Title

Selective Numeracy: Effects of Numeracy, Popular-Science Reportsand Personal Experience on Data-Based Decision Making

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 42(0)

Authors
Liu, Wan
Penner, Marcie

Publication Date
2020

Peer reviewed

Selective Numeracy: Effects of Numeracy, Popular-Science Reports and Personal Experience on Data-Based Decision Making

Wan Liu (wliu365@uwo.ca)
Department of Psychology, 266 Epworth Avenue
London Ontario, N6A 2M3 Canada

Marcie Penner (mpennerw@uwo.ca)
Department of Psychology, 266 Epworth Avenue
London Ontario, N6A 2M3 Canada

Abstract

In the current research, we investigated whether numeracy, scientific reports in the popular press, and personal experience were associated with people's data-based decision making. We collected data from English-speaking adult participants (N = 187), residing in the United States and Canada, who were recruited through Amazon Mechanical Turk and completed the online study. Results showed that participants with higher numeracy were more likely to make the correct data-based decision. However, participants used their numeracy selectively. They seemed to use their numeracy skills to confirm their own desire rather than to objectively evaluate the data or confirm reported scientific findings. No significant association was found between personal experience and data-based decision making. Future research may examine decision making across other, general-life domains to examine the replicability of the current results.

Keywords: Numeracy; Decision-making; Judgement

Everyday cognitive judgements involve predictions and decision making from limited available data (Griffiths & Tenenbaum, 2006). How do people make data-based decisions? Kahneman (2011) posited two cognitive systems used in making judgements and decisions; System 1 thinking is automatic, intuitive and quick, whereas System 2 thinking is slow, analytic and effortful. System 1 thinking relies on heuristics, which although generally effective can lead to systematic bias in decision making (Kahneman, 2011).

In today's world, where data analysis skills are becoming more and more important, the ability to make rational data-based decisions is a highly-relevant life skill (Xia & Gong, 2015) with implications for one's wellbeing (Gurmankin, Baron, & Armstrong, 2004; Hamm, Bard, & Scheid, 2003; Låg, Bauger, Lindberg, & Friborg, 2014). For example, in deciding whether to request a specific drug treatment to combat a virus, one might look to the widely-reported scientific data on the drug's effectiveness. Chen et al., (2020) reported that 87% of the treatment group (receiving hydroxychloroquine) and 93% of the control group recovered from COVID-19 within seven days. Based on this evidence, would you want to take the drug to treat COVID-19? Making accurate, data-based decisions requires deliberate application of numeracy skills, or System 2 thinking, rather than going with your gut (System 1; Kahneman, 2011). People's ability and willingness to engage System 2 to make accurate data-based decisions are impacted by a variety of factors including their numeracy skills (Peters et al., 2006), exposure to reports by experts (Hovland, Janis, & Kelley, 1953), and prior personal experience (Weber, 2006). People engage System 2 thinking selectively, displaying confirmation bias by taking evidence that confirms their existing views at face value and reinterpreting disconfirming evidence to diminish its impact (Baron, 2008).

In the current research, several variables that can influence people's data-based decision making were examined: numeracy, science reports in the popular press, and personal experience. People's numeracy skills, that is, the ability to process and work with numbers and probabilities, would logically influence their ability to make data-based decisions. Science reports here refer to the articles people read specifically in popular-science magazines, as those articles are considered reliable and accurate sources for people to base their decisions on. Personal experience means knowledge people acquire through first-hand or second-hand experience. For example, they could learn from their friends' experience, and base their decisions on that.

Numeracy

Numeracy refers to the ability to process probabilities and numerical concepts; making good decisions in daily life requires numerical ability (Peters, Västfjäll, Slovic, Mertz, Mazzocco & Dickert, 2006). Individuals with high numeracy are more likely to make use of numerical principles to remain less susceptible to framing effects (e.g., labeling pork as "25% fat" or "75% lean") and, thus, are more accurate in decision making than low numeracy individuals (Peters et al., 2006). Studies in the health field also found that numeracy was a unique predictor of probability judgment tasks related to medical decision making (Låg, Bauger, Lindberg, & Friborg, 2014). Lower numeracy individuals were more likely to trust verbal risk information from physicians (i.e., physicians' qualitative description of the risk) than numeric risk information, which was identical with verbal risk information, but included numerical information (e.g., fractions or percentages), whereas higher numeracy individuals were more likely to trust numeric risk information than verbal risk information

(Gurmankin, Baron, & Armstrong, 2004). Higher numeracy was also found to be associated with more accurate judgment about probabilities related to prostate cancer screening (Hamm, Bard, & Scheid, 2003).

In the political decision-making field, Kahan, Peters, Dawson, and Slovic (2017) also found that higher numeracy participants were more likely to perform better on data-based decision making. However, their performance depended on context. In their study, two sets of data interpretation problems were given to participants: one context was about gun control, which is a polarizing political issue, whereas the other context was about a skin cream, which was relatively neutral. Participants were found to interpret the two sets of data differently, even though they were identical. In the skin cream context, high numeracy individuals were more likely to interpret the data correctly; however, in the gun control context, high numeracy individuals were more likely to use their numeracy skills selectively to interpret the data in a way that aligned with their own political outlooks. Although numeracy is critical in people's data-based decision making, it seems to be used selectively rather than consistently.

Popular-Science Reports

Popular-science reports are defined as articles written by science journalists or scientists themselves with the general public as the target reader. Common examples include articles from magazines such as *Scientific American* and *Psychology Today*. Popular-science reports can be an important source for the general public to learn about scientific findings across all disciplines. Scientific findings should be a reliable source for people to base decisions on. However, people do not seem to make decisions primarily based on scientific findings (Kahan et al., 2017). Stanovich and West (2007) found that participants who smoked were less likely to acknowledge the negative health influences of second-hand smoke, and participants who consumed more alcohol were less likely to admit the consequences of alcohol consumption. It is very likely that most participants were exposed to the scientific knowledge about smoking and drinking alcohol, as they were all undergraduate students. Nonetheless, Kahan et al. (2017) also found that even in the face of compelling scientific evidence, the public can still show science-blindness when making empirical decisions related to politically-charged policies.

Personal Experience

Another source for people to base their decisions on is their personal experience. People might trust their own experiences more than scientific findings. People anchor their beliefs to their own experience and perspective, which biases their decision making (Epley, Keysar, Van Boven, & Gilovich, 2004). In the revolutionary paper by Tversky and Kahneman (1974), they demonstrated that people make decisions using anchoring and adjustment heuristics. In one of their experiments, participants were asked about the percentage of African counties in the United Nations. For each participant, a wheel of fortune with numbers from 0 to 100 was spun in front of them. They then were asked if the real number was higher or lower than the percentage spun. After that, participants were asked to estimate the real number by moving upwards or downwards the numbers on the wheel. The results demonstrated that their estimations were strongly influenced by the random number initially spun on the wheel. Therefore, people were likely to anchor their decisions to an initial value and produced final estimations that were not too far away from this initial starting point, despite the fact that this initial value was random and meaningless.

The tendency to be egocentrically biased is evident in early childhood (Epley, Keysar, Van Boven, & Gilovich, 2004). For example, children are likely to say, "The cup is to your right" when the cup is actually to your left, but to their right. Adults are less egocentrically biased, yet do not outgrow this tendency; they still anchor their beliefs to their own perspectives. For example, people might think their partners can easily read their minds (i.e., that their internal state is transparent to their partner) whereas it actually is not (Epley, et al., 2004). They know their own internal state and believe their partner knows it too.

Borrowing from anchoring and adjustment heuristics (Tversky & Kahneman, 1974), Epley et al. (2004) demonstrated that the assessment of another's perspective was a process of substituting one's own perception and adjusting as needed. In addition, the adjustment was insufficient in that it stopped once a plausible conclusion was reached, which was still egocentrically biased. As their study showed, participants' assessments of others' perspectives were still skewed in an egocentrically-biased direction even after adjustments were made. Therefore, people's personal experience is likely to influence their data-based decision making.

Current Study

The purpose of the current research was to determine whether these three variables, numeracy, scientific reports in the popular press, and personal experience were associated with people's data-based decision making. In previous research, the relations between general cognitive ability or general intelligence and data-based decision making have been studied (Låg et al., 2014; Stanovich & West, 2007; Stanovich, West, & Toplak, 2013), but few studies (Kahan et al., 2017; Peters et al., 2006; Reyna, Nelson, Han, & Dieckmann, 2009) specifically examined the relation between numeracy and data-based decision making. There is a research base of studies on data-based decision making in specific areas, such as politics and health (Kahan et al., 2017; Reyna et al., 2009), with few studies examining decision making in more general life domains (Peters et al., 2006). The current research filled these literature gaps by examining numeracy and data-based decision making in a more general life domain – studying – and also included popular science reports and personal experience as potential influences on data-based decision making.

We chose the topic of *whether cramming helps students do better/worse on exams* for the data-based decision making task, which is arguably more neutral than previously studied political topics such as gun control. The correct interpretation in the data-based decision making task was manipulated across participants (i.e., cramming was helpful/cramming was not helpful). Moreover, it was possible to manipulate both the position of the scientific report (i.e., cramming was helpful/ not helpful) and the outcome of the simulated personal experience related to cramming. Popular-science reports were operationalized as researcher-crafted *Psychology Today* articles. Personal experience was operationalized as a simulated five-minute cramming experience with "feedback" (not of true performance, but fake feedback crafted by the researcher), which indicated that cramming was helpful/ not helpful.

We predicted that numeracy would be associated with data-based decision making, such that individuals with higher numeracy skills would typically make more accurate decisions. For popular-science reports, if the report has an effect on data-based decision making, it is expected to bias participants' interpretation of the data in the direction of the popular-science conclusion. If this hypothesis is correct, then when the popular-science report says cramming is helpful, participants will be more likely to interpret subsequent data as showing that cramming is helpful; when the popular-science report says cramming is not helpful, participants will be more likely to interpret data as showing that cramming is not helpful. Similarly, if personal experience has an effect on data-based decision making, then participants' accuracy on the data-interpretation task should be influenced by the "feedback" they receive on the mock test. Therefore, we predicted that there would be interactions between popular-science reports and correct interpretation of data and between personal experience and correct interpretation of data such that participants would be more accurate when the correct interpretation of the data aligned with personal experience and the popular-science report findings.

Method

Participants

Data were collected from 297 participants recruited from Amazon Mechanical Turk (MTurk) for this study. The eligibility requirements were that participants lived in North America and were at least 18 years old. Participants were compensated $0.50 for participation. Data from 110 participants were removed due to participants' failing attention checks. The remaining sample for analysis included 187 participants (Mean age = 39.1 years, range = 22-75 years, 101 females).

Design

The design of the experiment was a 2 (popular-science argument: helpful, not helpful) × 2 (personal cramming experience: helpful, not helpful) × 2 (correct data interpretation: helpful, not helpful) between-subjects factorial design with accuracy on the data-interpretation task (correct or incorrect) as the dependent variable.

Measures

Abbreviated Numeracy Scale (Weller et al., 2013). The Abbreviated Numeracy Scale measures individual differences in numeracy skills. It consists of eight items in short-answer format (e.g., Imagine that we roll a fair, six-sided die 1000 times. Out of 1000 rolls, how many times do you think the die would come up as an even number?). The Cronbach's α is .71. For the current study, the context of one question was changed to be about strep throat rather than cancer at the request of the Research Ethics Committee. The number of questions correctly answered for each participant was used as his or her numeracy score. This measure of numeracy was selected to allow us to compare our results to those of Kahan et al. (2017).

Popular-Science Report. The instructions stated that the purpose of this task was to learn from a scientific report whether cramming was helpful on exams. Participants were randomly assigned to read one of two versions of a simulated *Psychology Today* article. One version presented the argument that cramming is helpful for students' performance on exams, i.e., "In most situations research has made it abundantly clear that massed trials greatly improves retention if the test follows immediately". The other version presented the argument that cramming is not helpful for students' performance on exams, i.e., "In most situations research has made it abundantly clear that spacing the learning over many shorter sessions is much more effective than trying to do it all in one big session".

Personal Cramming Experience. The instructions stated that the purpose of this task was to give participants a simulated cramming experience and see if cramming helped them to perform better on exams. Participants first were asked to read a one-page article on the topic of Confucianism for five minutes. The article was then removed from view. Participants were asked to answer four multiple-choice questions based on the content of the article, e.g., "For what period did the Shang Dynasty last? A. 551 BCE–479 BCE, B. 206 BCE–220 CE, C. 1600 BCE–1046 BCE, D. 1046 BCE–256 BCE". Following the test, participants were given "feedback" on their performance; however, the researcher manipulated this feedback. Participants were randomly assigned to get feedback that their cramming was helpful on the exam (feedback = 100% accuracy on test) or that their cramming was not helpful on the exam (feedback = 25% accuracy on test).

Data-Interpretation Problem. The data-interpretation measure was adapted from Kahan, Peters, Dawson, and Slovic (2017) by changing the topic to cramming. Participants were shown a description of an experiment testing the effectiveness of cramming and the results were displayed in a 2 × 2 contingency table. Participants were asked to indicate "whether the experiment shows that cramming is likely to make students do better or worse on an exam". Two different tables (See Figure 1 and 2) were created: the cell numbers across the two tables were held constant, but the correct interpretation of the data (i.e., cramming was helpful/cramming was not helpful) was manipulated by changing the column labels (did better/did worse). Participants were randomly assigned to receive data correctly interpreted as crammers did better or did worse.

Mark / People	Did Better on Test	Did Worse on Test
Students who did cram	223	75
Students who did not cram	107	21

Figure 1. Correct data interpretation- cramming not helpful.

Mark / People	Did Worse on Test	Did Better on Test
Students who did cram	223	75
Students who did not cram	107	21

Figure 2. Correct data interpretation- cramming helpful.

Procedure

The entire study took place on Amazon Mechanical Turk using the Qualtrics survey platform. Participants were first asked to complete the Abbreviated Numeracy Scale and provide demographic information including sex and age. Participants were then randomly assigned to one of two orders: personal cramming experience followed by reading popular-science report or vice versa. Finally, participants completed the data-interpretation measure, following which they were shown a debriefing screen that explained the true purpose of the study and, importantly, that psychological research strongly suggests cramming is not an effective studying method. Lastly, participants obtained a code for Amazon Mechanical Turk compensation.

Results

A binary logistic regression was conducted, predicting accuracy on the data-interpretation task from the independent variables (numeracy, personal cramming experience, popular-science argument), 2-way interaction terms with data-interpretation (numeracy × correct data-interpretation, personal cramming experience × correct data-interpretation, popular-science argument × correct data-interpretation), and demographic variables (age, gender).

Table 1: Regression Coefficients for Final Model.

	β	Wald	df	p	Odds Ratio
Numeracy	.435	7.795	1	.005	1.544
Age	.027	4.594	1	.032	1.027
Popular-science × data-interpretation	1.772	7.497	1	.006	5.883
Constant	19.559	.000	1	1.000	

Regression results indicated that the overall model of three predictors (numeracy, age, and the interaction between popular-science argument and correct data interpretation) was statistically reliable in determining accuracy on the data-interpretation task, -2 Log Likelihood = 229.508, $\chi^2(9) = 28.524$, $p = .001$, $R^2 = .189$. No other predictors were significant. The model correctly classified 67.9% of the cases. Moreover, each of the three predictors accounted for significant unique variance in data-interpretation. Regression coefficients for the final model are presented in Table 1. For numeracy, participants were 1.54 times more likely to correctly answer the data-interpretation question for each additional correct answer on the numeracy test, β = .435, $p = .005$. For age, for each year older, participants were 1.03 times more likely to correctly answer the data-interpretation question, β = .027, $p = .032$. The effect of the correct data interpretation was different across popular-science arguments, β = 1.772 , $p = .006$, odds ratio = 5.88.

To explore the data interpretation × popular-science interaction term, chi-squared tests of association were performed. There was neither a main effect of popular-science report on accuracy on the data-interpretation task, $x^2(1) = 0$, $p = .984$; nor a main effect of correct data-interpretation on accuracy on the data-interpretation task, $x^2(1) = 1.43$, $p = .232$ As shown in Figure 3, there was a significant qualitative interaction between popular-science report and correct data-interpretation on accuracy on the data-interpretation task. When the popular-science report said that cramming was helpful, participants were more likely to accurately solve the data-interpretation problem when the correct data-interpretation was that cramming was helpful (i.e., when the science report and correct data-interpretation were consistent) than not helpful, $x^2(1) = 8.40$, $p =.004$. When the popular-science report said that cramming was not helpful, there was no significant difference across the correct data-interpretation conditions, $x^2(1) = 1.26$, $p = .262$.

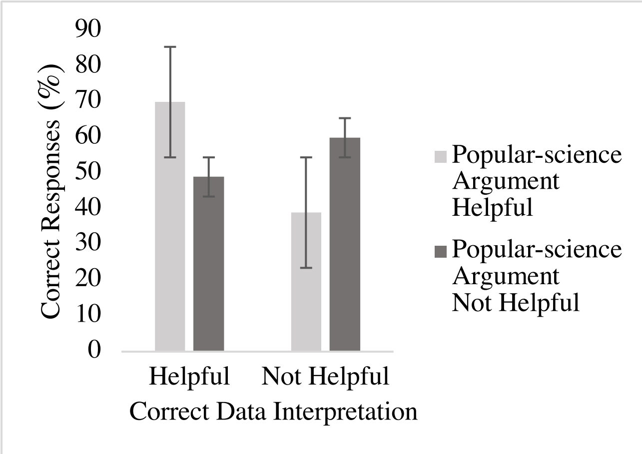

Figure 3. Percentage of participants who made the correct response on the data-interpretation question as a function of correct data interpretation and popular-science argument. Error bars represent the standard error of each condition.

Discussion

We found that participants with higher numeracy skills were more likely to make the correct data-based decision. This finding is consistent with past literature in health and politics domains (Hamm, Bard, & Scheid, 2003; Låg, Bauger, Lindberg, & Friborg, 2014; Kahan, Peters, Dawson, & Slovic, 2017). We also found that older participants were more likely to make the correct data-based decision than younger adults. This finding is consistent with research showing that younger adults have lower numeracy levels than older adults (LeFevre et al., 2014). Finally, we found that the interaction term of popular-science report x correct interpretation of data predicted the likelihood of making the correct data-based decision.

The effect of the correct data interpretation was different across popular-science arguments. When the popular-science report said that cramming was helpful, participants were significantly more accurate when the correct data-interpretation was also that cramming was helpful. In contrast, when the popular-science report said that cramming was not helpful, this same positive effect of congruency between the popular-science report and correct data interpretation was not evident. In other words, when the popular-science report said that cramming was helpful, more participants were using their numeracy skills to confirm the popular-science argument and decide that cramming was helpful.

We did not find a significant interaction between personal cramming experience and correct interpretation of data. This finding is consistent with that of Kahan et al. (2017), which suggests that personal experience might not be a predictor of data-based decision making on more neutral topics such as cramming. However, our finding might also be due to the ratio between personal experience and exposure to popular-science reports. In people's daily lives, they will likely have had multiple personal experiences with cramming for exams, but have read fewer scientific findings on the topic. In our experiment, we only simulated a one-time experience for cramming. The ratio of personal experience and popular-science reports was 1:1; whereas, in real life, the ratio is much larger than 1:1. This could contribute to weak power. However, in the debriefing, no participants indicated that they had not believed the feedback provided in the personal cramming experience.

We found that numeracy predicted success in data-based decision making, but that people did not use their numeracy skills consistently across all contexts. Instead, participants seemed to use numeracy selectively. The finding that people were more likely to interpret the data correctly when both the popular-science argument and the correct interpretation said that cramming was helpful did not reflect an overall reliance on scientific findings. If participants used their numeracy skills consistently to confirm science, or showed a positive effect of congruency between the popular-science report and the data before them, they should also have been more likely to interpret the data correctly when both the popular-science argument and the correct data interpretation said that cramming was not helpful. However, this was not the case. One might instead argue that participants tried to use their numeracy skills consistently, but allowed their own calculations to be overruled by scientific reports when the two results were in conflict. However, if participants were deferring to the science report, when the report and correct data-interpretation were in conflict, then participants should be as likely to decide incorrectly that cramming was not helpful in the conflicting conditions, which they were not. Moreover, if participants used their numeracy skills to evaluate the data, but did not base their decision on that calculation, or if they failed to engage their numeracy skills at all in making their decision, then numeracy should not emerge as a predictor of accuracy, which it did. Therefore, it is possible that participants used their numeracy skills selectively to confirm their own desire, in that they wanted cramming to be effective.

Consistent with this interpretation, some participants spontaneously expressed their desire for cramming to be helpful during the debriefing stage of the study. Thus, our results suggest that participants may selectively interpret data to align with their own desired outcome. If it is participants' desire that cramming be effective, then the selective numeracy that we observed may reflect confirmation bias, which refers to "the tendency to search for, interpret, favor, and recall information in a way that confirms one's pre-existing beliefs or hypotheses" (Plous, 1993, p. 233). Another term, myside bias, can be regarded as a subclass of confirmation bias that may be especially applicable here (McKenzie, 2004). As Stanovich et al. (2013, p. 259) put it, myside bias refers to the phenomenon wherein "people evaluate evidence, generate evidence, and test hypotheses in a manner biased toward their own prior beliefs, opinions, and attitudes".

Past research has explained participants' selective use of numeracy by reference to System 1 and System 2 thinking

(Kahan et al., 2017). In the Kahan et al. (2017) study, the authors noted that higher-numeracy participants were more likely to use System 2 thinking to calculate the percentage of each condition in order to obtain the correct answer, whereas lower-numeracy participants were more likely to use heuristics, or System 1 thinking, and jump to incorrect interpretations. However, participants used System 2 thinking selectively, depending on the context of the problem.

The current experiment used the efficacy of cramming as a study technique as the topic of decision making. Cramming is likely more relevant to university students than MTurk workers, so we are currently conducting a replication study with a university student sample. If the pattern of results from the current MTurk sample holds, we might expect a more pronounced interaction between popular-science argument and the correct data interpretation in student samples. Thus, when the popular-science report says that cramming is not helpful, there may be no significant difference across the correct data-interpretation conditions (consistent with the current findings). However, when the popular-science report says that cramming is helpful, students may be even more accurate than our MTurk sample when the correct data-interpretation is congruent (cramming is helpful), as students might have a stronger desire for cramming to be effective.

Conclusion

In conclusion, participants with higher numeracy were more likely to make the correct data-based decision. However, participants used their numeracy selectively; they may have been more likely to use their numeracy skills to confirm their own desired outcome. No significant association was found between personal experience and data-based decision making. People should develop more awareness of their "selective numeracy" and make use of their numeracy skills when making important data-based decisions.

Acknowledgements

This research was supported by a grant from King's University College at Western University to MP.

References

Baron, J. (2008). *Thinking and deciding* (4th ed.). New York, NY, US: Cambridge University Press.

Chen et al., (2020). Efficacy of hydroxychloroquine in patients with COVID-19: Results of a randomized clinical trial. medRxiv. https://doi.org/10.1101/2020.03.22.20040758

Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology*, *87*(3), 327–339. doi:10.1037/0022-3514.87.3.327

Griffiths, T. L., & Tenenbaum, J. B. (2006). Optimal predictions in everyday cognition. *Psychological science*, *17*(9), 767-773. doi:10.1111%2Fj.1467-9280.2006.01780.x

Gurmankin, A.D., Baron, J., & Armstrong, K. (2004). The effect of numerical statements of risk on trust and comfort with hypothetical physician risk communication. *Medical Decision Making*, *24*, 265–271.

Hamm, R.M., Bard, D.E., & Scheid, D.C. (2003, November). *Influence of numeracy upon patient's prostate cancer screening outcome probability judgments*. Paper presented at the annual meeting of the Society for Judgment and Decision Making, Vancouver, British Columbia, Canada.

Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion: Psychological studies of opinion change.* New Haven, Connecticut, US: Yale University Press.

Kahan, D. M., Peters, E., Dawson, E., & Slovic, P. (2017). Motivated numeracy and enlightened self-government. *Behavioural Public Policy*, *1*(1), 54–86. doi:10.1017/bpp.2016.2

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

Låg, T., Bauger, L., Lindberg, M., & Friborg, O. (2014). The role of numeracy and intelligence in health-risk estimation and medical data interpretation. *Journal of Behavioral Decision Making*, *27*(2), 95–108. doi:10.1002/bdm.1788

LeFevre, J., Penner-Wilger, M., Pyke, A., Shanahan, T., Deslauriers, W. A., Trbovich, P., & Roberts, M. A. (2014). Putting two and two together: Declines in arithmetic fluency among young Canadian adults. *Carleton University Cognitive Science Technical Report*, 1. Retrieved from http://www.carleton.ca/ics/TechReports.

McKenzie, C. R. M. (2004). Hypothesis testing and evaluation. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment & decision making* (pp. 200–219). Malden, MA: Blackwell.

Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological Science*, *17*(5), 407–413. doi:10.1111/j.1467-9280.2006.01720.x

Plous, S. (1993). *The psychology of judgment and decision making*. New York, NY: Mcgraw-Hill

Reyna, V. F., Nelson, W. L., Han, P. K., & Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making. *Psychological Bulletin*, *135*(6), 943-973. doi:10.1037/a0017327

Stanovich, K. E., & West, R. F. (2007). Natural myside bias is independent of cognitive ability. *Thinking and Reasoning*, *13*(3), 225–247. doi:10.1080/13546780600780796

Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology*, *94*(4), 672–695. doi:10.1037/0022-3514.94.4.672

Stanovich, K. E., West, R. F., & Toplak, M. E. (2013). Myside bias, rational thinking, and intelligence. *Current Directions in Psychological Science*, *22*(4), 259–264. doi:10.1177/0963721413480174

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131. doi: 10.1126/science.185.4157.1124

Weber, E. U. (2006). Experience-based and description-based perceptions of long-term risk: Why global warming does not scare us (yet). *Climatic Change*, *77*, 103–120 doi: 10.1007/s10584-006-9060-3

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. doi:10.1002/bdm.1751

Xia, B. S., & Gong, P. (2015). Review of business intelligence through data analysis. *Benchmarking*, *21*(2), 300-311. doi:10.1108/BIJ-08-2012-0050