 10 times. (*1 or more scored as 1*)
2. How often have you used a line of credit to cover household expenses, such as purchasing items for your home, vacations, or going out for dinner? Options: Never, Once, 2–5 times, 6–10 times, or More than 10 times. (*1 or more scored as 1*)
3. Have you ever taken out a very short-term (i.e., payday) loan? Options: Never, Once, 2–5 times, 6–10 times, or More than 10 times. (*1 or more scored as 1*)
4. I budget my income as follows: ("*under*" *scored as 0*, "*all*" *scored as 0.5*, *and* "*more*" *scored as 1*)

❑ I spend under the amount of disposable income I have available each month
❑ I spend income I have available each month
❑ I spend more than the amount of disposable income I have available each month

How many times have you intentionally signed up for recurring payments that you later regretted? Options: Never, Once, 2–3 times, 4–5 times, or More than 5 times. (*1 or more scored as 1*)
Have you ever missed a mortgage or a car payment? Options: Never, Once, 2–3 times, 4–5 times, or More than 5 times. (*1 or more scored as 1*)
If yes, do you normally pay your credit card bill in full each month? Options: Yes or No (*1is scored as No*)
How much of a balance do you carry, unpaid, each month on average? (*More than 0 is scored as 1*)

Gambling Behavior

1. How many times a week do you gamble? Options: Never, Once, 2–3 times, or More than 3 times. (*Once or more scored as 1*)
2. What is the largest amount of money you have ever gambled with on any one day? Options: Never gambled, $1 or less, More than $1 up to $10, More than $10 up to $100, More than $100 up to $1000, More than $1000 up to $10 000, or More than $10 000. (*More than $1000 scored as 1*)
3. How many times have you gambled more than you intended to? Options: Never, Once, 2–3 times, or More than 3 times. (*Once or more scored as 1*)
4. Have you ever gambled with borrowed money or from a source that you should not have used? Options: Never, Once, 2–3 times, or More than 3 times. (*Once or more scored as 1*)

The Outcomes Low Frequency Score was based on the following 10 questions: suspension of driver's license, driven without a driver's license, driven an uninsured car, charged with driving while intoxicated, taken out a very short-term pay day loan, if participants spent more than their available disposable income each month, whether participants have been fired from employment (see below), missed a mortgage or car payment, gambled more than intended, and gambled with borrowed money.

1. How many times have you been fired from employment? Options: Never, Once, 2–3 times, 4–5 times, or More than 5 times. (*1 or more scored as 1*)

DOI: 10.1002/bdm