DEVIATIONS FROM BAYES' THEOREM DURING BELIEF UPDATING IN YOUNGER
AND OLDER ADULTS: EVIDENCE FROM BEHAVIOUR AND NEURAL ACTIVITY

by

Bonnie Andrea Armstrong

M. A., Psychology, Ryerson University, 2015

B. A. (Honours), Psychology, University of Waterloo, 2013

A dissertation

presented to Ryerson University

in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

in the program of

Psychology

Toronto, Ontario, Canada 2019

©Bonnie Andrea Armstrong, 2019

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this dissertation. This is a true copy of the dissertation, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this dissertation to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this dissertation by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my dissertation may be made electronically available to the public.

**Deviations from Bayes' Theorem During Belief Updating in Younger and Older Adults:
Evidence from Behaviour and Neural Activity**

Doctor of Philosophy, 2019

Bonnie Andrea Armstrong

Psychology, Ryerson University

Abstract

Updating prior information with new information in accordance with Bayesian principles is a difficult task. Younger adult decision makers deviate from Bayes' theorem by either overweighting prior information (i.e., using a conservatism heuristic) or overweighting new information (i.e., using a representativeness heuristic) on decision tasks without feedback. Similar to younger adults, older adults make decisions that require belief updating. Given age-related decrements in cognitive control, older adults may be at a disadvantage compared with younger adults when updating beliefs. Prior research shows no age differences when making decisions under risk, however older adults perform worse than younger adults when making decisions under ambiguity. Currently it is unknown how older adults use heuristics when updating beliefs about risk and ambiguous information compared with younger adults.

The primary aim of this dissertation was to examine age-related differences in the use of heuristics during belief updating, as well as the cognitive processes and neural correlates that underpin behaviour. In three experiments, younger and older adults completed a belief updating task with and without feedback using an urn-ball paradigm. The main results showed that both younger and older adults committed the representativeness error more than the conservatism error, with no age differences observed when updating beliefs without feedback but with younger adults updating beliefs more accurately than older adults with feedback. Further, age differences in the neural correlates that underlie belief updating showed evidence that older adults recruit

additional resources in frontal regions of the brain to facilitate performance compared with younger adults. Event-related potentials showed evidence of cognitive control in response to conflicting information in both age groups, but a diminished neural response to feedback in older compared with younger adults. Additionally, while younger adults were not influenced by ambiguous information, older adults avoided committing the representativeness error only when new information was ambiguous. Last, individual differences in numeracy and cognitive reflection, but not thinking disposition, modulated belief updating performance. Together, the results show that younger and older adults can learn to update beliefs with feedback but with younger adults learning to a greater degree than older adults, especially when information is ambiguous.

Acknowledgements

I would first like to express my sincere gratitude to my supervisor, Dr. Julia Spaniol, for her continuous support throughout my PhD. Her expertise and breadth of knowledge helped to formulate both the research questions and methodology of this dissertation. In addition, Dr. Spaniol's patience and encouragement allowed me to grow as a research scientist in a challenging and supportive environment. My PhD experience has been extremely positive and fulfilling, and for this I thank Dr. Spaniol for affording me this opportunity. I would also like to thank my internal committee members, Drs. Ben Dyson and Todd Girard for their thorough reviews of this dissertation and thoughtful comments. I would especially like to thank Carson Pun for his invaluable help with numerous tasks across the series of experiments such as his assistance with programming, script writing and ERP analysis. I would also like to thank Dr. David Flora for his help writing the multi-level modeling script and for our discussions concerning analysis and interpretation.

I would like to thank Natalie Ein, a fellow graduate student, who has grown with me academically. I am extremely grateful that we experienced graduate school together and can look back on fond memories of struggle and success. Additionally, I would like to thank my parents, Debbie and Ian Armstrong, for their support and whose good examples have taught me to work hard for the things that I aspire to achieve. I would also like to thank my good friend Carly Pearce, who has been a constant source of support and encouragement during the challenges of graduate school. Last, a most sincere thank you to my beloved husband, Tim Rossy, whose patience, continued support of my work, and shared interest in research has helped make graduate school that much more enjoyable. He has shared this experience with me over the last six years, so it only seems right that I dedicate this dissertation to him.

Table of Contents

Abstract	iii
List of Tables	ix
List of Figures	xi
List of Appendices	xiii
List of Abbreviations	xiv
Chapter 1: General Introduction	1
Chapter 2: Literature Review	10
Bayesian Inference.....	10
Heuristic-Based Decision Making	12
Cognitive and Neurophysiological Changes with Age	14
Age Differences in Decision Making	17
Age Differences in Bayesian Inference	21
Dual-Process Theory: Cognitive Architecture of Information Processing	24
Reinforcement Learning	30
Decisions Under Risk and Ambiguity	33
Individual Differences in Cognitive Reflection	36
Chapter 3: Synthesis of Dissertation Goals	42
Dissertation Experiments	44
Chapter 4: General Method	46
Participants.....	46
Materials	48
Belief updating task.	48

Experimental Design.....	61
Data Analysis	62
Chapter 5: Experiment 1.....	65
Hypotheses.....	69
Method	71
Results.....	74
Discussion	91
Chapter 6: Experiment 2.....	99
Hypotheses.....	111
Method	116
Results.....	125
Discussion	166
Chapter 7: Experiment 3.....	178
Hypotheses.....	183
Method	185
Results.....	193
Discussion	221
Chapter 8: General Discussion	227
Prevalence of Base-Rate Neglect.....	230
Cognitive Architecture of Information Processing	237
Decisions Under Risk vs. Ambiguity.....	242
Limitations	246
Future Directions	248

Conclusions.....	249
Appendices.....	251
References	294

List of Tables

Table 1. Posterior Odds for all Decision Situations.....	54
Table 2. Colour Counterbalancing of Urn Distributions	56
Table 3. Hypotheses for Experiment 1	70
Table 4. Characteristics of the Final Sample in Experiment 1	72
Table 5. Means and Standard Deviations for Accuracy in Experiment 1	75
Table 6. Self-Assessment for Experiment 1	84
Table 7. Multi-Level Model Comparison on Accuracy in Experiment 1	86
Table 8. Berlin Numeracy Test as a Predictor of Accuracy in Experiment 1	88
Table 9. Hypotheses for Experiment 2	115
Table 10. Characteristics of the Final Sample in Experiment 2	117
Table 11. Means and Standard Deviations for Accuracy in Experiment 2	126
Table 12. Self-Assessment for Experiment 2	135
Table 13. Multi-Level Model Comparison on Accuracy in Experiment 2	137
Table 14. Berlin Numeracy Test as a Predictor of Accuracy in Experiment 2	139
Table 15. Lipkus Numeracy as a Predictor of Accuracy in Experiment 2	141
Table 16. Cognitive Reflection Test-Reflection as a Predictor of Accuracy in Experiment 2 ...	143
Table 17. Hypotheses for Experiment 3	185
Table 18. Characteristics of the Final Sample in Experiment 3	187
Table 19. Means and Standard Deviations for Accuracy in Experiment 3	194
Table 20. Self-Assessment for Experiment 3	207
Table 21. Multi-Level Model Comparison on Accuracy in Experiment 3	209
Table 22. Lipkus Numeracy as a Predictor of Accuracy in Experiment 3	211

Table 23. Cognitive Reflection Test-Reflection as a Predictor of Accuracy in Experiment 3 ...	214
Table 24. Faith in Intuition as a Predictor of Accuracy in Experiment 3	217

List of Figures

Figure 1. Sample trial of the representativeness heuristic conflicting with Bayes' theorem in Experiment 1	53
Figure 2. Accuracy proportion for each decision situation across trial block in Experiment 1	77
Figure 3. Reaction time for each decision situation across trial block in Experiment 1	81
Figure 4. Association of accuracy in conflict situations in Experiment 1 for younger adults	83
Figure 5. Association of accuracy in conflict situations in Experiment 1 for older adults	83
Figure 6. Schematic of the 64-channel International 10/20 system	118
Figure 7. Sample trial of the representativeness heuristic conflicting with Bayes' theorem in Experiment 2	121
Figure 8. Accuracy proportion for each decision situation across trial block in Experiment 2..	128
Figure 9. Reaction time for each decision situation across trial block in Experiment 2	131
Figure 10. Association of accuracy in conflict situations in Experiment 2	133
Figure 11. LRP: Grand averaged ERPs from C3-C4, time-locked to urn distribution onset	147
Figure 12. N2: Grand averaged ERPs from FCz and Fz, time-locked to sample onset	153
Figure 13. pP: Grand averaged ERPs from AF3, AFz, and AF4, time-locked to sample onset.	154
Figure 14. P3b: Grand averaged ERPs from Pz and Cz, time-locked to sample onset	158
Figure 15. FRN: Grand averaged ERPs from Cz, time-locked to sample onset	163
Figure 16. Sample trial of the representativeness heuristic conflicting with Bayes' theorem in Experiment 3	192
Figure 17. Accuracy proportion for each decision situation and condition in Experiment 3	200
Figure 18. Reaction time for each decision situation and condition in Experiment 3	203
Figure 19. Association of accuracy in Experiment 3 in CPUS condition for younger adults	204

Figure 20. Association of accuracy in Experiment 3 in UPCS condition for older adults	204
---	-----

List of Appendices

Appendix I: Consent Form for Experiment 1	251
Appendix II: Debriefing Form for Experiment 1	255
Appendix III: Consent Form for Experiment 2.....	257
Appendix IV: Debriefing Form for Experiment 2	261
Appendix V: Consent Form for Experiment 3.....	263
Appendix VI: Debriefing Form for Experiment 3	267
Appendix VII: Experiment 1 Instructions.....	269
Appendix VIII: Experiment 2 Instructions	271
Appendix IX: Experiment 3 Instructions	273
Appendix X: Posterior Probabilities and Corrected Odds of Decision Situations.....	276
Appendix XI: Demographic Questionnaire	277
Appendix XII: Self-Assessment for Experiments 1 and 2.....	278
Appendix XIII: Self-Assessment for Experiment 3.....	280
Appendix XIV: R Script Used for Multi-Level Modeling.....	283
Appendix XV: MATLAB Script Used for Signal Processing and ICA	286
Appendix XVI: ERP Latency Results.....	292

List of Abbreviations

ANOVA	Analysis of variance
CconfB	Conservatism heuristic conflicts with Bayes' theorem
DASS-21	21-item Depression Anxiety and Stress scale
DSST	Digit Symbol Substitution Test
EEG	Electroencephalography
EOG	Electrooculogram
ERP	Event-related potential
fMRI	Functional magnetic resonance imaging
FRN	Feedback-related negativity
ICA	Independent components analysis
LRP	Lateralized readiness potential
PANAS	Positive and Negative Affect Schedule
PFC	Prefrontal cortex
pP	Prefrontal positivity
PPV	Positive predictive value
RalignB	Representativeness heuristic aligns with Bayes' theorem
RconfB	Representativeness heuristic conflicts with Bayes' theorem

Chapter 1: General Introduction

Imagine being at home and the 10-year-old carbon monoxide (CO) detector alerts dangerous CO levels. One must decide whether to evacuate one's home or not on the basis of a) knowing that high levels of CO can be poisonous and severely affect one's health, and b) the probability of the alarm being false given that it is 10 years old. To complicate matters, imagine it is the summer, the windows are open in the house, the furnace is not running, and you do not have a gas stove (i.e., a false alarm is plausible). By conducting further research, you discover that even older CO detectors have low false alarm rates (i.e., a CO leak is plausible). On balance, this information makes it difficult to decide whether to evacuate your house or not. In order to judge the posterior probability (i.e., the likelihood of a leak given that the CO alarm went off), prior beliefs (i.e., CO leak in your house) must be updated with new information (i.e., the CO alarm going off).

Judging posterior probabilities requires Bayesian inference. Bayes' theorem is used as the formal standard for belief updating, in which one must update the odds of an event in light of new evidence (Bayes & Price, 1763; Knight, 1921). Although formal Bayesian statistics involve calculation, in the real world abiding by the conceptual logic of Bayes' theorem can help avoid errors in probabilistic judgment. However, research shows that following the conceptual principles of Bayes' theorem is a difficult task (Gigerenzer, Gaissmaier, Kurz-Milcke, Schwartz, & Woloshin, 2007; Goodie & Fantino, 1999; Kahneman & Tversky, 1972; Simon, 1972). Instead, simpler strategies that promote fast and efficient information processing such as heuristics are commonly used in real world settings to judge probabilities (Grether, 1980, 1992; Kahneman & Tversky, 1972).

Heuristics are general decision making strategies that ignore part of the available information, diminish the work of retrieving and storing information in memory, and are based on minimal need for information search or analytic evidence evaluation (Gigerenzer, Todd, & ABC Research Group, 1999). Heuristics streamline the decision making process by reducing the amount of integrated information in order to make decisions quickly, frugally and without a heavy dependence on cognitive resources compared with complex decision making strategies (Gigerenzer & Gaissmaier, 2011). For example, when walking to work one may notice a construction worker hauling up a pallet of bricks on a pulley. Likely a snap judgment would be made to choose to walk around the construction area instead of walking directly under the bricks to avoid danger. This decision would be made without assessing the entire situation such as calculating the probability of the bricks falling while walking under the pallet or the chances of survival if the bricks fell. That is, a heuristic would likely be used to make a quick decision without the cost of much mental effort.

Although heuristics speed up decision making processes, they can introduce errors and biased judgments (Todd, Gigerenzer, & ABC Research Group, 2011). Heuristics lead to error in judgment particularly when the heuristic is not suited for the structure of the environment, such as when the problem at hand requires the application of normative strategies (e.g., applying the conceptual principles of Bayes' theorem; Gigerenzer et al., 1999). In fact, errors in judgment are commonly observed when heuristics are used to make Bayesian inferences (Savage, 1954). One such error involves overweighting prior information in which a base-rate only strategy is used. This heuristic used to make decisions has been termed conservatism (Edwards, 1968), and relates to the investment and commitment to prior beliefs (Klayman, 1995), also found in the anchoring phenomenon (Mussweiler & Strack, 1999). A second error involves overweighting new

information in which a base-rate neglect strategy is used, a decision making heuristic termed representativeness (Grether, 1980, 1992; Kahneman & Tversky, 1972). The representativeness heuristic has been described as making a comparison between new information and a parent population (Juslin, 2015; Juslin, Nilsson, & Winman, 2009; Juslin & Persson, 2002), engaging a “matching intuition” (Glöckner & Witteman, 2010).

There is a vast literature showing that younger adults have difficulty making Bayesian inferences because they tend to base decisions on either prior or new information rather than the product of their integration (Grether, 1980, 1992; Juslin, Nilsson, Winman, & Lindskog, 2011; Kahneman & Tversky, 1972). Like younger adults, older adults are also confronted with decisions that require Bayesian inference. It has been postulated that older adults may be required to make Bayesian inferences that inform important decisions more often than younger adults, such as in medical and financial contexts that require informed decision making based on probabilistic information (Chen & Sun, 2003). Although research shows that older adults use heuristics to make decisions more than younger adults (Blanchard-Fields, Hertzog, Stein, & Pak, 2001; Hess, 2001; Peters, Hess, Auman, & Västfjäll, 2007), much less is known about how age affects the use of heuristics during Bayesian inference.

More generally, in the heuristics and biases literature cognitive reflection tasks are commonly used to examine the prevalence of making judgments based on automatic, intuitive, fast and effortless processing (i.e., Type 1 processing) compared with analytical, slow and effortful processing that burdens executive cognitive resources (i.e., Type 2 processing; Kahneman, 2011). Typically, cognitive reflection tasks are formatted as word problems and are deliberately designed to pit an automatic response generated by Type 1 processing against a normative response generated by Type 2 processing (Kahneman, 2011). In these tasks, intuitive

answers that are automatically generated by Type 1 processing are typically suboptimal responses (although there are exceptions; see Evans & Stanovich, 2013). In order to make a correct judgment, a conflict must be detected between the automatic response and a normative response (Stanovich, 2011). Inhibitory control is then engaged to suppress and override Type 1 processing, in order to allow Type 2 processing to actively simulate hypothetical, alternative responses that may lead to a more optimal response than the response generated by Type 1 processing (Evans, 2003, 2007, 2010; Evans & Stanovich, 2013; Oldrati, Patricelli, Colombo, & Antonietti, 2016). Importantly, simulating alternative responses must not be confused with the real world. Therefore, decoupling operations must be sustained to avoid confusion (Stanovich, 2011). Simultaneously suppressing automatic responses while generating alternative responses is largely demanding of executive functions, specifically inhibitory control (Best, Miller, & Jones, 2009; Miyake & Friedman, 2012) and working memory (Feldman Barrett, Tugade, & Engle, 2004). In the context of a belief updating task in which Bayes' theorem is the formal standard, Type 1 processing may automatically generate a conservatism or representativeness heuristic leading to an error in judgment. However, if a conflict is detected between using a conservatism or representativeness heuristic and a normative strategy, such as the principles of Bayes' theorem, Type 2 processing may be engaged to simulate alternative responses that may lead to a more optimal response.

Within the aging literature, there is consensus regarding the general pattern of age-related changes in cognition, with a linear decline particularly in working memory and executive functions such as cognitive control (Salthouse, 2019). Cognitive control is an umbrella term that describes the processes that are involved in controlling or managing other cognitive processes such as maintaining and updating information in working memory, sustaining attention when

confronted with interference, as well as inhibitory control (Diamond, 2013; Gassaley & Nobre, 2012; Lenartowicz, Kalar, Congdon, & Poldrack, 2010). Age differences on behavioural measures of cognitive control have been associated with the anatomical and functional deterioration of the prefrontal cortex (PFC) with age (Braver & Barch, 2002). In particular, the inhibitory deficit hypothesis proposes that age-related declines in inhibitory efficiency due to weakened cognitive control with age, underlies the deficits observed in attentional and working memory tasks (Hasher & Zacks, 1988). Given the deficits in cognitive control in old age, older adults may perform worse than younger adults on tasks that require inhibiting automatic, heuristic-based responses in order to simulate alternative responses that follow normative principles, a process that involves working memory. However, how older adults use heuristics relative to normative strategies to update beliefs compared with younger adults has never been empirically examined.

As previously described, cognitive reflection tasks usually present a few word problems with no feedback. Research shows age differences in reinforcement learning, with younger adults using feedback more successfully than older adults (Eppinger, Kray, Mock, & Mecklinger, 2008; Fernandes et al., 2018; Nieuwenhuis et al., 2002). Some research attributes this age discrepancy to a decline in dopamine with age (see Bäckman, Nyberg, Lindenberger, Li, & Farde, 2006; Braver & Barch, 2002 for reviews), a neurotransmitter pivotal for learning from feedback (Mathewson, Dywan, & Segalowitz, 2005). Other research suggests that older adults may have difficulty learning from feedback compared with younger adults due to an age-related asymmetry in the processing of valence, such that older relative to younger adults show a preference for positive over negative information (Eppinger et al., 2008; Mather & Carstensen, 2005; Mather &

Johnson, 2000). Whether feedback would promote the use of normative strategies in younger and older adults over automatically generated heuristics when updating beliefs is unknown.

In addition, cognitive reflection tasks typically require making decisions under risk, in which all relevant information is available and can be calculated (Volz & Gigerenzer, 2012). Bayesian inference in the real world requires updating beliefs based on ambiguous information, with some information unavailable or unattainable (Volz & Gigerenzer, 2012). Referring to the CO leak example described above, although one could research the false-alarm rate of a CO detector that is 10 years old, the probability of a CO leak is unknown. Older adults show an aversion to ambiguity (Tymula et al., 2013), possibly due to a deficit in making inferences about ambiguous information (Hämmerer et al., 2019) or a deficit in understanding how to assign weight to ambiguous information when making decisions (Nassar et al., 2016). The literature is currently mixed regarding whether there are age differences in decisions made based on ambiguous information. Some research shows that younger adults make more accurate decisions under ambiguity than older adults (Hämmerer, et al., 2019; Nassar et al., 2016). However, other research shows that younger adults have difficulty making decisions based on ambiguity and, similar to older adults, show an aversion to ambiguous information leading to decision errors (Seer, Lange, Boos, Dengler, & Kopp, 2016). Whether there are age-related differences when making Bayesian inferences based on risk information compared with ambiguous information is unknown.

Beyond age differences, there is a voluminous literature on the individual differences associated with cognitive reflection. Specifically, the Cognitive Reflection Test (Frederick, 2005) has previously been used to measure conflict detection and index the tendency to default to autonomous processing. In order to detect conflict between an automatic and normative

response, or to generate an alternative response that is better than the response generated by automatic processing, one must have the “mindware” to support these processes such as numeracy or probabilistic reasoning skills (Clark, 2001; Perkins, 1995). Additionally, research shows that cognitive reflection is modulated by thinking dispositions (i.e., intuitive versus reflective thinking styles) that stem from one’s goals, attitudes toward forming and changing beliefs, as well as one’s willingness to problem solve (Baron, Scott, Fincher, & Metz, 2015; Bruine de Bruin, McNair, Taylor, Summers, & Strough, 2015; Cacioppo, Petty, Feinstein, & Jarvis, 1996). Research shows that higher Cognitive Reflection Test and numeracy scores, and reflective thinking dispositions relate to higher decision accuracy on cognitive reflection tasks (Cacioppo & Petty, 1982; Cokely et al., in press; Pacini & Epstein, 1999; Toplak, West, & Stanovich, 2011; Tversky & Kahneman, 1983). At present however, it is unclear whether feedback would close the gap that would otherwise exist between those of high versus low numeracy, or those with more intuitive versus reflective thinking dispositions. Additionally, no research has examined whether individual differences in numeracy and thinking dispositions modulate performance on a belief updating task when decisions are based on risk information compared with ambiguous information.

Examining the use of heuristics when normative strategies are required, such as in the case of Bayesian inference, is an important research endeavor with societal implications. For instance, work from my research program in the domain of medical decision making shows that when asked to judge the probability of having a disease given a positive test result (i.e., the positive predictive value; PPV), younger and older adults tend to make estimates that reflect high probabilities when the true probability is low. These results reflect heuristic-based judgments such that responses reflect the general belief that diagnostic tests are highly accurate

in lieu of considering all relevant information, that if integrated correctly would point towards a low probability (Armstrong & Spaniol, 2017; Armstrong, Sparrow, & Spaniol, under review). This argument is strengthened by results reported by Wegier, Armstrong, and Shaffer (2019) who showed that when asked to judge the PPV, high probability judgments were made even when a low PPV was explicitly presented (i.e., the answer was provided!). In the real world, erroneously judging the PPV to be high when it is actually low, as in the case of maternal serum screening for Down syndrome which has a PPV of 2.2%, may have negative effects on patients such as stress or overtreatment (Wegier et al., 2019).

Beyond medical decision making, errors in probability judgment have also been found in pathological gambling (e.g., Toplak, Liu, Macpherson, Toneatto, & Stanovich, 2007), as well as in financial decision making resulting in problems of money management and economic life planning (e.g., Lusardi & Mitchell, 2014). Examining behaviour, as well as the neural underpinnings of belief updating amongst younger and older adults will shed new light on which processes support Bayesian inference. Interpreting data through the lens of dual-process theory of information processing, as well as the cognitive aging literature will also elucidate where there may be a breakdown in the information processing system in older age that results in using erroneous heuristics to make judgments.

Dissertation Overview

In the heuristics and biases literature, it is well-documented that younger adults default to Type 1 over Type 2 processing leading to the use of heuristics in lieu of normative strategies during Bayesian inference (Grether, 1980, 1992; Kahneman & Tversky, 1972). Although older adults use heuristics to make decisions more than younger adults (Blanchard-Fields et al., 2001; Hess, 2001; Peters et al., 2007), whether age affects the use of heuristics when heuristics conflict

with Bayes' theorem has never been empirically tested. Further, whether there are age or individual differences in belief updating when feedback is presented, or when information is based on risk relative to ambiguity, are avenues of research that have never been empirically explored. In light of the gaps in the literature described, the overarching goal of this dissertation is to investigate how age and individual differences modulate sensitivity to the principles of Bayesian inference by pitting normative strategies (i.e., Bayes' theorem) against error-prone heuristics (i.e., conservatism or representativeness) when feedback is provided compared with when feedback is not provided. Further, I examine the neural correlates that underlie decisions to elucidate the processes that take place before decisions are made.

The primary goal of Chapter 2 is to orient the reader to the literature that is most relevant to the topic of the dissertation and to build the rationale for the series of experiments presented in the ensuing chapters. First, I will introduce the concept of Bayesian inference while highlighting some of my own work on this topic and describe the literature on the role heuristics play in decision making. Building on this foundation, I will then summarize the literature on how age may influence Bayesian inferences, as well as heuristic-based decisions. I will also briefly introduce the dual-process account of information processing and decision making. In addition, I will summarize the research on how age impacts reinforcement learning, as well as how age affects decisions made under risk compared with decisions made under ambiguity. Last, I will describe the research that has examined individual differences in cognitive reflection.

Chapter 2: Literature Review

Bayesian Inference

As briefly mentioned in Chapter 1, Bayes' theorem can be used as a normative strategy of optimization for updating beliefs (Bayes & Price, 1763; Knight, 1921; Stigler, 1983, 1986). In order to generate a posterior judgment, prior information is combined with new information using the following formula:

$$P(H|D) = \frac{P(H) P(D|H)}{P(H) P(D|H) + P(\sim H) P(D|\sim H)}$$

In this formula “H” is the hypothesis under investigation, and “D” represents the data or new information that is relevant to the hypothesis. $P(H)$ reflects the probability that the hypothesis is true prior to acquiring new information, and $P(\sim H)$ is the probability that the alternative hypothesis is true prior to having acquired new information. The posterior probabilities in the formula include, $P(H|D)$ which represents the probability that the hypothesis is true given new information, $P(D|H)$ which represents the probability of observing new information given that the hypothesis is true, and $P(D|\sim H)$ which represents the probability of observing new information given that the alternative hypothesis is true.

It has been well documented that Bayesian inference is a difficult task (Gigerenzer et al., 2007; Goodie & Fantino, 1999; Grether, 1980; Kahneman & Tversky, 1972). Even with the support of visual aids used to alleviate or simplify mental computation such as graphs, diagrams, or icon arrays (Garcia-Retamero & Cokely, 2013; Garcia-Retamero & Galesic, 2010), successful use of Bayes' theorem is uncommon, and can require high levels of graph literacy and numeric ability (Galesic & Garcia-Retamero, 2011). It is important to note however that in the heuristics and biases literature, when errors in Bayesian inference are made, this does not reflect errors in calculation, but instead reflects large qualitative errors concerning the conceptual logic of the

rules of probability. It is rare that an individual would know or have access to Bayes' theorem. Without access to explicit probabilities, Bayesian inferences are derived from beliefs about states of the world. Tasks that require Bayesian inference index whether individuals understand the principles of Bayes' theorem, such as attending to the appropriate variables and integrating relevant information that approximate responses derived through normative strategies. In this sense, Bayes' theorem is used as the formal standard of belief updating. Because people have difficulty following the principles of Bayes' theorem, systematic violations are commonly observed (Fiedler, 2000; Garcia-Retamero, Galesic, & Gigerenzer, 2010; Grether, 1980; Kahneman & Tversky, 1972).

To provide an example, below is the classic cab problem used in seminal studies of Bayesian inference (Bar-Hillel, 1980; Tversky & Kahneman, 1982).

A cab was involved in a hit-and-run accident at night. Two cab companies, the Green and the Blue, operate in the city in which the accident occurred. There are a total of 85% Green cabs in the city, and 15% Blue cabs. A witness of the accident identified the cab as Blue. The court tested the reliability of the witness under the same conditions as the night of the accident by presenting the witness with a sample of cabs (half of which were Blue and half of which were Green), with the witness correctly identifying the colour of the cabs 80% of the time and erring 20% of the time. What is the probability (ranging from 0% to 100%) that the cab involved in the accident was Blue rather than Green?

Bayes' theorem provides a normative strategy for integrating relevant information in order to obtain an optimal answer. The equation below describes how a Bayesian inference could be made based on the information described in the cab problem:

$$P(H|D) = \frac{P(H) P(D|H)}{P(H) P(D|H) + P(\sim H) P(D|\sim H)}$$

$$P(H|D) = \frac{(.15) (.80)}{(.15) (.80) + (.85)(.20)} = .41$$

Without access to the formula however, a solution to the cab problem can be reasoned as follows (Stanovich, West, & Toplak, 2016): 15% of the cabs are Blue and the witness, who is 80% accurate, identified the cab as Blue. Given this information, it is known that in 100 accidents involving cabs, 15 were Blue and the witness would identify 80% of the cabs as Blue (12/15). On the other hand, out of 100 accidents involving cabs, 85 were Green and the witness would identify 20% of the cabs as Blue (17/85). Therefore, the witness would identify 29 (12 + 17) cabs as Blue, but only 12 would actually be Blue. That is, 12 out of 29 cabs (41%) that were identified as Blue, were actually Blue. Thus, it is more likely that the accident involved a Green cab (59%).

Previous research shows that the most common answer to the cabs problem is ~80%, reflecting the witness' accuracy of identifying the cab without integrating prior probability information (the base-rate probability of 15%), resulting in an overestimation of individual-case evidence (Bar-Hillel, 1980; Hamm, 1993). The systematic error of using the probability of the witness' identification accuracy exemplifies the use of simple strategies to make decisions in lieu of normative strategies in a context that requires formal analysis, a common observation due to the difficulty of Bayesian inference (Kahneman, 2003).

Heuristic-Based Decision Making

Heuristics are simple strategies used to make judgments and decisions (Gigerenzer, & Gaissmaier, 2011). They are short-cuts or rules of thumb that do not follow normative methods of problem solving (Albar & Jetter, 2009). As briefly mentioned in Chapter 1, although

heuristics are usually effective and lead to good decisions (Gigerenzer, Hertwig, & Pachur, 2011), there is a wealth of literature showing that they can also lead to errors in judgment if used in an environment that does not support simple cognitive strategies (Gigerenzer et al., 1999). That is, heuristics become the source of systematic bias when the environment is structured for the use of probability theory (Goldstein & Gigerenzer, 2002). When making probabilistic judgments, heuristics are commonly used to save effort such that the computational load is minimized at the expense of accuracy (Gigerenzer et al., 2011).

Deviating from the principles of Bayes' theorem can result in two errors: overweighting prior information reflecting use of a conservatism heuristic (i.e., a base-rate only rule; Edwards, 1968) and overweighting new information reflecting use of a representativeness heuristic (i.e., a base-rate neglect rule; Kahneman & Tversky, 1972). When deciding whether an object belongs to Population A or to Population B, the conservatism heuristic is applied if the decision is based on the relative frequency of A and B without regard for the object. The conservatism heuristic demonstrates a commitment to prior beliefs (Klayman, 1995), also found in the anchoring phenomenon (Mussweiler & Strack, 1999). Suboptimal decisions result when only prior information is considered in situations that require the integration of information, such as Bayesian inference. In the cabs problem for example, a conservatism error is committed when the decision maker judges the probability that the cab in the accident was Blue as 15%, without regard for new information (i.e., 80% accuracy of witness identification) when making a decision.

In contrast, the representativeness heuristic reflects the overweighting of new information in which a base-rate neglect rule is applied. Using a similar example as the above, when deciding whether an object belongs to Population A or to Population B, the representativeness heuristic is

applied if the decision is based on the degree to which the object seems representative of Population A or Population B, and the relative frequencies of A and B are ignored (i.e., the decision is based on a “matching intuition”; Glöckner & Witteman, 2010). Critically, it is not wrong to consider how similar an object is to a parent population, but only considering representativeness does not account for prior probabilities (Barbey & Sloman, 2007; Kahneman & Tversky, 1973; Tversky & Kahneman, 1974). As previously mentioned, a common error committed by decision makers in the cabs problem involves using the witness’ identification accuracy (i.e., 80%) as the probability that the cab in the accident was Blue, without regard for prior information (i.e., 15% of the cabs in the city are Blue) when making a decision.

The voluminous literature on heuristics and biases has been built on the younger adult population. Bayesian inference is required to make important decisions particularly in the domains of medicine and finance (Chen & Sun, 2003; Worthy, Gorlick, Pacheco, Schyner, & Maddox, 2011). The prevalence of medical and financial decisions increase with age (Chen & Sun, 2003; Worthy et al., 2011). However, normal age-related cognitive decline may affect how decisions are made compared with decisions made by those who have not undergone cognitive decline.

Cognitive and Neurophysiological Changes with Age

The cognitive aging literature shows that multiple cognitive domains are affected by age, including declines in working memory (De Neys & Verschueren, 2006), processing speed (Salthouse, 1996), executive function (Mather & Knight, 2005), inhibitory control (Hasher, Lustig, & Zacks, 2007) and sustained attention (McDowd & Craik, 1988) – cognitive processes that place a heavy demand on cognitive control (Braver & Barch, 2002). In particular, the inhibitory deficit hypothesis proposes that older relative to younger adults are more susceptible

to interference effects and have more difficulty inhibiting information that is irrelevant to one's goal (Hasher & Zacks, 1988). Baddeley's (1996) characterization of the function of working memory refers to the capacity to sustain selective attention while inhibiting the disrupting effects of task-irrelevant information. It has been argued that the decline in working memory capacity commonly observed in old age is related to the increase in distractibility due to age-related declines in inhibitory control (Salthouse, 2019). In addition to the declines in cognitive control mechanisms, fluid intelligence declines markedly with age after peaking in the 20s or 30s, (Cattell, 1971; Hartshorne & Germine, 2015; Horn, 1970; Horn & Cattell, 1967; Horn & Donaldson, 1976; Salthouse, 2019). Fluid intelligence refers to the ability to reason and problem solve in novel contexts (Cattell, 1971; Gray, Chabris & Braver, 2003; Sternberg, 1985), and relates to metacognition (i.e., reflecting about one's own ongoing mental processes; Sternberg, 1985), working memory and attentional control (Kane & Engle, 2002). Fluid intelligence is commonly indexed using speeded measures such as the Digit Symbol Substitution Test (Wechsler, 1997), and shows a reliable slowing in processing speed with age (Salthouse, 2019).

In addition to cognitive changes, subcortical and cortical neural circuits undergo structural, functional, and neuromodulatory changes as a function of healthy aging (Rajah, et al., 2009). For example, overall brain volume decreases with age, and an attenuation of white-matter tract connectivity particularly in the frontal lobes (Madden et al., 2012), a brain region associated with executive function and strategic planning (Fjell & Walhovd, 2010). The PFC theory of aging (Park, 2000; West, 1996) argues that goal-oriented functions of the PFC (e.g., integrating information, executive control, inhibiting distracting or interfering information) are most susceptible to aging due to the neurophysiological changes that occur in this region. Further, some neurotransmitter systems, such as the dopaminergic system associated with reward

processing (i.e., learning from feedback) and motivation, also undergo decline with age (Erixon-Lindroth et al., 2005).

Although the research described above paints a “doom and gloom” picture of cognitive aging, age-related cognitive decline is not uniform across cognitive domains despite brain-level changes associated with age. That is, some aspects of cognition are spared from age-related decline (Cabeza, Nyberg, & Park, 2005). Cognitive functions that remain intact or improve with age include speech and language (Bopp & Verhaeghen, 2005), semantic memory (St-Laurent, Abdi, Burianova, & Grady, 2011), as well as crystallized intelligence (Cattell, 1971; Hartshorne & Germine, 2015; Horn, 1970; Horn & Cattell, 1967; Horn & Donaldson, 1976; Salthouse, 2019; Strough, Karns, & Schlosnagle, 2011). Crystallized intelligence refers to knowledge of language, information and concepts specific to a culture that have been acquired over time (Horn & Cattell, 1967; McGrew, 2009), and is commonly indexed by vocabulary measures such as the Mill Hill Vocabulary Test (Raven, 1982) or general information measures that peak in late life (Salthouse, 2019). Additionally, changes in functional brain networks have been shown to compensate for neural degeneration (Cabeza et al., 2005), preserving certain cognitive abilities. For example, studies using neuroimaging have found an increase in bilateral brain activation in older adults for tasks that activate unilateral brain regions in younger adults (Grady, Bernstein, Beig, & Siegenthaler, 2002; Reuter-Lorenz et al., 2000). The increase of bilateral activation with age has been interpreted as a compensatory activity, in which the aging brain reorganizes regional recruitment in order to compensate for age-related losses (Cabeza et al., 2018; Grady, 2012).

In summary, aging is associated with neurophysiological changes of the PFC (West, 1996) which has been linked to declines observed in cognitive control processes (Braver &

Barch, 2002) and fluid intelligence (Salthouse, 2019), as well as neurochemical changes such as the decline in dopamine function with age, a system important for learning from feedback (Erixon-Lindroth et al., 2005). In contrast, general knowledge acquired over time (i.e., crystallized intelligence) is preserved or improves with age (Salthouse, 2019).

Age Differences in Decision Making

It has been projected that the number of older adults over the age of 80 will double by 2030 in North America (Carstensen & Hartel, 2006). As the global population ages, older adults will be required to take more responsibility for making decisions that concern their physical, psychological and financially well-being. The majority of older adults are unaffected by severe pathology such as dementia and are living at home well into the eighth decade of life (Qiu & Fratiglioni, 2018). Maintaining independence and well-being in old age requires the ability to make good decisions (Worthy et al., 2011). Further, the complexity and frequency of important decisions, especially within the domains of finance and medicine, increases with age (Banks & Oldfield, 2007; Chen & Sun, 2003; Worthy et al., 2011). While big decisions, such as choosing insurance coverage, have obvious long-term consequences, small everyday decisions also have a cumulative impact on quality of life (Worthy et al., 2011).

Analytical processing of information to make decisions declines with age (Braver & Barch, 2002; Hanoch, Wood, & Rice, 2007). This is especially the case when tasks are complex and cognitively demanding (Yoon, Cole, & Lee, 2009). For example, Finucane, Mertz, Slovic, and Schmidt (2005) investigated the association between age and decision quality by varying the complexity of the task. As task complexity increased (i.e., number of choices increased), decision errors increased, with older adults showing more difficulty making decisions than younger adults. In addition, older adults have been shown to seek less information to make

decisions (Mather, 2006), apply simpler decision rules (Bruine de Bruin, Parker, & Fischhoff, 2007, 2009; Johnson, 1990), are more likely to be swayed by changes in the framing of decision problems (Bruine de Bruin et al., 2007, 2009), and are less able to control the impact of automatic processing when making judgments (Hess, Waters, & Bolstad, 2000). For example, older adults have shown to exhibit larger belief-bias effects in syllogistic reasoning tasks (Gilinsky & Judd, 1994), such that they are less able to judge the logical validity of arguments independently of their prior knowledge about the truth of the premise and conclusion.

Notably, an important skill that can underly the quality of decisions is numeracy (Peters et al., 2006). Numeracy has been defined as the “micro-strategies” that comprise one’s understanding and ability to assign meaning to mathematical concepts (Nelson, Reyna, Fagerlin, Lipkus, & Peters, 2008; Peters, 2012). Previous research shows that old age is correlated with worse performance on numeracy measures and numerical decision tasks, even after accounting for age-related differences in education and fluid abilities (Bruine de Bruin, Parker, & Fischhoff, 2012; Bruine de Bruin et al., 2009; Finucane et al., 2005; Weller et al., 2013). However, Bruine de Bruin and colleagues (2015) recently showed that motivation to problem solve, measured by the Need for Cognition scale, mediated the negative relationship between age and numeracy, suggesting that older adults may show low numeracy due to lack of motivation.

Although cognitive aging can negatively impact decision quality when making decisions that require analytical processing, older adults’ knowledge and experience seem to benefit their decisions in familiar life situations enabling them to avoid biased decisions. Previous research has shown that older adults are better than younger adults at discontinuing unprofitable investments reflecting a reduced susceptibility to sunk-cost effects with age (Bruine de Bruin et al., 2009). Through experience older adults may have learned when to adhere to sunk-cost rules,

which reduced reliance on analytical processes such as fluid abilities in order to avoid sunk-cost decision biases (Stanovich & West, 2008). Additionally, Tentori, Osherson, Hasher, and May (2001) reported that in a grocery-store context, older adults were less likely than younger adults to be influenced by a dominated option. The authors argued that older adults' greater experience with a grocery-store context (i.e., more experience purchasing groceries) may have helped them discount irrelevant information (also see Kim & Hasher, 2005 for similar results). Additionally, Mata, Schooler and Rieskamp (2007) reported that older adults performed similarly to younger adults on a task that required learning conditional probabilities to make a decision. Although analysis of behavioural choices revealed no age differences, younger and older adults applied different strategies to make decisions, with older adults using more heuristic-based strategies compared with younger adults. This suggests that reliance on heuristics may play a compensatory role for the aging decision maker.

Fuzzy-trace theory is another dual-process account suggesting that with age, decisions made are increasingly based on the "gist" or bottom line meaning of information, rather than basing decisions on verbatim details (Reyna, 2004). Similarly, Yates and Patalano (1999) posit an age-related shift from processing information analytically, to following a rule-based process in which a general rule is extracted from cumulative experience, to a final automatic style of making decisions in which judgment and decision rules become habitual. Moreover, Mata and Nunes (2010) conducted a meta-analysis that showed that older adults are more likely to use heuristic-based strategies to make decisions (e.g., seek less information to make a decision compared with younger adults), further supporting the prevalence of using heuristics when making decisions in late life.

Although the increase in reliance on experience and heuristic-based approaches to making decisions in old age may reflect compensation for cognitive decline, it may also reflect a developed skill for making decisions that have been optimized over the course of the lifespan. Liu, Wood and Hanoch (2015) investigated whether older adults' accumulated crystallized intelligence could provide an alternate route for making good decisions in order to compensate for declines in fluid intelligence. Results support this idea by showing that older adults performed as well or better than younger adults on a variety of decision making tasks.

Importantly, not all empirical evidence supports the notion that older adults use heuristics to make decisions. Delaney, Strough, Parker and Bruine de Bruin (2015) used a cluster-analytic approach to investigate whether rational, intuitive, spontaneous, dependent, and avoidant styles of decision making combined to form specific decision making profiles amongst younger and older adults. Results showed that older adults were less likely to belong to an experiential decision style cluster that was defined by rapid decisions based on heuristics and were more likely to belong to a cluster defined by a controlled, independent decision style. In light of this, it cannot be argued that older adults are universally more likely to use heuristic-based strategies to make decisions.

In summary, older adults may be able to make good decisions if the decision requires experiential processing in familiar settings, or if the gist of information can be extracted in order to make a decision (Bruine de Bruin et al., 2009). However, if a decision requires analytical processing drawing on fluid intelligence or cognitive flexibility, decisions made in late life may be poor (Weller et al., 2013). Although younger adults have difficulty making Bayesian inferences, literature concerning how age affects the accuracy of decisions that require Bayesian

inference is scarce. Moreover, how heuristics are used to make Bayesian inferences in old age is unknown.

Age Differences in Bayesian Inference

Although most research has investigated Bayesian inference in younger adults, older adults are also confronted with risk information on which to base decisions. For instance, with age there is an increase in making medical decisions such as choosing medications or opting for surgery (Galesic, Gigerenzer, & Straubinger, 2009), as well as financial decisions such as retirement planning, selling one's house, and organizing one's will, all of which involve consideration of risk information (Chen & Sun, 2003). Evidence concerning whether age affects Bayesian inference is mixed. For example, Fiske (2005) employed a Bayesian inference task and showed that older adults performed better on two of three measures of Bayesian inference compared with younger adults, with higher working memory capacity and fluid intelligence correlating with lower decision accuracy for younger adults' only (Fisk, 2005). The correlation observed suggests that working memory and fluid intelligence may underlie performance amongst younger adults, however older adults may not have relied on these mechanisms to make judgments due to age-related changes in cognition (Nyberg, Lovden, Riklund, Lindenberger, & Backman, 2012). Although counterintuitive, similar findings have been reported in previous research. For example, Stanovich and West (2000) showed that those with higher intellectual ability (i.e., operationalized through fluid intelligence and academic aptitude measures) produced less normative responses than those with lower intellectual ability. Those who are highly intellectual may have applied analytical strategies that led their judgments astray (e.g., when an incorrect normative rule is applied to make decisions such as incorrectly combining base-rate and diagnostic information).

A study conducted by Spaniol & Bayen (2005) examined younger and older adults' conditional probability judgments by presenting participants with fictitious patients with one of two diseases and one of eight symptoms. The amount of time patient information was presented on the screen varied (i.e., 1 s, 2.5 s, or 5 s). Subsequent to encoding patient information, participants were presented with a multiple-choice quiz presenting patient information and were asked to choose the disease or symptom of the last patient they saw by choosing one of the response options. Results showed age-related differences in conditional probability judgments with older adults showing stronger judgment biases compared with younger adults. The results of this study were attributed to an age-related deficit in memory encoding, with age differences emerging when demands on memory were high (shorter encoding duration), and age differences disappearing when memory demands were reduced (longer encoding duration).

A study conducted by Galesic and colleagues (2009) showed that, regardless of numeracy level, younger and older adults made more accurate Bayesian inferences when information was presented in a natural frequency format that preserved base-rate information (e.g., “48 out of 50”) compared with a conditional probability format that presented single-event probabilities (e.g., “96%”). Younger and older adults showed more accurate Bayesian inferences in the former relative to the latter format. Although a difference in accuracy was observed between formats, judgments were low in accuracy overall (i.e., even in a natural frequency format, about 45% of younger and older adults made erroneous judgments).

Building on the study by Galesic and colleagues (2009), my Master's thesis showed that accurate Bayesian inferences were made when information was experienced compared with when information was described in a verbal summary, regardless of age or numeracy level (Armstrong & Spaniol, 2017). The experience format presented a slideshow of patient cases to

participants in which absolute frequency information was presented. Joint frequency information regarding each patients' a) health status (i.e., actually having a disease or not) and b) test result (i.e., a positive or negative test result) was presented for each patient. In contrast, the description format presented a verbal summary of relative frequency information that could be used to make a Bayesian inference (i.e., the same natural frequency format used by Galesic et al., 2009).

Whereas the experience format produced judgments that approximated correct posterior probabilities, the description format produced judgments that reflected the use of heuristics (i.e., making judgments that reflected probabilities explicitly provided in the summary such as the sensitivity of the diagnostic test). Making Bayesian inferences as if using heuristics have also been observed in younger adults in other experiments that comprise my research program (Armstrong, Sparrow, & Spaniol, under review; Wegier, Armstrong, & Shaffer, 2019), as well as expert populations, such as medical residents (Armstrong, Spaniol, & Persaud, 2018).

In summary, the literature on aging and Bayesian inference is mixed, with some studies showing that older adults make more accurate judgments than younger adults (Fiske, 2005), some studies showing that older adults make less accurate judgments than younger adults (particularly when demand on memory is high; Spaniol & Bayen, 2005), and some studies showing no differences between age groups (Armstrong & Spaniol, 2017; Galesic et al., 2009). For the purpose of this dissertation, it is important to have a conceptual understanding of the cognitive architecture of information processing when confronted with conflict, in order to examine belief updating when error-prone heuristics conflict with Bayes' theorem. The following section provides a summary of the processes that underlie cognitive reflection through the lens of dual-process theory.

Dual-Process Theory: Cognitive Architecture of Information Processing

The dual-process theory of rational decision making postulates two processes that underlie decisions (Evans, 2008; Osman, 2004). One process reflects an experiential or affective mode of processing information, in which decisions are effortless, fast, automatic, heuristic-based, and are typically based on intuition and experience (i.e., Type 1 processing; Epstein, 1994; Kahneman, 2003). The other process reflects a more deliberative and analytical style of processing, and is described as effortful, conscious, time consuming, and demanding of cognitive resources (i.e., Type 2 processing; Kahneman, 2011). Both modes can be activated simultaneously and are interdependent, such that analytical processes use products of experiential processes as input (Kahneman, 2003) and experiential processes use products of analytical processes as input (Liu, Wood, & Hanoch, 2015).

Prominent characteristics of Type 1 processing include automatic execution, detachment from high-level control systems, low dependence on central processing capacity, associative processing, with automatic processes working in parallel (Kahneman, 2011). Type 1 processing underlies emotion regulation, face recognition, frequency estimation, and implicit learning with little variance in these processes across individuals (Barrett & Kurzban, 2006; Carruthers, 2006; Evans, 2008, 2009; Moors & De Houwer, 2006). In contrast, Type 2 processing is slow, demanding of high-level control systems, computationally expensive, and processes information serially (Kahneman, 2011). Type 2 processes are recruited in situations of importance such as when making financial, employment or medical decisions.

Although Type 1 processing provides an approximated response to solving problems, Type 2 processing is engaged when detailed analysis of a problem is required. Although a rational response can be generated by Type 1 processing if normative principles are overlearned

(Stanovich, 2011; Stanovich et al., 2016), engaging Type 1 processing can be costly when normative strategies have not been overlearned and Type 2 processing is required (Arkes & Ayton, 1999). In situations in which Type 1 processing produces a suboptimal response, a conflict might be detected between the automatic response generated by Type 1 processing and a normative response. Type 2 processing can then be engaged to suppress and override Type 1 processing (Stanovich, 2011). In order to override Type 1 processing, Type 2 processing must be able to interrupt the automaticity of Type 1 processing requiring inhibitory control (Best, Miller, & Jones, 2009; Hasher, Lustig, & Zacks, 2007; Miyake & Friedman, 2012). Once Type 1 processing has been suppressed, hypothetical alternative responses can then be generated by Type 2 processing, with the goal of generating a better response than the response generated by Type 1 processing. The generation of alternative responses is derived through hypothetical reasoning or “cognitive simulation” (Evans, 2007, 2010; Evans & Stanovich, 2013).

Cognitive simulations allow one to mentally test alternative responses in order to gauge the fit of the response to the situation and the decision maker’s goals (Stanovich, 2011). Importantly however, representations of the real world must not be confused with representations generated from cognitive simulations. Sustaining separation between real and simulated representations is a process distinct to Type 2 processing that has been termed “cognitive decoupling operations” (Stanovich, 2011). Decoupling operations are taxing and must be continually enforced while generating hypothetical responses. Executive functions such as working memory capacity have been shown to predict one’s ability to sustain decoupling operations (Feldman Barrett, Tugade, & Engle, 2004), with fluid intelligence highly correlated with executive functioning and Type 2 processing (Burgess, Gray, Conway, & Braver, 2011;

Chuderski, 2014; Duncan et al., 2008; Hicks, Harrison, & Engle, 2015; McVay & Kane, 2012; Salthouse, Atkinson, & Berish, 2003).

To summarize, when Type 1 processes produce an automatic and suboptimal response, a conflict between this automatic response and a normative response may be detected (i.e., conflict detection). Type 2 processes may then suppress and override Type 1 processes (i.e., suppression and override). Subsequently, Type 2 processes may then generate and simulate alternative responses, while also decoupling simulated and real representations of response options.

Override operations for Type 2 processing require both procedural (inhibiting Type 1 processing) and declarative knowledge (generating alternative responses based on stored knowledge; Stanovich et al., 2016). Detecting a conflict between automatic and normative responses, as well as generating alternative responses that are better than responses generated automatically requires proficient knowledge (Stanovich et al., 2016). The knowledge, strategies, and rules acquired from experience that are used for decoupling operations have been termed “mindware” (Clark, 2001; Perkins, 1995). There is large variance in the quality of mindware across individuals because this process is based on specialized knowledge derived from experience in domains such as probabilistic, causal, and scientific reasoning, as well as numeracy (Stanovich et al., 2016)¹. As previously mentioned, just as Type 1 processing can lead to a normative response, Type 2 processing can lead to a normatively incorrect response due to errors in knowledge or “contaminated mindware”. An example of contaminated mindware is the well-documented gambler’s fallacy, in which one believes the probability of an event is lowered when that event recently occurred, even though the probability of the event is independent from trial to

¹ Of note, crystallized intelligence as traditionally measured is nonspecialized, and does not relate to mindware (Stanovich et al., 2016).

trial (Jarvik, 1951). Although Type 2 processing is engaged in this example, the knowledge of the rules of probability are incorrect. Importantly, no task purely separates problems in processing from problems in mindware (Stanovich et al., 2016). However, inferences can be made regarding what the task may depend on more. The Cognitive Reflection Test (discussed more later), for example, depends more on conflict detection and override, whereas numeracy tasks depend more on proficient mindware.

Control mechanisms of Type 2 processing have been differentiated into two positively correlated components: algorithmic and reflective processing (Stanovich, 2009). Different from algorithmic processing that reflects *what* cognitive mechanisms underlie behaviour such as computational processing, reflective processing reflects *why* processes were engaged such as one's goals and beliefs that give reason to behaviour. Algorithmic processing can be assessed by fluid intelligence measures and cognitive ability such as numeracy tasks, whereas reflective processing can be assessed by thinking dispositions that capture belief structure, attitudes towards beliefs and goals, such as the Need for Cognition or Faith in Intuition scales (Cacioppo & Petty, 1982; Cacioppo, Petty, & Feinstein, 1996; Epstein, Pacini, Denes-Raj, & Heier, 1996; Pacini & Epstein, 1999). Individual differences in thinking dispositions reflect variance in goal management, values and self-regulation (Stanovich & West, 1998).

It is important to note that simulating alternative responses does not mean all alternative responses are generated. A cognitive simulation may for instance be inflexible and associative, and thus generate an alternative response that is easily accessible without constructing a different response that may challenge the first response generated (i.e., partial cognitive simulation; Stanovich, 2009). Alternative responses are usually simple, reflect one's beliefs or have previously been generated and are easily accessible (Stanovich et al., 2016). An example of this

phenomenon is when heuristics are used to make decisions about probabilistic information (e.g., base-rate neglect; Kahneman & Tversky, 1973). This concept has been termed serial associative cognition with a focal bias (Stanovich, 2009). Therefore, although there is a general preference to process information automatically, when situations require Type 2 processing such as in novel contexts, decision makers commonly default to serial associative processing with a focal bias. For instance, framing effects show that the way information is framed is taken as focal and subsequent processing is based on how that information was framed with alternative ways of framing information left unexplored. Reflective processing has been shown to initiate the interruption of serial associative cognition by either triggering a new or more comprehensive serial associative process in which more alternative responses are generated (Stanovich et al., 2016).

In summary, when Type 1 processes produce a suboptimal response, a conflict may be detected between an automatic and normative response. Detecting a conflict can be initialized by reflective processing, which can be captured by thinking disposition measures such as the Need for Cognition or Faith in Intuition scale (Bruine de Bruin et al., 2007; Toplak et al., 2011, 2014a, 2014b), as well as measures of cognitive reflection such as the Cognitive Reflection Test (Frederick, 2005). Reflective processing also initiates the override of Type 1 processes, requiring inhibitory control to suppress automatic processes so that Type 2 processes can be engaged (Toplak et al., 2011). Generating alternative responses that are superior to the automatic response heavily depends on proficient mindware such as numeracy. Sustaining decoupling operations while simultaneously overriding Type 1 processes and engaging Type 2 processes to generate responses is supported by algorithmic processes and is largely demanding of executive functions, particularly inhibitory control and working memory (Stanovich, 2011). Additionally, rather than

exhausting all possible alternative responses, decision makers typically default to serial associative cognition with a focal bias, a phenomenon commonly observed when heuristics are used to make decisions when formal analysis is required.

Dual-process theory: Base-rate neglect. The representativeness heuristic has been shown to be a strong influencer of probabilistic judgment (Barbey & Sloman, 2007; Brainerd, 2007; Kahneman & Tversky, 1973; Tversky & Kahneman, 1974). For instance, if a male U.S. citizen was randomly sampled and described as being shy and introverted, prior research shows that it is more prevalent to judge him to be a librarian than a farmer, even though there are 20 times more farmers in the United States than librarians (Kahneman, 2011). Committing a representativeness error (i.e., using a base-rate neglect strategy) has been found to be more prevalent than committing a conservatism error (i.e., using a base-rate only strategy; see Barbey & Sloman, 2007, for review).

Dual-process theorists have argued that base-rate neglect occurs because individual-case evidence cues an intuitive response that reflects more concrete and tangible information that is vivid and difficult to override (Kahneman & Frederick, 2002; Kahneman & Tversky, 1973; Tversky & Kahneman, 1974). In contrast, prior probabilities may require some degree of slower, analytical processing (Bonner & Newell, 2010; De Neys & Glumicic, 2008; Ferreira, Garcia-Marques, Sherman, & Sherman, 2006; Kahneman & Frederick, 2002), with most decision makers showing a general preference to forego analytical processing for less cognitively expensive processing (Evans, 2008; Evans & Stanovich, 2013; Kahneman, 2003; Sloman, 1996; Thompson, 2009). Similarly, fuzzy-trace theorists argue that base-rate neglect occurs because new information cues intuitive gist-based associations in memory, whereas base-rates are processed through more analytical, verbatim-based reasoning (Reyna, 2004; Reyna & Brainerd,

2008; Wolfe & Fisher, 2013). That is, the discrepancy between the ease with which base-rates and individual-case evidence are processed is the source of base-rate neglect².

As previously discussed, it is unknown how older adults would perform on a belief updating task that pits error-prone heuristics against Bayes' theorem relative to younger adults. Given the literature that details the cognitive architecture of reflective processing in combination with the literature on cognitive aging, one may expect older adults to be at a disadvantage when updating beliefs compared with younger adults. However, it is unknown whether there are age differences in the stages of cognitive reflection such as conflict detection, inhibiting automatic processes, or generating alternative responses. Additionally, whether younger and older adults can learn to engage reflective processes to improve decision making is a topic that remains unexplored.

Reinforcement Learning

Tasks that measure cognitive reflection in the heuristics and biases literature are usually structured as non-repeatable word problems related to a fictitious situation, without the decision maker learning whether they are correct or not. One reason for this may be that investigators do not want to popularize answers to assessment items. For instance, an assessment commonly used as a predictor of rational thinking is the three-item Cognitive Reflection Test (Frederick, 2005). However, the items on the original Cognitive Reflection Test have become well known and have been included in classroom demonstrations, magazines, and best-seller books such as Daniel Kahneman's *Thinking, Fast and Slow* (Kahneman, 2011). An extended 15-item version of the Cognitive Reflection Test that measures rational thinking has been proposed by Toplak, West,

² But see De Neys, 2007; 2012; Pennycook & Thompson, 2012, who argue that conflicts between prior and new information may tap a conflict between two Type 1 processes.

and Stanovich (2014a) in order to avoid potential confounds of familiarity. However, due to cognitive reflection tasks not including feedback or numerous trials to observe whether performance changes with more opportunities, it is currently unknown whether more effortful, reflective processing could be learned in order to improve judgments and decisions.

Reinforcement learning theory posits that actions followed by positive outcomes are more likely to be repeated in the future, whereas actions followed by negative outcomes are less likely to reoccur (Sutton & Barto, 1998). Dopamine plays a critical role in reinforcement learning, such that dopaminergic neurons signal how much an outcome deviates from a predicted outcome (Schultz, Dayan, & Montague, 1997). It has been argued that learning takes place when an outcome is better than predicted, reflected in a phasic increase in dopamine neuron activity, whereas an outcome that is worse than expected leads to the extinction of the learned behaviour, reflected in a phasic decrease of dopamine neuron activity (Eppinger et al., 2008).

Age differences in reinforcement learning. Aging is associated with a marked decline in mesencephalic dopamine neurons, which may contribute to the impairments in cognitive control observed in old age, such as conflict monitoring and error processing (Braver & Barch, 2002; Eppinger, Hämmerer, & Li, 2011; Eppinger et al., 2008; Paus, 2001). Critically, some research suggests that older adults show deficits in processing errors and reinforcement learning due to attenuated signaling in the mesencephalic dopamine system (de Boer, et al., 2017; Mathewson, Dywan, & Segalowitz, 2005; West, 2004). In contrast, other research suggests that these deficits stem from an age-related asymmetry in the processing of feedback valence³ (Eppinger, & Kray, 2011; Eppinger et al., 2008; Mather & Carstensen, 2005). More specifically,

³However, some studies show no age differences between positive and negative feedback (see Di Rosa et al., 2017; Pietschmann, Endrass, Czerwon, & Kathmann, 2011).

positive information is processed more than negative information as older adults put more effort into regulating their emotions by preferring to process positive information and diminishing the processing of negative information, a phenomenon termed the positivity bias (Charles, Mather, & Carstensen, 2003; Reed & Carstensen, 2012). Evidence supporting the latter theory shows that older adults learn better in errorless learning conditions (Heldmann, Markgraf, Rodríguez-Fornell, & Münte, 2008).

The heuristics and biases literature typically use word problems such as the bat-and-ball problem from the Cognitive Reflection Test. The bat-and-ball problem is described as follows: A bat and a ball cost \$1.10 in total. The bat costs \$1.00 more than the ball. How much does the ball cost? (Frederick, 2005). The correct answer is \$0.05, with prior research showing that only a small portion of participants provide the correct answer (e.g., about 30.3% of Canadian undergraduates; Pennycook, Cheyne, Koehler, & Fugelsang, 2016). The most common answer to this question is \$0.10 (e.g., approximately 64.9% among Canadian undergraduate students provide this answer; Pennycook et al., 2016). However, a cursory check indicates that the ball cannot cost \$0.10 because the bat would then cost \$1.10, with the bat and ball totalling \$1.20. Many respondents give the wrong answer because \$0.10 easily comes to mind and is an automatic response. Under these circumstances, Type 2 processing may not be engaged at all, or so little that a discrepancy between automatic and normative responses goes undetected. Conflict detection is more likely to occur when participants receive feedback on their responses.

Although research shows that both younger and older adults have a preference for making heuristic-based decisions (Peters et al., 2007), it is unknown whether feedback would bolster the avoidance of using error-prone heuristics to make decisions and encourage reflective processing in a belief updating task, or whether age differences in performance would emerge.

Given the difficulty of making Bayesian inferences, it is unclear whether younger adults would use feedback to make decisions that closer reflect normative responses. Additionally, in light of previous literature showing that older adults learn less from feedback than younger adults possibly due to a decline in the dopaminergic system with age (de Boer et al., 2017) or due to a tendency to devalue negative feedback in late life (Reed & Carstensen, 2012), it may be expected that older adults would benefit from feedback less than younger adults in a belief updating task. However, this too is a topic that remains unexplored and deserves empirically testing.

Many decisions made in real life are based on ambiguous information (Savage, 1954; Volz & Gigerenzer, 2012). Thus far, I have discussed decisions made under risky conditions, in which all information is known and the answer can be derived through calculation. Decisions made under ambiguous conditions, however, have received far less empirical study in the heuristics and biases literature.

Decisions Under Risk and Ambiguity

The terms risk and ambiguity carry specific meaning in the decision making literature. Decisions made under risk are decision situations in which all relevant information is known and calculation of the correct choice is possible. Decisions made under ambiguity are decision situations in which some relevant information is unknown or unknowable (i.e., calculation of the correct choice is impossible), and the future is uncertain (Knight, 1921; Savage, 1954). Savage (1954), the author of Bayesian Decision Theory, referred to situations involving risk as “small worlds” in which all information is known, and possible alternatives and consequences can be enumerated and ordered according to subjective preferences, particularly reflecting simplified environments such as when making monetary gambles or applying Bayes’ theorem to determine an optimal choice (Binmore, 2009; Savage, 1954). In contrast, “large worlds” refer to situations

involving ambiguity, in which some of the information is unknown, and/or judgments must be made from small samples (Binmore, Stewart, & Voorhoeve, 2012; Savage, 1954). The majority of decisions made in real life are based on uncertain situations (Volz & Gigerenzer, 2012), with known probabilities and outcomes being quite rare. Examples of making decisions under ambiguity include deciding whom to marry or how to raise your children (Volz & Gigerenzer, 2012).

Bayes' theorem can only be applied to small worlds in which probabilities are known and alternatives (e.g., probabilities for multiple options) are minimal. Typically, in large world settings, Bayes' theorem cannot be applied (Savage, 1954). Other strategies must be used such as heuristics to make optimal decisions when some information is ambiguous (Volz & Gigerenzer, 2012). The financial crash in 2008 is a strong example demonstrating how statistical estimates of risk do not always predict the uncertain world of finance (Taleb, 2010). Therefore, applying the calculus of probability in uncertain environments such as the real world may not always lead to good decision making.

Typically researchers studying decision making investigate decision abilities in small-world environments, with a heavy reliance on gambling paradigms in which the individual must make binary decisions on monetary gambles with payoffs precisely defined (i.e., varying magnitudes of reward and probabilistic outcomes; Gigerenzer, 2016; Volz & Gigerenzer, 2012). However, when a portion of the information is removed from the situation, ambiguity increases, and probabilistic judgments must be represented appropriately. The focus of decisions under risk is problematic considering decisions made under ambiguity are more prevalent in real life than decisions made under risk (Gigerenzer et al., 2011). The strategies used in small worlds are different than those used in large worlds. In other words, the underlying cognitive processes that

support statistical manipulation of information (e.g., monetary gambles) are different from the processes that support judging ambiguity (e.g., whom to marry). Although decisions made under risk compared with ambiguity should be distinguished, there may also be a distinction between how younger and older adults use heuristics in situations of risk and ambiguity to inform decisions.

Age differences in decisions under risk vs. ambiguity. Of course, decisions made under risk and ambiguity apply to both younger and older adults. Older adults for example, may have to assess the risks and benefits when deciding whether to undergo surgery (decision under risk) or to sell their house and move into an assisted living facility (decision under ambiguity). Interestingly, younger and older adults show similar levels of performance when making decisions under risk (Armstrong & Spaniol, 2017; Mata et al., 2007; Zamarian, Sinz, Bonatti, Gamboz, & Delazer, 2008). However, older compared with younger adults show deficits in learning when information is ambiguous, a finding that has been associated with an increased aversion to ambiguous information with age (Eppinger, et al., 2008; Hämmerer, Li, Müller, & Lindenberger, 2011; Herbert, Eppinger, & Kray, 2011; Pietschmann, Endrass, Czerwon, & Kathman, 2011; Samanez-Larkin, Worthy, Mata, McClure, & Knutson, 2014). Older adults may be averse to spending the mental effort to maintain accurate representations of ambiguity because resources are limited relative to processing risk information that is explicitly described (Westbrook, Kester, & Braver, 2013).

Recent research investigating age differences in decisions made under ambiguity have shown that older adults relative to younger adults show a deficit in making inferences about information that is masked in order to reduce uncertainty (Hämmerer et al., 2019). Other research has reported that older adults show a reduced ability to represent ambiguous

information such that older adults assign more weight to unmasked information and less weight to masked information, whereas younger adults are more likely to consider masked information and develop accurate subjective representations of ambiguity (Nassar et al., 2016).

Decisions under ambiguity have been studied at the brain level using fMRI to examine the brain regions associated with processing ambiguous information. Specifically, there is evidence suggesting the prefrontal regions including the anterior PFC (aPFC) and orbitofrontal cortex (OFC) are involved in representing uncertainty (Badre, Doll, Long, & Frank, 2012; Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006; Schultz et al., 2008). As described earlier, the structure and function of the frontal lobe deteriorates with age (Nyberg et al., 2010). Eppinger, Heekeren and Li (2015) showed an age-related deficit in learning how to predict future rewards that were ambiguous. The results were attributed to the under-recruitment of prefrontal regions in older relative to younger adults. McGuire, Nassar, Gold and Kable (2014) have argued that the under-recruitment of prefrontal areas may limit the function of the cortical learning system, making it difficult for older adults to represent ambiguous information.

Although older adults show more difficulty learning from ambiguous information and making decisions under ambiguity compared with younger adults, it is currently unknown how ambiguous information impacts cognitive reflection in a belief updating task with younger and older adults. This is an important research question considering the daily use of heuristics when making decisions based on ambiguous information in real life.

Individual Differences in Cognitive Reflection

Investigating individual differences in judgment and decision making have important implications for the real world because this heterogeneity indicates that some people may make better judgments or decisions in domains such as medicine, law, or policy than others.

Identifying those who make better judgments and decisions than others may improve outcomes and benefit the broader society by placing good decision makers in leadership positions (Nelson et al., 2008; Reyna & Farley, 2006). Although there is a wealth of evidence showing that cognition declines with age (Salthouse, 2019), just as age-related changes in brain function are not uniform across individuals, age-related changes in cognition are also not uniform across individuals (Morse, 1993). In fact, previous research shows that cognitive abilities are more heterogeneous in older compared with middle-aged adults (Morse, 1993; Spreng, Wojtowicz, & Grady, 2010). Therefore, investigating individual differences in judgment and decision making across both younger and older adults is important. In light of this, assessments commonly used to predict individual differences in cognitive reflection in the heuristics and biases literature such as numeracy scales, the Cognitive Reflection Test, and thinking disposition scales, are also included in this dissertation (Frederick, 2005; Reyna, Nelson, Han, & Dieckmann, 2009).

Numeracy. Numeracy has been shown to be important in a range of everyday tasks, such as making medical decisions (Galesic, Gigerenzer, & Straubinger, 2009). Numeracy has been used to predict performance on decision making tasks within the heuristics and biases literature, with low numeracy associated with increased susceptibility to a variety of cognitive biases, even when general intelligence is partialled out (Cokely et al., in press; Peters, Västfjäll et al., 2006). In the context of cognitive processing, numeracy measures used in the literature heavily depend on declarative knowledge in which percentages and probabilities are derived (see Cokely, Galesic, Schultz, Ghazal, & Garcia-Retamero, 2012; Lipkus, Samsa, & Rimer, 2001; Schwartz, Woloshin, Black & Welch, 1997) and have been associated with the generation of alternative responses that are superior to heuristic-based, automatic responses when confronted with conflicting information on cognitive reflection tasks (Stanovich et al., 2016).

The numeracy assessments index the ability to effectively solve everyday problems such as evaluating medical treatment or financial options, and political claims (Cokely et al. 2012; Cokely, Ghazal, & Garcia-Retamero, 2014; Reyna et al., 2009). Particularly the Berlin Numeracy Test (Cokely et al., in press) has demonstrated that numeracy (i.e., practical probabilistic reasoning) across varying levels of expertise uniquely predicts decision making skill and risk literacy across tasks that tap real world settings (e.g., HIV prevention, cardiovascular risk mitigation), as well as tasks that elucidate thinking dispositions (e.g., framing resistance, sunk-cost biases; Cokely et al., 2014; Cokely et al., in press; Garcia-Retamero, & Cokely, 2013; Garcia-Retamero, Wicki, Cokely, & Hanson, 2014). Numeracy level is a robust predictor of both numerical and nonnumerical decisions, with many causal mechanisms linking numeracy to decision making skill such as heuristic, intuitive, gist-based and number-sense processes (Cokely & Kelley, 2009; Ghazal, Cokely, & Garcia-Retamero, 2014; Peters & Bjälkebring, 2015; Traczyk & Fulawka, 2016). Cokely and colleagues (in press) have argued that the reason for this robust predictive power is because numeracy tests are reflective of judgment and decision making tasks that engage inductive reasoning and cognitive control in settings of risk and uncertainty.

In the context of Bayesian inference, Armstrong and Spaniol (2017) showed that amongst younger and older adults, numeracy did not modulate the accuracy of Bayesian inferences. Specifically, correlational analyses were run on the scores from the 12-item Lipkus Numeracy Task and the mean absolute error of the participants' judgment of the PPV and the true PPV (with large deviations observed as younger and older adults vastly overestimated the PPV when the true PPV was small). That is, correlational analyses were conducted between numeracy scores and only one judgment. Future work should further examine the link between numeracy

and Bayesian inference by obtaining a more representative measure of Bayesian inference. For example, rather than Bayesian inference being measured by one question, multiple questions could be used to obtain a measure of average performance that may better index belief updating performance.

Cognitive Reflection Test. The Cognitive Reflection Test is a widely used assessment that is designed to measure cognitive reflection. Cognitive reflection has been defined as one's ability or disposition to suppress a response that is based on an idea that has automatically come to mind (Frederick, 2005). The Cognitive Reflection Test is comprised of word problems that evoke intuitive, yet erroneous responses. If the decision maker engages reflection, they may be able to realize the error and generate a better response (Frederick, 2005; Stanovich et al., 2016). Oechssler, Roider, and Schmitz (2009), and Hoppe and Kusterer (2011) showed that higher test scores on the Cognitive Reflection Test were correlated with lower incidences of falling prey to certain biases such as the conjunction fallacy (Campitelli & Labollita, 2010; Cokely & Kelley, 2009; Tversky & Kahneman, 1983), with low Cognitive Reflection Test scores indicating a tendency for one to act on impulse and provide intuitive responses (Toplak, et al., 2011). In the context of cognitive reflection, performance on the Cognitive Reflection Test depends heavily on conflict detection and the overriding Type 1 processes and depends less on specific knowledge (Stanovich et al., 2016).

Thinking dispositions. Thinking dispositions relate to belief formation and decision making. Baron (1985) argued that examples of thinking dispositions that relate to rational decision making include weighting new evidence against one's prior beliefs, the time spent on a problem before giving up, or assigning weight to others' opinions when forming one's own opinions. Common thinking disposition scales used in the literature include the Faith in Intuition

and Need for Cognition scales, which have been used to predict a wide range of psychological measures in the judgment and decision making literature (e.g., Cacioppo & Petty, 1982; Cacioppo et al., 1996; Epstein et al., 1996; Pacini & Epstein, 1999).

The Faith in Intuition assessment was developed to assess how much individuals trust their intuitions and instincts (i.e., preference for Type 1 processing; Epstein, et al., 1996; Pacini & Epstein, 1999). In contrast, the Need for Cognition scale was developed to assess how much a person engages in and prefers to think effortfully (i.e., preference for Type 2 processing; Cacioppo & Petty, 1982; Cacioppo, et al., 1996). Both assessments are self-report measures and typically emerge as separate factors and are generally not negatively correlated as one may assume (Epstein et al., 1996). Shiloh, Salton, and Sharabi (2002) showed that individual differences in thinking dispositions predicted the likelihood of heuristic-based responses, as well as the framing effect. As previously described, Bruine de Bruin and colleagues (2015) showed that older adults were associated with both lower numeracy and lower Need for Cognition, with the negative relationship between age and numeracy mediated by Need for Cognition. This suggests that older adults' lower numeracy performance may be driven by a lack of motivation or willingness to perform the task (Hess, 2014; Hess & Queen, 2014; Norris, McGeown, Guerrini, & Castronovo, 2015).

However, this effect does not always replicate when these scales are associated with probabilistic judgment tasks. For instance, Alós-Ferrer and Hügelschäfer (2012) reported that the Faith in Intuition scale correlated with susceptibility to the representativeness, but not the conservatism heuristic. Alós-Ferrer, Garagnani, and Hügelschäfer (2016) reported that these scales did not predict performance on tasks used to study decision biases in probability judgments, and Lu (2015) showed that both scales did not correlate with performance on the

conjunction fallacy problem. That is, although literature is mixed concerning whether thinking disposition scales predict performance on tasks that require cognitive reflection, it is important to glean whether performance is affected by motivation or the decision maker's beliefs.

In sum, individual differences in numeracy, cognitive reflection, and thinking dispositions have been found to modulate performance on heuristics and biases tasks (although there is mixed evidence concerning whether thinking dispositions underlie decision biases). The current dissertation also includes these measures to examine whether these effects extend to a belief updating task of probability information, as well as older individuals.

Chapter 3: Synthesis of Dissertation Goals

The primary goal of this dissertation is to continue my research program of examining the processes that underlie belief updating in the context of Bayesian inference and illuminate how age affects these processes. However, rather than investigate the ways in which Bayesian inference can be improved as I have previously (e.g., Armstrong & Spaniol, 2017; Armstrong, Spaniol, & Persaud, 2018; Armstrong, Sparrow, & Spaniol, under review; Wegier, Armstrong, & Shaffer, 2019), in this dissertation I examine the cognitive processes that underlie decisions that pit error-prone heuristics against Bayes' theorem. That is, similar to the cognitive reflection tasks used in the heuristics and biases literature, I examine decisions in a context in which using heuristics leads to error, and using normative strategies leads to correct decisions. Specifically, I examine use of the representativeness and conservatism heuristics when younger and older adults make decisions under risk and ambiguity, with and without feedback in a belief updating task, as well as examine the individual differences and neural correlates that underlie these decisions.

To briefly reiterate the relevant literature: 1) Bayes' theorem is used to update prior information with new information (Bayes & Price, 1763; Knight, 1921) and has been shown to be a notoriously difficult task (Grether, 1980; Kahneman & Tversky, 1972). 2) Due to the complexity of making Bayesian inferences, heuristics are used (Kahneman & Tversky, 1972). However, using heuristics to make Bayesian inferences when formal analysis (i.e., abiding by the principles of Bayes' theorem) is required leads to systematic biases in judgments such as committing a representativeness or conservatism error (Achtziger, Alós-Ferrer, Hügelschäfer, & Steinhauser, 2014). 3) Cognitive aging is associated with the increased use of heuristics when making decisions (Peters, Hess, Västfjäll, & Auman, 2007). Older adults may be required to make decisions that require Bayesian inference more often than younger adults such as in the

domains of finance or medicine (Chen & Sun, 2003). 4) Evidence is mixed concerning whether age affects Bayesian inferences (Armstrong & Spaniol, 2017; Fiske, 2005; Spaniol & Bayen, 2005). 5) In cognitive reflection tasks, an intuitive, yet suboptimal response is pit against a normative, optimal response (Kahneman, 2011). Type 1 processes generate automatic responses and can be interrupted by Type 2 processes when a conflict is detected between an automatic and normative response, with Type 2 processes also engaged to generate alternative responses (Stanovich, 2011). These stages of information processing are dependent on reflective processing, proficient mindware (e.g., knowledge of numeracy) and thinking dispositions (Stanovich et al., 2016). 6) Committing a representativeness error is more common than committing a conservatism error because diagnostic evidence (i.e., new information) has been argued to be more intuitive and concrete than base-rate (i.e., prior) information (Barbey & Sloman, 2007). 7) Cognitive reflection tasks do not provide the decision maker with feedback. Although older adults learn less successfully from feedback than younger adults potentially due to a deficiency in the dopaminergic system (Eppinger et al., 2008; Paus, 2001), or due to devaluing negative feedback with age (Mather & Carstensen, 2005), it is unknown whether feedback would provoke reflective processing in a belief updating task. 8) Cognitive reflection is commonly examined when making decisions under risk. However, most decisions made in the real world are based on ambiguous information. Although older adults show an aversion to ambiguous information relative to younger adults (Ellsberg, 1961; Seer et al., 2016), it is unknown how ambiguous information affects how heuristics are used by younger and older adults during belief updating. 9) There is now a voluminous literature on how individual differences in younger adults relate to performance on cognitive reflection tasks (Stanovich et

al., 2016). However, there is minimal research on how individual differences relate to performance on a belief updating task of probability amongst younger and older adults.

Several research questions have emerged from this literature. Specifically, it is unclear how older adults use heuristics compared with younger adults, how feedback affects decisions that depend on reflective processing, whether the neural correlates that underlie belief updating differ as a function of age, how ambiguity impacts belief updating amongst younger and older adults, or how individual differences relate to belief updating.

Dissertation Experiments

To address these questions, a series of experiments was conducted using a belief updating task. Specifically, participants were required to make a binary decision about which of two urns a sample of balls was drawn from on the basis of a) the likelihood a sample was drawn from each urn (prior information), and b) the sample of balls drawn (new information). In Experiment 1, I employed the urn-ball paradigm without feedback to examine the use of heuristics in younger and older adults, as well as the role of potential predictors of decision accuracy including numeracy, inhibitory control, and thinking dispositions. Experiment 2 was similar, however, participants received feedback following each response, and both behavioural responses and event-related potentials (ERPs) were examined to compare cognitive processes between younger and older adults. Specifically, I tested the hypothesis that younger adults would use feedback to improve decision accuracy more than older adults. I also sought to elucidate potential age differences in the neural processes that underlie belief updating. Finally, in Experiment 3, I employed a modified version of the belief updating task in which prior and new information varied in certainty (i.e., certain vs. uncertain), in order to investigate how ambiguity affects the use of heuristics by younger and older adults.

In the ensuing chapters, I first provide a General Methods section that describes the methods that are common across all experiments. I then describe each experiment in a separate chapter. Within each experiment chapter, a literature review is provided that is specific to that experiment in order to describe the relevant literature in detail, highlight the rationale for the experiment and list the respective a-priori hypotheses. The General Discussion section links the main results observed across all experiments, situates the main findings in the existing literature, and discusses the theoretical and empirical implications of the results. Finally, limitations of the experiments are discussed, directions for future research are suggested and a brief conclusion synthesizing the dissertation is provided.

Chapter 4: General Method

The purpose of this chapter is to provide an overview of the methods, procedures and analyses that are common to all three experiments in the dissertation. Details specific to experiments are described in each respective experiment chapter. All procedures met ethical standards for conducting human psychological research and were approved by the Research Ethics Board at Ryerson University (see Appendices I-VI for consent and debriefing forms). The decision situations of interest are abbreviated in the experiments for simplicity. That is, the situation in which the representativeness heuristic conflicts with Bayes' theorem is abbreviated as "RconfB", the situation in which the conservatism heuristic conflicts with Bayes' theorem is abbreviated as "CconfB", and the situation in which the representativeness heuristic aligns with Bayes' theorem is abbreviated as "RalignB".

Participants

Power analysis. Sample-size estimates were calculated a priori using G*Power 3.1.9.2 (Faul, Erdfelder, Buchner, & Lang, 2013). Using a mixed ANOVA with one between-subject factor with two levels (age: young, old), and one within-subject factor with three levels (decision situation: RconfB, CconfB, RalignB), and with a two-tailed alpha level set to .05, G*Power analysis estimated that a sample of $n = 22$ per group would provide 95% power to detect a small-to-medium between-within interaction effect ($f = .25$). To increase power for detecting effects, samples above this estimate were collected for each experiment.

Participant recruitment. A total of 85 healthy younger (ages 18 to 35) and 85 healthy older adults (ages 65 and above) participated in Experiments 1 to 3. Younger adults were recruited from the community through paper flyers posted around Ryerson University campus and online flyers posted on Facebook, Craigslist and Kijiji. Older adults were recruited from the

Ryerson Senior Participant Pool. Participants were compensated for their participation based on experiment duration. For a 1.5-hour session (Experiment 1) participants received \$20 (\$15 + \$5 bonus payout) CAD. For a 2.5-hour session (Experiment 3) participants received \$30 (\$25 + \$5 bonus payout) CAD. For a 3-hour session (Experiment 2) participants received \$35 (\$30 + \$5 bonus payout) CAD⁴. Informed consent was collected from all participants prior to participation. Participants in Experiments 1 and 3 completed the study in a testing room in the Memory and Decision Processes Lab, and participants in Experiment 2 were tested in the EEG lab in the Psychology Research and Training Centre at Ryerson University.

Inclusion and exclusion criteria. Participants' data were excluded and later replaced based on the following health-related criteria: (1) history of neurological abnormalities, neurodegenerative disorders or psychiatric disorders (e.g., traumatic brain injury, stroke, dementia, prolonged periods of unconsciousness); (2) a current diagnosis of uncontrolled medical conditions that might affect cognitive performance (e.g., diabetes, cardiovascular diseases, autoimmune disorders); (3) a current diagnosis of a mood disorder (e.g., depression or anxiety); (4) current prescription of medications that impact mental functioning; and (5) visual or hearing impairments (including awareness of colour-blindness). Only participants who met the above criteria were included in analyses for each experiment.

Additionally, participants' data was also excluded and later replaced if their performance did not meet certain criteria on specific cognitive and affective assessments, such as: (1) scoring below 26 on the Mini-Mental Status Examination (MMSE; Folstein, Folstein, & McHugh, 1975) suggesting mild cognitive impairment (only older adults completed this assessment); (2) scoring

⁴ Participants were told that if they made 80% or more correct decisions on the task, they would receive an extra \$5 in each experiment. Unbeknownst to participants, all participants were told they made over 80% of their decisions correctly and received the extra \$5.

28 or above on the depression subscale of the 21-item Depression Anxiety Stress Scale (DASS-21; Lovibond & Lovibond, 1995) suggesting extremely severe depressive symptoms within the past week; and/or (3) scoring 20 or above on the anxiety and stress subscales of the DASS, suggesting extremely severe anxiety symptoms within the past week. Participants who provided extreme scores on these assessments were excluded as previous research shows that extreme depression, anxiety, and stress impact attention and memory (Dalglish et al., 2003; MacLeod & McLaughlin, 1995), potentially confounding results. Last, to ensure participants included in analyses understood the belief updating task, younger and older adults were excluded if they scored below 70% in decision accuracy on the alignment decision situations, as the difficulty level of these decisions was low and participants' performance would reveal whether task instructions were comprehended or not.

Materials

Belief updating task. All experiments used an urn-ball paradigm, similar to those used by Achtziger et al. (2014) and Grether (1980, 1992). The task required participants to make a binary decision about which of two urns a sample of balls was drawn from on the basis of a) the likelihood of each urn chosen to draw a sample from (prior information), and b) the sample of balls drawn (new information). Upon arrival to the lab, participants were introduced to the experiment and informed consent was obtained. The experiment commenced with the belief updating task conducted on the lab computer, which was programmed and presented using Presentation version 20.0 software (Neurobehavioral Systems, Inc., Berkeley, CA). The experimenter provided each participant with a demonstration of the urn task using two jars that contained blue and green gumballs with their proportions reflecting that of the task. The experimenter explained the colour proportions in each urn and stated that there was a 75%

likelihood that they would draw a sample from the left urn, and a 25% likelihood they would draw a sample from the right urn. The experimenter drew four predetermined gumballs out of one of the jars without the participant looking, showed the sample to the participant, and asked them to decide which urn they believed the experimenter had drawn the sample from⁵. Once the participant chose an urn, the experimenter asked them if they had any questions. The experimenter then showed the participant printed sheets of paper of what the stimuli would look like on the computer and explained how the stimuli was representative of the colour proportion of balls in each jar used in the demonstration and answered any questions the participant had. Participants were then asked to carefully read through instructions presented on the computer screen and were subsequently asked to repeat the instructions of the task in their own words to the experimenter. Once the experimenter felt the participant understood the task, participants began the practice task consisting of 15 practice trials that included each combination of prior and sample information. If there were no questions after completing the practice task, participants began the belief updating task (see Appendices VII-IX for belief updating instructions for each experiment).

The task consisted of 424 trials (i.e., 120 trials per decision situation, with 64 filler trials), with a 2-minute break in the middle of the computer task separating counterbalancing blocks, and three optional 30 second breaks within each block (six 30 second breaks total). Specifically, the 2-minute break was presented after the 212th trial, and 30-second breaks were presented within each block on the 53rd, 106th, and 159th trial. After the 2-minute break, the majority colour of the urns and sample changed. Trials were presented in random order within each counterbalanced block.

⁵ The representativeness situation was presented in the demonstration.

In a single trial sequence, participants were presented with a) the distribution of coloured balls within two urns, b) prior information (i.e., the probability the computer would choose either urn), and c) sample evidence (i.e., 4 balls drawn from one of the two urns by the computer). The belief updating task employed by Achtziger et al. (2014) presented prior and sample information independently of one another. The modified belief updating task used in this dissertation series presented information on each trial sequentially, layering on the previous information presented, such that the urn distributions were presented first, then the prior, then the sample, and then the decision prompt, for two reasons: 1) to ensure participants paid attention to each individual component of information, and 2) to avoid effects being confounded with memory constraints (e.g., attenuated memory for prior information). After the sample was presented, a decision prompt was presented requiring participants to decide which urn they believe was more probable that the sample was drawn. In Experiment 1 a blank screen followed by a fixation cross was presented subsequent to a response that separated trials, and in Experiment 2 and 3 a feedback window followed by a fixation cross was presented subsequent to a response that separated trials.

The belief updating task used causal rather than noncausal prior information. In the cabs problem described above, noncausal base-rates were used (i.e., base-rates with no obvious causal link to the new evidence; Ajzen, 1977; Barbey & Sloman, 2007; Bar-Hillel, 1980; Koehler, 1996; Tversky & Kahneman, 1982). The causal version of the cabs problem states that the two cab companies have the same number of cabs, however 85% of accidents in the city involving cabs are Green cabs and 15% involve cabs that are Blue, rather than stating that 85% of the cabs in the city are Green. The former statement emphasizes the potential involvement of a Green cab in the accident compared with the latter statement. Research has shown that causal base-rates are used more than noncausal base-rates because they seem more related to the new evidence

presented. The belief updating task employed in the current dissertation presented the number of urns, the number of balls in each urn, the proportion of coloured balls and the likelihood the computer will choose either urn (i.e., causal base-rate information).

In keeping with the method introduced by Achtziger and colleagues (2014), in order to measure reliance on the representativeness heuristic, six decision situations were of interest: 2 *alignment* situations in which the representativeness heuristic led to the correct urn choice (i.e., the heuristic and Bayes' theorem were aligned), 2 *conflict* situations in which the representativeness heuristic led to the incorrect urn choice (i.e., the heuristic and Bayes' theorem were in conflict), and 2 *neutral* situations in which the representativeness heuristic did not apply. Further, decision situations were also designed to test for use of the conservatism heuristic. Specifically, there were two decision situations (i.e., the neutral decision situations described above) in which the most likely urn was *not* the urn with the higher prior probability. That is, the neutral situations did not involve a conflict between the representativeness heuristic and Bayes' theorem but did involve a conflict between the conservatism heuristic and Bayes' theorem. Participants who use prior information to make decisions and underweight sample evidence are likely to choose the incorrect urn in these decision situations. In line with Achtziger et al. (2014), across Experiments 1-3, dependent variables included decision accuracy and reaction times (RTs).

Single trial sequence. The trial began when a fixation cross was presented cueing the beginning of the trial for 1000 ms. Next a representative distribution of 4 blue and green balls in two urns ("urn distributions") were presented vertically on the left and right sides of a grey background on the computer screen with a fixation cross in the center of the screen for 1500 ms. After 1500 ms, the probability of either urn being chosen by the computer to draw a sample of 4

random balls (“the prior”) was presented for 3000 ms with the urn distributions. Achtziger et al. (2014) used a rule to define the prior probabilities such that the computer would draw a random number between 1 and 4 that would not be revealed to participants. For example, if the computer randomly draws a 1 the left urn will be chosen (25% chance), however if the computer randomly draws a 2, 3, or a 4, the right urn will be chosen (75% chance). In order to reduce cognitive load, the current dissertation diverged from Achtziger et al.’s (2014) design, and presented prior probabilities as percentages. Specifically, prior probabilities for the left urn consisted of 25% (75% right urn), 50% (50% right urn), and 75% (25% right urn), and were presented with equal frequency. Next, the sample of 4 balls drawn by the computer (“the sample”) from one of the two urns was presented in the middle of the screen in vertical alignment in addition to the urn distributions and prior information. The sample remained on the screen until the participant pressed one of two corresponding keys (“F” for the left urn and “J” for the right urn) to indicate their urn choice. There was no response deadline and no feedback. Once the participant made a decision a blank screen was presented for 1000 ms⁶, and the sample was put back into the urn it was drawn from (i.e., sampling 4 balls with replacement). See Figure 1 for a schematic of a single trial sequence.

⁶The blank screen was used as a placeholder in Experiment 1 for the feedback screen presented in Experiment 2 and 3.

Figure 1. Sample trial of the representativeness heuristic conflicting with Bayes' theorem in Experiment 1.

Decision situations. The combinations of prior and sample information were designed to form decision situations that would test the use of the representativeness and conservatism heuristics. Prior-sample combinations included three possible priors (i.e., 25% chance of left urn and 75% chance of right urn; 50% chance of left urn and 50% chance of right urn; 75% chance of left urn and 25% chance of right urn) and five sample outcomes (0-4 majority colour balls), for a total of 15 possible decision situations. The representativeness heuristic is used when a decision is made on the basis of the match between the sample and the parent population. In light of this, the representativeness heuristic influenced the *correct* urn in two decision situations (aligned: prior 50% left urn and 2 blue balls drawn; prior 50% left urn and 3 blue balls drawn), the *incorrect* urn in two decision situations (conflict: prior 75% left urn and 2 blue balls drawn; prior 25% left urn and 3 blue balls drawn), and did not influence a decision in two decision

situation (neutral: prior 75% left urn and 1 blue ball drawn; prior 25% left urn and 4 blue balls drawn). The conservatism heuristic is used when a decision is made on the basis of prior information. To examine use of the conservatism heuristic, there were two decision situations in which the higher prior probability was *not* indicative of the correct choice (i.e., Bayes' theorem conflicts with the conservatism heuristic). These decision situations were composed of the following prior-sample combinations: prior 75% left urn and 1 blue ball drawn, and prior 25% left urn and 4 blue balls drawn. Following Achtziger et al. (2014), the posterior odds⁷ for the *left urn* for each prior-sample combination, and the prescriptions of Bayesian updating, using the representativeness or conservatism heuristic are presented in Table 1.

Table 1. *Posterior Odds for all Decision Situations*

Prior	Number of majority colour balls in the sample									
	0		1		2		3		4	
75%	0.19	R	0.56	R	1.69	L	5.06	L	15.19	L
	L	-	L	-	L	R	L	L	L	-
50%	0.06	R	0.19	R	0.56	R	1.69	L	5.06	L
	-	-	-	-	-	R	-	L	-	-
25%	0.02	R	0.06	R	0.19	R	0.56	R	1.69	L
	R	-	R	-	R	R	R	L	R	-

Note. Prior = prior probability for the left urn. For each prior-sample combination, the posterior odds are given for the left urn (upper left), and the remaining three entries are the prescriptions for Bayesian updating (upper right), conservatism heuristic (lower left) and the representativeness heuristic (lower right). Dark grey and black shaded cells indicate cases in which Bayesian updating conflicts with the representativeness and conservatism heuristics, respectively, and light-grey cells indicate cases in which Bayesian updating aligns with the representativeness heuristic.

⁷ The posterior odds were calculated by multiplying the prior odds ratio by the sample odds ratio.

Importantly, the six shaded decision situations of interest approximate the same level of difficulty (see Appendix X for the posterior probabilities and corrected odds for each decision situation of interest).

Stimuli. Blue and green balls (image size 60×60 pixels) were shown on the computer screen to reflect the probabilistic distributions of the urns and sample evidence. Numbers were presented on the computer screen indicating the rule of the trial (i.e., prior information), and were presented in black (RGB value: 0, 0, 0) 24-point Arial font against a grey (RGB value: 100-100-100) background. All stimuli were presented in the center of the computer screen within a space of 220×220 pixels. In addition, to control for brightness, coloured stimuli were equated on luminance (blue, [RGB value: 0-0-255]; green [RGB value 0-255-0]).

Both urns were always presented in the same location on the computer screen (i.e., vertically on the left and right sides). The original urn task paradigm conducted by Grether (1980, 1992) used two urns filled with 6 balls each. However, Achtziger et al. (2014) validated a task version with urns filled with 4 balls each, which was the design employed in the current series of studies. The majority colour was counterbalanced, with the majority colour always presented at the top, and the minority at the bottom. Counterbalancing Condition 1 used 3 blue balls and 1 green ball in the left urn, and 2 blue balls and 2 green balls in the right urn. Counterbalancing Condition 2 used 3 green balls and 1 blue ball in the left urn, and 2 green balls and 2 blue balls in the right urn (see Table 2 for an illustration). The distribution of the balls within each urn was constant throughout each trial (i.e., the left urn always reflected a 75% - 25% colour distribution and the right urn always reflected a 50% - 50% colour distribution).

Table 2. *Colour Counterbalancing of Urn Distributions*

	Left Urn	Right Urn
Counterbalancing 1
Counterbalancing 2

Demographic, health, cognitive and affective assessments. A battery of paper-and-pencil assessments were completed after the belief updating task in all experiments to measure individual differences related to demographic information, as well as cognitive and affective characteristics. Information obtained from these assessments was used to ensure that the younger and older adult samples in each experiment were representative of these age groups (e.g., expected age-related differences), and was used to determine whether individual differences in these variables were associated with the dependent variables being investigated. The assessments were administered to participants in the order listed below in each experiment.

Demographic and health assessment. All participants completed a demographic and health questionnaire administered by the experimenter via phone or email to screen for eligibility. Concerning demographic information, participants were asked about their age, gender, and years of formal education. Health-related queries included prior and/or current diagnoses of medical conditions (e.g., heart conditions, neurological disorders, or psychiatric disorders), visual or hearing impairments (e.g., awareness of colour blindness), and current medication use (see Appendix XI).

Cognitive assessments.

Numeracy. As the primary goal of the current dissertation is to examine the underlying processes of belief updating, it was important to determine whether numeracy level contributes to decision accuracy in the belief updating task. In order to measure numeracy, two numeracy

assessments, namely the Lipkus Numeracy Task (Lipkus et al., 2001; Schwartz et al., 1997) and the Berlin Numeracy Test (Cokely et al., 2012) were included.

Lipkus Numeracy Task. The 11-item Lipkus Numeracy Task (Lipkus et al., 2001) with one coin toss item (12-items total) by Schwartz and colleagues (1997) was used to assess numeric ability (e.g., “If the chance of getting a disease were 10%, how many people would be expected to get the disease out of 100?”). Questions were presented in a multiple-choice (2-items) or fill-in-the-blank (10-items) format. Correct answers were summed to derive a score out of 12. Higher scores indicated higher levels of numeracy.

Berlin Numeracy Test. A fill-in-the-blank version of the 4-item Berlin Numeracy Test (Cokely et al., 2012) was also used to assess levels of numeracy (e.g., “Imagine we are throwing a five-sided die 50 times. On average, out of these 50 throws how many times would a five-sided die show an odd number [1, 3, or 5]?”). Correct answers were summed together to derive a score out of 4. Higher scores indicated higher levels of numeracy.

Digit Symbol Substitution Test (DSST). The DSST is a neuropsychological test from the revised version of the Wechsler Adult Intelligence Scale (WAIS-R; Wechsler, 1997) and measures processing speed, a component of fluid intelligence. In the task, nine digits are paired with unique symbols. Participants are asked to match as many of the symbols with their corresponding digits as possible in 2-minutes, working from left to right down each row. The total number of correct number-symbol pairings provided an index of processing speed.

Mill Hill Vocabulary Test. The Mill Hill Vocabulary Test assessment (Raven, 1982) is a 34-item vocabulary measure that served as an index of crystallized intelligence. During the task, participants are instructed to choose one of six words that has the same meaning as a prompt

word. Item difficulty gradually increases from items 1 to 34. Correct answers are summed, with higher scores indicating higher verbal intelligence.

Mini-Mental Status Examination (MMSE). The MMSE (Folstein et al., 1975) was used as a screening test for potential cognitive impairment amongst older adult participants. The 11-item assessment with 30 possible points to earn required the experimenter to verbally ask questions in which the participant had to provide a verbal response (e.g., “What is today’s date?” or “How do you spell the word ‘world’ backwards?”), and required the participant to answer questions via paper-and-pencil (e.g., “Write a complete sentence.”) or perform an action (e.g., “Take this paper in your right hand, fold it in half, and lay it on your lap.”). Scores for correct answers were added across items, for a maximum score of 30. Participants scoring below 26 were excluded from the experiments.

Thinking disposition assessments.

Faith in Intuition. A 15-item version of the Faith in Intuition scale developed by Keller, Bohner and Erb (2000), a subscale from the Rational-Experiential Inventory (REI; Epstein et al., 1996), was used to measure how much decisions are based on gut feelings and intuition. Participants were asked to indicate the degree to which each of the 15 items related to them using a 10-point Likert scale ranging from 1 (“completely false”) to 10 (“completely true”) on items such as “When it comes to making decisions, I often follow my gut feelings”. Following Alós-Ferrer and Hügelschäfer (2012), participant responses across items were summed and divided by 15 to obtain a total score ranging from 0 to 10. Higher scores indicated higher faith in intuition.

Need for Cognition. The Need for Cognition assessment (Cacioppo & Petty, 1982), a subscale from the REI (Epstein et al., 1996), was used to measure motivation to engage in effortful cognitive activities. Participants were asked to indicate the degree to which each

statement related to them using a 5-point Likert scale ranging from 1 (“extremely uncharacteristic of me”) to 5 (“extremely characteristic of me”) on items such as “I like to have the responsibility of handling a situation that requires a lot of thinking”. Nine of the 18 items are reverse-score items. The appropriate items were first reverse-scored and were then added together to obtain a summary score out of 90. Higher scores indicated higher need for cognition.

Preference for Intuition and Deliberation (PID). The PID assessment was used to measure the extent to which decisions are made based on intuition or deliberation. Although the PID was completed by participants across all experiments, the authors of the PID reported several weaknesses of the assessment such as the explicit use of the terms intuition and deliberation artificially increasing the correlation between scales and the subscales did not capture individual differences in decision mode that were identified by a factor analysis (Betsch & Iannello, 2010). The authors developed a new scale (i.e., the Unified Scale to Assess Individual Differences in Intuition and Deliberation [USID]; Betsch & Iannello, in preparation), that combines the PID scale with the Rational-Experiential Inventory (Pacini & Epstein, 1999), the General Decision Making Style inventory (GDMS; Scott & Bruce, 1995), the Cognitive Style Indicator (CoSI; Cools & Van den Broeck, 2007), and the Perceived Modes of Processing Inventory (PMPI; Burns & D’Zurilla, 1999). Due to the critical weaknesses of the PID, the scale was dropped from analysis and will not be discussed in the ensuing chapters.

Affective assessments.

Depression Anxiety Stress – 21 item (DASS-21) scale. The DASS-21 scale was designed to measure depression, anxiety, and stress over the past week (Lovibond & Lovibond, 1995). The depression subscale includes seven items that measure dysphoria, sense of hopelessness, self-deprecation, devaluation of life, lack of interest/motivation, and anhedonia. The anxiety subscale includes items that measure autonomic arousal, skeletal/muscle effects, and subjective

experience of anxious affect. The stress subscale includes items that target difficulty with relaxing, the tendency to be irritable and/or over-react, and nervousness. Participants were asked to indicate the degree to which each statement has related to them over the past week on a 4-point Likert scale ranging from 0 (“did not apply to me at all”) to 3 (“applied to me very much or most of the time”). Scores for each scale were summed and multiplied by two with a possible range of 0 to 42 for each subscale. These scores were then grouped into classifications of severity for each subscale to classify symptoms as normal, mild, severe, or extremely severe. In the current dissertation, only estimates of depression and anxiety were analyzed, given prior research suggesting that such symptoms can elicit different emotional biases during cognitive processing (e.g., Dalgleish et al., 2003; MacLeod & McLaughlan, 1995).

Positive and Negative Affect Schedule (PANAS). The PANAS quantifies experience with two dimensions of mood—positive and negative affect (Watson, Clark, & Tellegan, 1988). The 20-item scale consists of 10-items reflecting positive emotions and 10-items reflecting negative emotions. Participants were asked to rate the extent to which they were experiencing each emotion in that moment on a 5-point Likert scale ranging from 1 “very slightly or not at all” to 5 “extremely”. The sum of ratings for positive items was used to index positive affect and the sum of ratings for negative emotion items provided an index of negative affect. Possible scores ranged from 10 to 50 for each dimension with higher scores indicating higher levels of positive or negative affect.

Self-Assessment. A self-assessment was completed following the belief updating task that required participants to reflect on their performance on the task. Participants were asked 1) what information influenced their decisions the most, 2) what strategy they used to make their decisions, 3) how accurate they believed their decisions were, 4) how difficult they found the

task, and 5) how confident they were working with numbers in general. Questions in the self-assessment were formatted in multiple-choice, and an open space was also provided for comments about the task. Participant responses were examined to gauge how participants assessed their own performance (see Appendices XII-XIII).

Experimental Design

The experiments employed different designs resulting in differing ANOVAs conducted. However, dependent variables included decision accuracy and RT for all experiments.

Dependent Variables

Decision accuracy and RT. Average accuracy and RT for each of the 15 decision situations in the experiment was calculated for each block. Of the 15 prior-sample combinations, 6 were decision situations of interest, corresponding to 2 representativeness situations, 2 conservatism situations, and 2 alignment situations (see shaded decision situations in Table 1). To examine whether decision accuracy statistically differed between similar situations, paired-sample *t* tests on accuracy were conducted for each pair of similar situations (e.g., 2 representativeness situations). Similar situations were aggregated regardless of whether these situations differed from each other in decision accuracy in order to yield 3 decision situations (representativeness, conservatism, and alignment). Alignment situations were examined to ensure participants understood the task (i.e., participants had to score a minimum of 70% or more in alignment decision situations, otherwise participants were excluded from the final sample). To investigate the rate of learning, trials were blocked (i.e., 4 blocks in Experiment 1, 2 blocks in Experiment 2, and with rate of learning not examined in Experiment 3). To investigate whether accuracy improved or whether RT became faster across trial blocks and whether this differed across age group or decision situation, $2 \text{ (age)} \times 3 \text{ (decision situation)} \times 4 \text{ (trial block)}$ mixed

ANOVAs were conducted. To unpack significant interactions, planned comparisons, derived from a-priori hypotheses, were conducted.

Data Analysis

Data analysis was conducted using both IBM SPSS version 24 (IBM Corp., 2014) and R (R Core Team). Across all experiments, results were interpreted in terms of statistical significance ($\alpha = .05$), effect sizes, and odds ratios. When ANOVAs were conducted, partial eta squared (η_p^2) values (i.e., the proportion of variance explained by a variable that cannot be explained by other variables in the model) were interpreted. When two means were compared via independent or paired-sample t tests, Cohen's d was used to estimate effect sizes, with .20 (small effect), .50 (medium effect), and .80 (large effect) thresholds used for interpretation. When multi-leveling modeling was conducted, odds ratios were used to estimate effect sizes, with 0 (perfect) and 1 (no effect) thresholds used for interpretation.

When a significant Mauchly's test indicated that the assumption of sphericity was violated, the Greenhouse-Geisser correction was used to correct degrees of freedom of the F -distribution. In addition, tests of normality and checks for outliers were conducted on decision accuracy and RT in each experiment. Data points that fell beyond three standard deviations from the mean were considered outliers. Outliers were removed from the dataset, and the analysis was rerun. Results from analysis with outliers included were reported when results were the same regardless of whether outliers were removed from the dataset or not, with outliers described where applicable in the ensuing experiment sections. To probe significant interactions, planned comparisons were conducted based on a-priori hypotheses. Type I errors were controlled by adjusting the alpha-levels using Benjamini-Hochberg corrections when post-hoc comparisons were made.

Decision strategy. Separate Pearson bivariate correlations were conducted to examine associations between accuracy in the representativeness and conservatism situations for each age group. Correlations were informative with respect to the relationship between the two heuristics across individuals.

Individual differences. Given that the data was clustered at the participant variable (i.e., each participant made repeated binary decisions), and to add multiple predictors to control for these variables and examine whether they would predict performance, logistic multi-level modeling was performed. Prior to running analyses, individual differences indexed by the assessments were used as predictor variables of decision accuracy. First, predictor variables were grand-mean centered. Second, the intraclass coefficient (ICC) was calculated in order to observe how much variability in decision accuracy was due to clustering (i.e., the ratio of the between-cluster variance of decision accuracy to the total variance of decision accuracy). The ICC is used to summarize the degree to which the values of a variable can be characterized by the grouping of individual observations into clusters (Flora, 2018). Specifically, the ICC represents the ratio of variance resulting from clustering relative to the total variance. Ranging from 0 to 1, an ICC of 0 indicates perfect independence (i.e., the outcome does not depend on cluster and a traditional regression can be used), and an ICC of 1 indicates perfect interdependence (i.e., the outcome only varies between clusters; Flora, 2018). The ICC was calculated by dividing the intercept variance by the sum of the intercept and residual variance. When a substantial ICC value was observed (e.g., 10%), a multi-level modeling (MLM) approach was used to examine predictor variables of accuracy data.

Next, empty models with no predictor variables added were run to assess how much variance in decision accuracy existed due to clustering. Then, a random-intercepts model and a

random-slopes model on accuracy within representativeness and conservatism decision situations, with age included as a Level 2 predictor were compared in order to observe which model fit the data better. A random-intercepts model assumes that the effect of decision situation is constant across participants, and a random-slopes model assumes that the effect of decision situation differs across participants. The model that derived the lowest Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and deviance values indicated the model that best fit the data. Predictor variables were then added to the model separately to determine the variance they contributed to decision accuracy. Of note, in the effort to keep the results concerning individual differences concise, only significant interactions and important or surprising null effects are reported. See Appendix XIV for the R script used for multi-level modeling.

Chapter 5: Experiment 1

Although there is a general tendency to forego Type 2 processing for Type 1 processing for all decision makers (Kahneman, 2011; Stanovich, 2011), older adults have been shown to rely on Type 1 processes more than younger adults in situations that require Type 2 processing (Peters et al., 2007). It has previously been argued that due to cognitive decline, older adults may lack motivation to engage processes that are cognitively taxing (Bruine de Bruin et al., 2015), and instead may process information in ways that save cognitive resources, which can result in the overuse of heuristics (Bruine de Bruin, Parker, & Fischhoff, 2012; Johnson, 1990; Peters et al., 2007). At present, how age affects the processes that underlie belief updating when Type 2 processing is required is unknown.

Dual-process literature uses simple data such as decision response time to investigate decision processes (Kahneman, 2011). Tasks used to examine cognitive reflection in the heuristics and biases literature typically use non-repeatable word problems to examine decisions. However, these tasks do not allow for repetition and after having read a short paragraph about a fictitious scenario, response times are typically long. Using paradigms that have a large number of trials for each participant affords the opportunity to compare average response times from different decision situations within an individual in order to make predictions about decision processes (Alós-Ferrer et al., 2016). For example, if a response is generated from an intuitive process and another response is generated from a normative process, one may predict the former response to be faster on average than the latter response because intuitive processes are faster than normative processes (Kahneman, 2011). Longer response times are observed particularly in situations in which the decision maker faces strong trade-offs and there is an inner struggle resulting in a slower decision (Dashiell, 1937). In contrast, if one alternative is clearly preferred

or generated automatically, a faster decision ensues. To examine process data associated with conflict and cognitive reflection, conflict situations should be compared with non-conflict situations. Whereas intuitive and normative processes prescribe different choices in conflict situations, intuitive and normative processes prescribe the same choice in non-conflict situations (i.e., intuitive and normative responses are aligned).

Response times in cognitive tasks slow with age (Brinley, 1965; Salthouse, 2019). As previously described however, older adults are more likely than younger adults to rely on automatic processes to make decisions (Peters et al., 2007), which have been linked to fast decisions (Kahneman, 2011). It is possible that even though older adults rely on automatic processes to make decisions, response execution may be slow. Given the wealth of evidence showing the general slowing in decision response time in older relative to younger adults (Brinley, 1965; Salthouse, 2019), it may be expected that older adults would make decisions slower than younger adults, regardless of whether a decision situation influences intuitive or reflective processing.

Additionally, older adults may be less likely to engage reflective processes compared with younger adults, especially on tasks that rely heavily on inhibitory control, working memory, and fluid intelligence – components of cognition that decline with age (Braver & Barch, 2002). As described in Chapter 2, individual differences in specialized knowledge such as numeracy, and thinking dispositions such as a preference to make decisions based on analytical reasoning or intuition, have been shown to modulate performance on cognitive reflection tasks. Less explored in the heuristics and biases literature are individual differences in inhibitory control and how these differences may relate to cognitive reflection.

Theories of conflict monitoring, and cognitive control have been primarily based on interference effects observed in the Stroop task (MacLeod, 1991; Stroop, 1935) and the Eriksen Flanker task (Eriksen & Eriksen, 1974), in which stimuli are either in conflict or not. Interference effects derived from incongruent stimuli are evident in response times with incongruent trials producing longer response times than congruent and neutral trials, and with congruent trials producing shorter response times than incongruent or neutral trials. Using neutral response times to index a baseline of processing speed, the difference in response times between incongruent and neutral trials (i.e., the Stroop effect) reflects a measure of inhibitory control (Stroop, 1935).

The Stroop task has been used to study the age-related declines in inhibitory processes (MacLeod, 1991), with older adults reliably showing a larger Stroop effect than younger adults, due to a decline in the efficiency of inhibitory processes (West & Alain, 2000). Given the literature showing that executive functions such as inhibition are required to successfully suppress Type 1 processing in order to engage Type 2 processing when confronted with conflict, those with greater inhibitory control may associate with higher decision accuracy on a belief updating task. However, at present it is unknown whether the Stroop effect would be associated with belief updating in younger and older adults.

Achtziger and colleagues (2014) conducted a study using an urn ball paradigm, in which prior information should be updated with new information in order to make a binary decision. Specifically, the authors related individual differences in representativeness and conservatism errors to individual differences in ERP amplitudes. The authors, however, did not report which heuristics were used most often or whether multiple heuristics were used by participants. Results of the study showed that younger adults' decisions improved over time without the presentation of feedback. This result is unsurprising as numerous studies have demonstrated younger adults'

ability to learn without feedback (e.g., Kelley & McLaughlin, 2012). In contrast, older adults show a deficit in the ability to learn without feedback (Eppinger, Heekeren, & Li, 2015; Herbert et al., 2011). However, older adults also show evidence of adapting to decision contexts (Gigerenzer et al., 2011; Mata, Josef, Samanez-Larkin, & Hertwig, 2011; Mata et al., 2007). Some research shows that the quality of decisions made by older adults have been shown to be comparable to younger adults' decisions, however the strategies used to make decisions differ between age groups (Mata et al., 2007). Currently it is unknown whether older adults would learn to update beliefs without feedback by showing an increase in decision accuracy over time. In addition, research shows that younger and older adults do not differ on decisions that are based on risk information (Zamarian et al., 2008), particularly on decision tasks that require Bayesian inference (Armstrong & Spaniol, 2017). Whether improvements in decision accuracy over time differ as a function of age on a belief updating task that requires cognitive reflection about risk information is unknown.

In light of the research described above, the current experiment used a version of Achtziger et al.'s (2014) paradigm to investigate the effects of age and individual differences in numeracy, thinking dispositions, and inhibitory control on decision accuracy and response times in a belief updating task in which error-prone heuristics were pit against Bayes' theorem. The primary goals of Experiment 1 were to, 1) elucidate how older adults use heuristics in a belief updating task relative to younger adults, 2) examine whether representativeness errors are more commonly made compared with conservatism errors, 3) examine whether there are age differences in decision accuracy over time without feedback, 4) test whether longer (shorter) response times associate with conflict (alignment) situations and whether age slows decision

time; and 5) investigate whether numeracy, thinking dispositions, and inhibitory control modulate performance on the belief updating task.

Hypotheses

As described above, older adults show evidence of relying on automatic processes more than younger adults, leading to the overuse of heuristics with age (Beseides, Deck, Sarangi, & Shor, 2012; Johnson, 1990; Peters et al., 2007) and with younger but not older adults improving decision quality without feedback over time (Eppinger et al., 2015). As such, results are expected to reveal age differences, such that older adults were predicted to associate more with decision errors than younger adults, with younger but not older adults predicted to show evidence of learning over time (i.e., increased decision accuracy and response time). Further, given the literature showing that diagnostic information (i.e., new information) is perceived as more concrete and intuitive than base-rate (i.e., prior) information (Barbey & Sloman, 2007), it was predicted that both younger and older adults would commit representativeness errors more than conservatism errors, but with older adults more likely to commit both types of errors than younger adults.

Additionally, it was hypothesized that decision situations that evoke reflective, Type 2 processing would be associated with slower response times compared with decision situations that evoke intuitive, Type 1 processing. Further, given the evidence of age-related slowing in decision response times (Brinley, 1965; Salthouse, 2019), it was also predicted that older adults would produce slower response times than younger adults overall.

In light of research showing that inhibition is required for reasoning ability (Stanovich & West, 2000), it was hypothesized that inefficient inhibitory control indexed by larger Stroop effects, would be associated with lower decision accuracy in conflict relative to alignment

situations. Given that inhibitory mechanisms decline with age (Hasher & Zacks, 1988), older adults are expected to have difficulty with decisions that evoke automatic processes that must be inhibited in the course of reasoning. Further, numeracy, processing speed (an index of fluid intelligence) and thinking dispositions were expected to modulate performance in belief updating. Specifically, higher scores on the Berlin Numeracy Test, Lipkus Numeracy task, DSST and Need for Cognition, and lower scores on the Faith in Intuition assessments were expected to predict higher decision accuracy. Lower scores on the Berlin Numeracy Test, Lipkus Numeracy task, DSST and Need for Cognition, and higher scores on the Faith in Intuition assessments were expected to predict lower decision accuracy (see Table 3 for the list of hypotheses for Experiment 1).

Table 3. *Hypotheses for Experiment 1*

List of Hypotheses	
	1 Younger adults were expected to show higher decision accuracy than older adults across decision situations.
	2 The RconfB situation was expected to produce lower decision accuracy than the CconfB and RalignB situations.
	3 Older adults were predicted to make less accurate decisions in the RconfB situation than younger adults, with no age difference in decision accuracy in the CconfB situation.
	4 Younger, but not older adults, were expected to show an increase in decision accuracy and faster response times across trial blocks.
	5 The RconfB and CconfB situations were expected to produce slower RTs than the RalignB situation, with older adults producing slower RTs than younger adults overall.
	6 Larger Stroop effects (i.e., lower inhibitory control) were hypothesized to associate with lower decision accuracy in conflict situations, with older adults showing larger negative correlations than younger adults.
Assessments	7 Higher (lower) scores on the BNT, Lipkus, DSST and NFC, and lower (higher) scores on the FI scales were expected to associate with higher (lower) decision accuracy, regardless of age.

Note. RconfB = Representativeness heuristic conflicts with Bayes' theorem; CconfB = Conservatism heuristic conflicts with Bayes' theorem; RalignB = Representativeness aligns with Bayes' theorem; BNT = Berlin Numeracy Test; Lipkus = Lipkus Numeracy Task; Digit Symbol Substitution Test; NFC = Need for Cognition; FI = Faith in Intuition; RT = response time.

Method

Participants

The final sample included 30 younger adults (ages 18-33) and 30 older adults (ages 66-84). Participants completed the battery of assessments described in the General Methods section (pages 56-61). These data show that older adults were more educated, had fewer symptoms of stress, higher positive affect, slower processing speed, higher vocabulary, and lower numeracy levels on the Lipkus numeracy task, compared with their younger counterparts. One younger adult was excluded and replaced for scoring too high on the depression and anxiety subscales of the DASS-21, and six older adults were excluded and replaced for not meeting the decision accuracy criteria for the alignment situations (i.e., scoring below 70%) described in the General Methods section (pages 47-48). The characteristics of the final sample are presented in Table 4.

Table 4. *Characteristics of the Final Sample in Experiment 1*

		Younger Adults (n = 30)	Older Adults (n = 30)		
		<i>M (SD)</i>	<i>M (SD)</i>	<i>p</i>	<i>d</i>
Demographics					
	Age in Years	23.27 (4.21)	74.00 (4.31)	<.001	-11.89
	Age Range	18-33	66-84	-	-
	Male/Female Ratio	10/20	7/23	.28	-
	Education in Years	15.30 (1.18)	16.67 (3.01)	.02	-0.59
	Right/Left Handed	27/3	28/2	.50	-
Emotion					
Assessments	Depression ^a	5.60 (4.91)	3.40 (4.90)	.09	0.45
	Anxiety ^a	3.73 (3.27)	3.13 (3.31)	.48	0.18
	Stress ^a	9.87 (5.92)	6.60 (6.24)	.04	0.54
	Positive Affect ^b	23.20 (6.26)	31.70 (8.38)	<.001	-1.15
	Negative Affect ^b	12.47 (3.25)	11.97 (3.02)	.54	0.16
Cognitive					
Assessments	MMSE ^c	-	29.20 (.714)	-	-
	DSST ^c	96.67 (14.81)	58.90 (17.99)	<.001	2.29
	Mill Hill ^c	14.43 (4.80)	22.57 (3.38)	<.001	-1.96
	BNT ^c	1.43 (1.19)	0.90 (1.37)	.11	0.41
	Lipkus Numeracy ^c	10.13 (1.74)	8.20 (2.76)	<.01	0.84
	Faith in Intuition	4.94 (1.30)	5.54 (1.59)	.12	-0.41
	Need for Cognition	61.87 (11.15)	61.53 (12.09)	.91	0.03

Note. Between-group comparisons were made using independent-sample *t* tests, aside from gender and handedness which were examined using Pearson's chi square. Standard deviations are in parentheses. ^aSubscales of the 21-item Depression Anxiety and Stress Scale; ^bSubscales of the Positive and Negative Affect Schedule; ^cScores reflect number of correct solutions; MMSE = Mini-Mental State Examination; DSST = Digit Symbol Substitution Test; Mill Hill = Mill Hill Vocabulary Test; BNT = Berlin Numeracy Test; Lipkus Numeracy = Lipkus Numeracy Task. *d* = Cohen's *d* for standardized effect size estimates of the mean difference between groups.

Experimental Design

The experiment was a 2 (age: young, old) x 3 (decision situation: RconfB, CconfB, RalignB) mixed factorial design. Factors and dependent variables included in this experiment are described in the General Methods section (pages 53-55, 61-62).

Procedure

Belief updating task. Instructions for the belief updating task were provided and participants completed the practice task, with trials in the belief updating task following the exact sequence described in the General Methods section (pages 48-53). See Figure 1 for a schematic of a single trial sequence.

Cognitive and affective assessments. All participants completed the battery of assessments outlined in the General Method section (pages 56-61), in addition to the Stroop task.

Stroop task. The Stroop task (MacLeod, 1991) was unique to Experiment 1 and was added to provide an index of inhibitory control and examine whether the Stroop effect associated with performance on the belief updating task. The Stroop colour-naming task was administered to participants to obtain a measure of prepotent response inhibition. Participants first completed a practice task consisting of 40 trials, in which participants had to press a key that corresponded to the colour of a string of X's on the computer screen. A second practice task consisting of 24 trials was completed, and built on the first practice trial by adding words of colours that were either congruent or incongruent with the colour they were presented in. Once the two practice tasks were complete and the experimenter clarified any questions the participant had, the experimental Stroop task, consisting of 216 trials, began. In the Stroop task, participants were required to make a decision in response to the colour of the word by pressing a corresponding key on the keyboard. The meaning of the word would either be congruent with the colour (the word "Blue" presented in the colour blue) or incongruent with the colour of the word (the word

“Blue” presented in the colour purple; Stroop, 1935). In neutral trials, participants were required to respond to the colour of a string of X’s by pressing the corresponding key on the keyboard. After every response, participants received feedback on whether they were correct or incorrect, as well as their RT. Upon completion of the Stroop task, the experimenter then debriefed the participant by explaining the goals of the study, followed by appropriate compensation.

Stroop stimuli. Concerning stimuli in the Stroop task, colour words were presented in Arial font, size 35, in colours: yellow (RGB value: 255, 255, 0), brown (RGB value:165, 42, 42), blue (RGB value: 0-0-255) and purple (RGB value: 128, 0, 128) on a black background (RGB value: 0-0-0), with instructions presented in white (RGB value: 255-255-255). The Stroop task in Experiment 1 was programmed and presented using E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA).

Data Analysis

Data analysis for Experiment 1 followed the procedures outlined in the General Methods section (pages 62-64), with the addition of analyzing data obtained from the Stroop task. First, RT data from neutral trials was subtracted from incongruent trials to obtain the Stroop effect (MacLeod, 1991). To test whether Stroop effect data was associated with decision accuracy in the belief updating task, separate Pearson bivariate correlational analyses for younger and older adults were conducted on Stroop effect data and decision accuracy in the RconfB situation, and Stroop effect data and decision accuracy in the CconfB situation.

Results

Decision Accuracy for All Decision Situations

Table 5 presents the average decision accuracy for each decision situation amongst younger and older adults. Collapsing across age group, independent sample *t* tests were

conducted across counterbalancing version on decision accuracy for each decision situation of interest within each condition. No differences in counterbalancing version were observed in the decision situations or conditions ($ps > .05$).

Table 5. Means and Standard Deviations for Accuracy in Experiment 1

Prior	Sample				
	0	1	2	3	4
75	89.17 (22.19)	82.94 (28.28)	53.67 (33.24)	97.78 (7.24)	97.22 (8.84)
	76.25 (31.72)	69.83 (34.55)	56.78 (39.58)	93.89 (12.75)	96.11 (11.32)
	8	60	60	6	6
50	96.25 (10.46)	96.25 (10.46)	91.56 (12.91)	96.39 (5.55)	97.50 (6.05)
	92.50 (19.59)	90.42 (17.58)	87.95 (14.45)	94.27 (8.73)	91.25 (17.11)
	8	8	60	60	8
25	96.11 (11.32)	93.33 (19.07)	93.89 (16.65)	49.55 (33.62)	73.94 (28.96)
	94.44 (11.01)	95.42 (8.36)	95.55 (9.72)	52.66 (37.94)	67.49 (30.95)
	6	8	6	60	60

Note. Values represent proportion (%) of correct decisions; Prior = prior probability for the left urn; Sample = number of majority colour balls of sample; Younger adults' accuracy is shown on the top line of each row, older adults' accuracy is shown on the second line. Standard deviations are in parentheses. The number of trials are indicated on the third line of each row. Bold font indicates decision situations of interest.

Combining like-situations. Collapsing across age groups, paired-sample t tests were conducted to examine whether decision accuracy of like-situations differed from each other. Results showed a significant difference between RconfB situations, such that the prior: 25, sample: 3 situation produced lower accuracy than the prior: 75, sample: 2 situation, $t(59) = -2.44$, $p = .02$, $d = -0.31$. A significant difference was also observed between CconfB situations, such that the prior: 25, sample: 4 situation produced lower accuracy than the prior: 75, sample: 1 situation, $t(59) = -2.14$, $p = .04$, $d = -0.29$. Last, a significant difference was also observed between RalignB situations, such that the prior: 50, sample: 2 situation produced lower accuracy than the prior: 50, sample: 3 situation, $t(59) = -3.56$, $p = .001$, $d = -0.50$. Decision situations

were collapsed regardless of these differences, as the primary interest was to investigate differences in decision accuracy across categories of situations.

Decision Accuracy

To assess whether participants' decision accuracy increased over time, trials were first segmented into 4 blocks for each decision situation. A 2 (age: young, old) $\times$ 3 (decision situation: RconfB, CconfB, RalignB) $\times$ 4 (trial block: Block 1, Block 2, Block 3, Block 4) mixed ANOVA on decision accuracy was conducted and results showed a main effect of decision situation, $F(1.20, 69.77) = 25.64, p < .001, \eta_p^2 = .31^8$. Planned comparisons showed that the RconfB situation ($M = .53, SD = .35$) promoted less accurate decisions than the CconfB situation ($M = .74, SD = .29, t(59) = -1.97, p = .01, d = 0.36$, and the RalignB situation ($M = .93, SD = .09, t(59) = -2.73, p < .001, d = 1.03$, with the CconfB situation promoting less accurate decisions than the RalignB situation, $t(59) = -1.39, p < .001, d = 0.71$. Critically, no significant main effect of age was observed, $F(1, 58) = 1.22, p = .27, \eta_p^2 = .02$, nor any other significant effects. These results suggest that decision accuracy did not change over time or as a function of age (see Figure 2).

⁸ Two outliers in the younger adult group were identified (see Figure 4, bottom right corner of scatterplot). The mixed ANOVA was rerun with these outliers excluded and results did not change from the description above. Therefore, the two outliers were included in analysis.

Figure 2. Accuracy proportion for each decision situation across trial block in Experiment 1.

Decision RT

To assess whether participants' decision RT increased across trial blocks, a 2 (age: young, old) $\times$ 3 (decision situation: RconfB, CconfB, RalignB) $\times$ 4 (trial block: Block 1, Block 2, Block 3, Block 4) mixed ANOVA on RT was conducted, with results revealing a main effect of age, $F(1, 58) = 18.49, p < .001, \eta_p^2 = .24$ such that younger adults ($M = 2112.27, SD = 1525.96$) made faster decisions than older adults ($M = 4053.07, SD = 3156.59$), a main effect of trial block, $F(1.50, 86.88) = 15.33, p < .001, \eta_p^2 = .21$, and a main effect of decision situation, $F(1.51, 87.79) = 24.01, p < .001, \eta_p^2 = .29$.

There was also a significant Decision Situation by Age interaction, $F(1.51, 87.49) = 5.03, p < .01, \eta_p^2 = .08$. In order to unpack the interaction, independent sample t tests were conducted to examine potential age differences in each decision situation, and repeated-measure ANOVAs with pairwise comparisons were used to probe decision situation effects within each

age group. Independent sample *t*-tests showed that younger adults ($M = 2298.76$, $SD = 1024.57$) made faster decisions in the RconfB situation compared with older adults ($M = 4326.97$, $SD = 2601.10$), $t(58) = -3.97$, $p < .001$, $d = 1.03$. Similar results were observed in the CconfB situation, such that younger adults ($M = 2344.21$, $SD = 1525.96$) made faster decisions compared with older adults ($M = 4833.71$, $SD = 3156.59$), $t(58) = -3.89$, $p < .001$, $d = 1.00$, and the RalignB situation, with younger adults ($M = 1692.18$, $SD = 746.41$) making faster decisions than older adults ($M = 2971.51$, $SD = 1500.61$), $t(58) = -4.18$, $p < .001$, $d = 1.08$. A repeated-measures ANOVA showed significant RT differences across decision situations for younger adults, $F(2, 58) = 7.75$, $p < .001$, $\eta_p^2 = .21$. Pairwise comparisons showed no RT differences between the RconfB and CconfB situations, $t(29) = -.24$, $p = .81$, $d = -0.03$. However, significant RT differences were observed between RconfB and RalignB situations, $t(29) = 2.51$, $p = .001$, $d = 0.76$, as well as between CconfB and RalignB, $t(29) = 2.91$, $p = .02$, $d = 0.55$. For older adults, results showed significant RT differences across decision situations, $F(1.37, 39.70) = 16.26$, $p < .001$, $\eta_p^2 = .36$. Differing from younger adults, pairwise comparisons showed differences in RT between the RconfB and CconfB situations, $t(29) = -1.80$, $p = .02$, $d = -0.44$. Similar to younger adults, pairwise comparisons showed differences between the RconfB and RalignB, $t(29) = 3.97$, $p = .002$, $d = 0.71$, as well as between CconfB and RalignB situations, $t(29) = 4.18$, $p < .001$, $d = 0.81$.

Further, a significant Decision Situation by Trial Block interaction on RT was observed, $F(2.51, 145.86) = 3.39$, $p = .003$, $\eta_p^2 = .06$. To probe the interaction, repeated-measures ANOVAs with pairwise comparisons were conducted to examine potential trial block effects within each decision situation on RT, as well as potential differences in RT for each decision situation within each trial block. For the RconfB situation, trial block was significant, $F(1.63,$

96.13) = 16.73, $p < .001$, $\eta_p^2 = .22$. Pairwise comparisons showed a significant difference between trial Block 1 ($M = 4234.38$, $SD = 2843.38$) and 2 ($M = 3383.04$, $SD = 2429.49$), $t(59) = 3.41$, $p = .01$, $d = 0.44$, between Block 2 and 3 ($M = 2989.23$, $SD = 2336.75$), $t(59) = 2.21$, $p = .03$, $d = 0.29$, and between Block 3 and 4 ($M = 2657.19$, $SD = 2224.24$), $t(59) = 3.55$, $p = .001$, $d = 0.46$, such that RT became faster over time. Concerning the CconfB situation, trial block was significant, $F(1.39, 82.05) = 9.53$, $p < .01$, $\eta_p^2 = .14$. Specifically, pairwise comparisons revealed a significant difference between trial Block 1 ($M = 4600.13$, $SD = 3913.97$) and 2 ($M = 3471.50$, $SD = 2692.45$), $t(59) = 3.24$, $p = .002$, $d = 0.42$, no difference between Blocks 2 and 3 ($M = 3305.13$, $SD = 3111.36$), $t(59) = .90$, $p = .37$, $d = 0.12$, as well as a significant difference between Blocks 3 and 4 ($M = 2987.07$, $SD = 2673.49$), $t(59) = 2.09$, $p = .04$, $d = 0.27$, such that RT became faster over time. Last, for RalignB situations, trial block was significant, $F(1.98, 116.58) = 9.68$, $p < .001$, $\eta_p^2 = .14$. Pairwise comparisons revealed no differences between trial Block 1 ($M = 2766.86$, $SD = 1612.12$) and 2 ($M = 2502.83$, $SD = 1856.31$), $t(59) = 1.86$, $p = .07$, $d = 0.24$, but a significant difference between Blocks 2 and 3 ($M = 2123.11$, $SD = 1453.61$), $t(59) = 2.19$, $p = .03$, $d = 0.28$, and no difference in RT between Blocks 3 and 4 ($M = 1971.49$, $SD = 1246.24$), $t(59) = 1.69$, $p = .10$, $d = 0.22$.

Additionally, a repeated-measures ANOVA was conducted with decision situation and trial block as factors on RT and showed a significant effect, $F(1.28, 75.84) = 19.13$, $p < .001$, $\eta_p^2 = .25$. To probe this effect further, within the first trial block, decision situation was significant, $F(1.27, 72.27) = 18.83$, $p < .001$, $\eta_p^2 = .25$, such that RTs did not differ between the RconfB and CconfB situations, $t(59) = -1.54$, $p = .13$, $d = -0.20$, but did significantly differ between the RconfB and RalignB situation, $t(59) = 5.65$, $p < .001$, $d = 0.73$, and the CconfB and RalignB situation, $t(59) = 4.43$, $p < .001$, $d = 0.57$. Similarly, within the second trial block, decision

situation was significant, $F(1.76, 100.56) = 12.29, p < .001, \eta_p^2 = .18$, such that RTs did not differ between the RconfB and CconfB situations, $t(59) = -.50, p = .62, d = -0.06$ but did show a significant difference between the RconfB and RalignB situation, $t(59) = 4.06, p < .001, d = 0.52$, and the CconfB and RalignB situation, $t(59) = 3.92, p < .001, d = 0.51$. Decision situation was also significant in the third trial block, $F(1.55, 88.19) = 13.83, p < .001, \eta_p^2 = .20$. Similar to the other two blocks, no difference was observed between the RconfB and CconfB situations, $t(59) = -1.77, p = .08, d = -0.23$ but a significant difference was observed between the RconfB and RalignB situation, $t(59) = 4.02, p < .001, d = 0.52$, and the CconfB and RalignB situation, $t(59) = 4.17, p < .001, d = 0.54$. Last, decision situation was significant in the fourth trial block, $F(1.62, 92.30) = 12.70, p < .001, \eta_p^2 = .18$. In line with the pattern of results above, no difference was observed between the RconfB and CconfB situations, $t(59) = -1.19, p = .10, d = -0.15$, but a significant difference was observed between the RconfB and RalignB situations, $t(59) = 3.22, p = .01, d = 0.42$, and the CconfB and RalignB situation, $t(59) = 4.23, p < .001, d = 0.55$ (see Figure 3).

Figure 3. Reaction time for each decision situation across trial block in Experiment 1.

Interestingly, accuracy did not correlate with RT in the RconfB situation, $r = .15$, $p = .27$, or in the CconfB situation, $r = -.17$, $p = .20$. However, accuracy significantly correlated with RT in the RalignB situation, $r = -.37$, $p = .004$.

Together, the results show that the RconfB situation produced the most inaccurate decisions, the CconfB situation produced intermediate decision accuracy, and the RalignB situation produced the most accurate decisions, with accuracy unaffected by age or trial block. Concerning RT, younger adults made faster decisions than older adults for each decision situation. Younger adults showed no difference in RT between conflict situations but did show slower RTs in conflict situations compared with the RalignB situation. Older adults showed the slowest RTs in the CconfB situation, intermediate RTs for the RconfB situation, and the fastest RTs for the RalignB situation. RTs became comparably faster across trial blocks for conflict situations but not for the RalignB situation. Last, only accuracy and RT in the RalignB situation correlated.

Relationship Between Representativeness and Conservatism

Pearson bivariate correlational analyses were conducted on decision accuracy between the RconfB and CconfB situations separately for younger and older adults to elucidate whether there was a relationship between reliance on representativeness and conservatism. A significant negative correlation between conflict situations emerged for younger adults, $r = -.43$, $p = .02$, however, after removal of the two outliers (bottom right of scatterplot; see Figure 4) the correlation was no longer significant, $r = -.21$, $p = .29$. Further, a significant negative correlation between the RconfB and the CconfB situation emerged for older adults, $r = -.58$, $p = .001$ (see Figure 5). The negative correlation implies that as accuracy increases in one decision situation, accuracy decreases in the other decision situation. This suggests that younger adults may have switched strategies to make decisions, whereas older adults may have chosen one decision strategy (e.g., overweighting new information).

Figure 4. Association of accuracy in conflict situations in Experiment 1 for younger adults.

Figure 5. Association of accuracy in conflict situations in Experiment 1 for older adults.

Concerning self-assessment data, chi-square tests were conducted on categorical items (i.e., items 1 and 2) of the self-assessment and independent-sample *t* tests were conducted on scaled items (i.e., items 3-5) of the self-assessment for each age group. Results showed that most participants reported using “both” the prior and sample to make decisions, “switching” decision strategies, reporting that they felt “somewhat accurate” in the decisions they made, that the task was “somewhat difficult”, and that they are “somewhat confident” or “confident” working with numbers in general. No age differences were observed (see Table 6).

Table 6. *Self-Assessment for Experiment 1*

Item	Younger Adults (n = 30)	Older Adults (n = 30)	<i>p</i> -value	<i>d</i>
1. Influence Decision				
Likelihood	5 (16.7%)	6 (20%)	.08	0.28
Sample	2 (6.7%)	8 (26.7%)		
Both	23 (76.7%)	16 (53.3%)		
2. Decision Strategy				
Consistent	6 (20%)	13 (43.3%)	.08	0.29
Switched	24 (80%)	16 (53.3%)		
Neither	0 (0%)	1 (3.3%)		
3. Accuracy				
Very inaccurate	1 (3.3%)	0 (0%)	.75	0.07
Somewhat inaccurate	4 (13.3%)	3 (10%)		
Neither inaccurate nor accurate	1 (3.3%)	7 (23.3%)		
Somewhat accurate	23 (76.7%)	20 (66.7%)		
Very accurate	1 (3.3%)	0 (0%)		
4. Task Difficulty				
Very difficult	0 (0%)	3 (10%)	.29	0.28
Somewhat difficult	13 (43.4%)	12 (40%)		
Neither difficult nor easy	11 (36.7%)	10 (33.3%)		
Somewhat easy	3 (10%)	4 (13.3%)		
Very easy	3 (10%)	1 (3.3%)		
5. Confidence with Numbers				
Not at all confident	4 (13.3%)	4 (13.3%)	1.00	0.94
Somewhat confident	14 (46.7%)	11 (36.7%)		
Confident	7 (23.3%)	12 (40%)		
Very confident	4 (13.3%)	3 (10%)		
Extremely confident	1 (3.3%)	3 (10%)		

In summary, correlational analyses showed no relationship in decision accuracy between conflict situations in younger adults after two outliers were removed, however the negative correlation observed for older adults suggests that decisions made may have been primarily based on new information (i.e., the sample). Finally, most participants reported using both prior and sample information to make decisions, making “somewhat accurate” decisions, reported that the task was “somewhat difficult” and that they are “somewhat confident” working with numbers in general, with no age differences emerging in the self-assessment questionnaire.

Individual Differences on Accuracy

First, the ICC was calculated to observe how much variability in the outcome was due to clustering. Results showed that approximately 11% of the variance in decision accuracy was due to clustering. Although there is no hard rule indicating how large an ICC is large enough to suggest non-negligible nesting, these results suggest a portion of the variability of accuracy is due to clustering. In light of this, a multi-level modeling approach was used to examine whether assessments were good predictors of decision accuracy on the belief updating task.

Next, a random-intercepts model and a random-slopes model on accuracy within RconfB and CconfB decision situations, with age included as a Level 2 predictor, were compared. Based on the AIC, BIC, and deviance values, the random-slopes model produced better model fit than the random-intercepts model. Therefore, individual difference measures of interest were included as predictors separately in the random-slopes models (see Table 7).

Table 7. *Multi-Level Model Comparison on Accuracy in Experiment 1*

Parameters		Model	
		Random-intercepts	Random-slopes
Fixed effects			
Intercept		1.27 (0.16)*	1.89 (0.43)*
	OR	3.57	6.65
Younger adults		-0.15 (0.23)	-0.36 (0.61)
	OR	0.86	0.69
RconfB		-1.01 (0.04)*	-1.5 (0.89)*
	OR	0.36	0.21
Interaction		-	-0.02 (1.27)
		-	0.98
Fit statistics			
deviance		16,719	11,187
AIC		16,727	11,199
BIC		16,758	11,245
df		4	6
Model comparison		$\chi^2 = 5,531.9^*$	

Note. Standard errors and deviations in fixed and random effects parentheses, respectively; RconfB = Representativeness heuristic conflicts with Bayes' theorem; Younger adults and RconfB were used as the reference group; OR = odds ratios; AIC = Akaike's Information Criterion; BIC = Bayesian Information Criterion; df = degrees of freedom; Subscript χ^2 denotes the two models being compared. * $p < .05$.

Berlin Numeracy Test. Including the Berlin Numeracy Test (BNT) as a predictor revealed a significant BNT by Decision Situation interaction on decision accuracy ($\beta = 1.24$, $SE = 0.84$; $p = .02$). To probe this interaction, BNT scores were centered around the median to examine how scores on the BNT related to decisions situations, specific to those who score around the median level on the BNT. To investigate the relation between BNT and decision situation on decision accuracy for those scoring below and above the median on the BNT, the model was re-estimated three times after centering the BNT variable at the 25th percentile, the median and the 75th percentile. As presented in Table 8, beta estimates for the effect of decision situation increase as scores on the BNT increase. This suggests that the effect of decision situation is strongest amongst those with lower BNT scores. The odds ratios suggest that among those with lower BNT scores, the odds of a correct response in the RconfB situation are only 4% of the odds of a correct response in the CconfB situation. However, among those with higher BNT scores, the odds of a correct response in the RconfB situation are 41% of the odds of a correct response in the CconfB situation. Together, the results indicate that the effect of decision situation becomes greater with lower scores on the BNT.

Table 8. *Berlin Numeracy Test as a Predictor of Accuracy in Experiment 1*

	BNT centered at 25 th percentile					BNT centered at median					BNT centered at 75 th percentile				
	β	SE	Z	<i>p</i>	OR	β	SE	Z	<i>p</i>	OR	β	SE	Z	<i>p</i>	OR
Intercept	2.45	0.64	3.81		11.56	2.06	0.43	4.75		7.85	1.67	0.45	3.71		5.33
BNT	-0.39	0.35	-1.10	.04	0.68	-0.39	0.36	-1.08	.04	0.68	-0.39	0.35	-1.11	.04	0.68
RconfB	-3.21	1.28	-2.51	.01	0.04	-2.06	0.87	-2.37	.01	0.13	-0.90	0.91	-0.99	.01	0.41
age	-0.70	0.81	-0.87	.36	0.49	-0.55	0.59	-0.93	.36	0.58	-0.39	0.68	-0.57	.36	0.68
BNT:RconfB	1.15	0.70	1.64	.02	3.16	1.15	0.72	1.59	.02	3.16	1.15	0.69	1.65	.02	3.16
BNT:age	0.16	0.47	0.34	.47	1.17	0.16	0.47	0.34	.47	1.17	0.16	0.46	0.35	.47	1.17
RconfB:age	0.65	1.61	0.40	.61	1.91	0.63	1.18	0.53	.59	1.87	0.61	1.36	0.44	.60	1.83
BNT:RconfB:age	-0.02	0.93	-0.02	.98	0.98	-0.02	0.95	-0.02	.98	0.98	-0.02	0.92	-0.02	.98	0.98

Note. BNT = Berlin Numeracy Test; RconfB = Representativeness heuristic conflicts with Bayes' theorem (reference group); β = beta estimate; SE = standard error; Z = z-score; OR = odds ratio

The Need for Cognition, Faith in Intuition, Lipkus Numeracy Task, Mill Hill Vocabulary Test assessments, as well as years of education did not significantly predict decision accuracy. In addition, the multi-level model with DSST as a predictor of decision accuracy failed to converge. Therefore, to examine whether DSST scores were related to decision accuracy, bivariate correlations were computed. DSST scores did not correlate with decision accuracy in the RconfB situation, $r = -.09$, $p = .52$, but DSST scores correlated with decision accuracy in the CconfB situation, $r = .31$, $p = .02$. I also examined correlations between DSST scores and decision accuracy within each age group. For younger adults, DSST scores did not correlate with accuracy in the RconfB situation, $r = -.17$, $p = .37$, or the CconfB situation, $r = .13$, $p = .49$. Similarly, for older adults, scores on the DSST did not correlate with decision accuracy in the RconfB situation, $r = -.04$, $p = .82$, or the CconfB situation, $r = .28$, $p = .14$.

Stroop performance. To test the hypothesis concerning whether the Stroop effect was associated with decision accuracy, an independent t-test on the Stroop effect in younger and older adults was first conducted to confirm age-related differences in inhibitory control. Results showed that older adults ($M = 211.04$, $SD = 134.27$) produced a larger Stroop effect than younger adults ($M = 89.27$, $SD = 73.45$), $t(58) = -4.36$, $p < .001$, $d = 1.13$. Next, Pearson bivariate correlational analyses were conducted across age group, as well as separately for younger and older adults⁹. Analysis revealed no correlation between the Stroop effect and accuracy in the RconfB situation, $r = -.08$, $p = .52$, or the CconfB situation, $r = -.08$, $p = .55$. Further, results showed no correlation between the Stroop effect and decision accuracy in the

⁹ Stroop effect data was first included into a multi-level model to predict decision accuracy. However, the model would not run due to the amount of data in the dataframe. Therefore, I resorted to correlational analyses to examine the relationship between the Stroop effect and decision accuracy.

RconfB situation for younger adults, $r = .10$, $p = .64$, nor for older adults, $r = -.25$, $p = .57$.

Similarly, results showed no correlation between the Stroop effect and decision accuracy in the

CconfB situation for younger adults, $r = .04$, $p = .32$, nor for older adults, $r = .19$, $p = .18$.

Correlational analyses were also conducted across age group, as well as separately for younger and older adults between the Stroop effect and decision RT on the belief updating task. Similar to analyses conducted on decision accuracy, no significant correlations were observed on decision RT.

In summary, the results showed that numeracy, as indexed by the Berlin Numeracy Test, is a good predictor of decision accuracy, such that those with higher Berlin Numeracy Test scores showed smaller differences in decision accuracy between RconfB and CconfB situations (higher accuracy in both), and those with lower Berlin Numeracy Test scores showed greater differences in decision accuracy between RconfB and CconfB situations (lower accuracy in the RconfB situation and higher accuracy in the CconfB situation). Additionally, collapsing across age, scores on the DSST correlated with accuracy in the CconfB situation only. No other assessments, including the Stroop task, were associated with decision accuracy on the belief updating task.

Summary of Results

The RconfB situation produced the most errors, followed by the CconfB situation and the RalignB situation. Accuracy was unaffected by age or trial block. Although younger adults did not show a correlation in accuracy between conflict situations, older adults did, such that reliance on the representativeness heuristic was associated with non-reliance on the conservatism heuristic. Concerning RT, younger adults made faster decisions than older adults overall, and responses became faster in conflict situations across blocks in both age groups. Whereas younger

adults showed no difference in RT in conflict situations, the CconfB situation produced longer RTs than the RconfB situation in older adults. Additionally, most participants reported using both the “prior and sample” to make decisions, making “somewhat accurate” decisions, reporting the task was “somewhat difficult” and that they were “somewhat confident” working with numbers in general, regardless of age. Last, scores on the Berlin Numeracy Test predicted decision accuracy in the belief updating task. Additionally, collapsing across age, scores on the DSST correlated with accuracy in the CconfB situation only. No other assessments were associated with belief updating performance, including the Stroop task.

Discussion

The heuristics and biases literature show that younger adults are prone to systematic biases in cognitive reflection tasks because heuristics are used in lieu of formal analysis (Kahneman & Tversky, 1972). For example, Achtziger and colleagues (2014) showed that younger adults commonly committed representativeness and conservatism errors rather than following the principles of Bayes’ theorem in a belief updating task. However, it remained unknown how age would affect the use of heuristics when updating beliefs, or whether individual differences such as numeracy, thinking dispositions, or inhibitory control modulate performance in belief updating in younger and older adults. In light of this gap in the literature, the goal of the current experiment was to test hypotheses about the influence of age, as well as individual differences in the use of heuristics during belief updating.

Given the evidence showing that older adults rely on automatic processes and use heuristics more than younger adults (Johnson, 1990; Peters et al., 2007), I hypothesized that reliance on heuristics to make decisions would be greater in older than in younger adults (Hypothesis 1). In contrast to this hypothesis, no age differences in the use of heuristics were

observed in this experiment. These results are the first to show that when updating beliefs, younger and older adults use heuristics to the same degree. The results also support previous findings that show no age differences when making decisions under risk (Armstrong & Spaniol, 2017). These results exemplify the difficulty of following the principles of Bayes' theorem and the tendency to resort to heuristics in both young and old age when updating beliefs.

Additionally, it has been argued that base rate-neglect is commonly observed because base-rates (i.e., prior information) are usually expressed as probabilities that engage Type 2 processes, and diagnostic (or new) information tends to evoke Type 1 processes because this information is perceived as more concrete and vivid (Barbey & Sloman, 2007). Thus, I hypothesized that the representativeness error would be committed more than the conservatism error over all (Hypothesis 2), but with older adults committing the representativeness error more than younger adults, and with no age differences in committing the conservatism error (Hypothesis 3). The results revealed a main effect of decision situation, such that susceptibility to the representativeness heuristic was greater than susceptibility to the conservatism heuristic. These results not only replicate prior findings that have examined younger adults' use of the representativeness and conservatism heuristics in belief updating (Achtziger et al. 2014; Grether, 1980; Dave & Wolfe, 2003), they extend this literature to the older adult population and show evidence that the representativeness heuristic is more influential than the conservatism heuristic. However, in contrast to Hypothesis 3, no Age by Decision Situation interaction was observed. Younger and older adults used both heuristics to the same degree.

The results partially supported Hypothesis 4. Prior literature shows that younger adults can learn without feedback (Achtziger et al. 2014; Kelley & Mclaughlin, 2012) whereas older adults have difficulty learning without feedback (Eppinger et al. 2015; Herbert et al. 2011). It

was therefore hypothesized that younger adults would show an increase in decision accuracy and faster RTs across trial blocks, whereas older adults would show stable levels of accuracy and RT across trial blocks. However, given that accuracy did not improve across trials, the results do not provide evidence of a learning effect¹⁰ in either age group. Although RT decreased in younger and older adults across trial blocks, this may reflect familiarity with the task over time. Why did younger adults show a learning effect in Achtziger et al.'s (2014) study but not in the current experiment? An important methodological difference between the two studies concerns how prior and new information were presented. Whereas Achtziger et al. (2014) presented the urn distributions and prior information separately from new information, the current study presented all information simultaneously. As a reminder, the representativeness heuristic has been described as comparing new information to an existing parent population (Grether, 1980). Therefore, simultaneous presentation of information may have made using the representativeness heuristic more compelling compared with presenting new information and then having to recall the urn distributions and prior information from memory. In other words, perhaps presenting relevant information together supported committing representativeness errors throughout the task, whereas forcing participants to recall the parent population distributions and prior information from memory gradually influenced reflective processing over time, as was observed in Achtziger et al. (2014).

Further, age differences may not have been observed because feedback was not presented. Prior research shows that younger adults are more successful at using feedback to improve performance compared with older adults (Eppinger et al., 2008; Nieuwenhuis et al., 2002). In the belief updating task, it may be too difficult to gauge performance or strategize

¹⁰ I refer to “learning” as making more accurate decisions over time.

about ways to improve performance without feedback. At present, it is unknown whether younger and older adults would show a learning effect in a belief updating task with feedback.

Supporting Hypothesis 5, younger adults showed faster RTs than older adults (Salthouse, 2019), with younger adults showing no difference in RT between conflict situations, but with older adults showing slower RTs when the conservatism heuristic conflicted with Bayes' theorem relative to the representativeness heuristic. Although speculative, this may be evidence that older adults reflect more when the conservatism heuristic conflicts with Bayes' theorem. More research is needed to further examine differences that may underlie use of the representativeness and conservatism heuristics when making decisions in old age. Last, both younger and older adults made faster decisions across trial blocks, which could be due to increased familiarity with the task.

Concerning decision strategies, after two outliers were removed, no relationship was observed in decision accuracy between conflict situations in younger adults. However, a negative correlation was observed in older adults indicating that reliance on the representativeness heuristic predicted non-reliance on the conservatism heuristic. Although initially it may seem as though younger and older adults are less susceptible to the conservatism heuristic compared with the representativeness heuristic, an alternative explanation may underlie these results. Drawing the reader's attention to the six decision situations of interest prior to collapsing across like-situations (Table 1), one may take the perspective of choosing one strategy over the other and applying this to each of the six decision situations of interest to the current experiment.

Consider first applying the conservatism heuristic to each decision situation (i.e., using only the prior with the higher probability to make each decision). One would make correct decisions in the RconfB situations, incorrect decisions in the CconfB situations, and given

chance performance due to equal priors, one would make 50% of the decisions correctly in the RalignB situations. Next, apply the representativeness heuristic to each decision situation (i.e., using only new information to make decisions by comparing the sample to each urn distribution). One would make incorrect decisions in the RconfB situations, correct decisions in the CconfB situations, and correct decisions in the RalignB situations. The critical point here is that if one consistently uses new information to make decisions across all decision situations, by default one would obtain 100% accuracy in the CconfB and RalignB situations. Mean decision accuracy in the CconfB situation may not be indicating an avoidance of the conservatism error, but rather indicating use of the representativeness heuristic, which in this situation leads to correct responses. That is, younger and older adults may be applying the representativeness heuristic more and the conservatism heuristic less than what the means suggest. This possibility was not discussed in Achtziger et al. (2014) who reported similar results concerning decision errors in younger adults.

Research shows that the ability to inhibit inappropriate responses is paramount for reasoning (De Neys, Schaeken & d'Ydewalle, 2005; Handley, Capon, Beveridge, Dennis & Evans, 2004; Houdé, 1997; Markovits & Doyon, 2004; Moutier, Plagne-Cayeux, Melot, & Houdé, 2006; Stanovich & West, 2000). It has been posited that older adults may make more heuristic-based decisions than younger adults (De Neys, & Van Gelder, 2008) because inhibitory control decreases with age (Hasher & Zacks, 1988; Lustig, Hasher, & Zacks, 2007). Concerning the processes that may underlie belief updating, it was of interest to examine whether inhibitory control, indexed by the Stroop effect, associated with performance (Hypothesis 6).

As expected, older adults showed a larger Stroop effect than younger adults (West, 2004; West & Alain, 2000), however the Stroop effect was not associated with performance on the

belief updating task. One reason for this result could be due to task differences. Although not a word problem, the belief updating task, like other dual-process tasks, requires reasoning and rational decision making, whereas the Stroop task specifically indexes inhibiting a prepotent response (Pennycook, 2017). In the current experiment, decisions in the belief updating task took approximately 3 seconds for younger and 5 seconds for older adults, whereas the Stroop effect was approximately 90 milliseconds for younger and 200 milliseconds for older adults. Decision accuracy or RTs on the belief updating task may not relate to Stroop effect scores because engaging inhibitory control is only one process of the many processes that underlie updating beliefs. As previously described, once Type 1 processing is suppressed (i.e., inhibited), Type 2 processing is engaged to generate alternative responses which depends on proficient mindware such as numeracy, and the ability to sustain cognitive decoupling operations (Stanovich, 2011). That is, inhibitory control captured by the Stroop effect may only partially underlie belief updating. However, because inhibitory control in the belief updating task was not captured in isolation, an association between the Stroop effect and performance on the belief updating task may have been washed out. Tasks that more closely index reflective processing such as the Cognitive Reflection Task (Frederick, 2005) or examining inhibitory control activity at the brain level when updating beliefs may further illuminate the processes that underlie this task.

Concerning individual differences captured by the battery of cognitive assessments, it was hypothesized that numeracy, and reflective rather than intuitive thinking dispositions would predict decision accuracy (Hypothesis 7). First, given that highly numerate individuals are less susceptible to some decision biases (Klaczynski, 2014; Peters et al., 2006), and are more likely to engage in deliberation (Pennycook, Trippas, Handley, & Thompson, 2013), it was expected that individual differences in numeracy would predict the use of heuristics in the current experiment.

Supporting this hypothesis, results showed that the Berlin Numeracy Test predicted decision accuracy on the belief updating task, such that higher numeracy was associated with less susceptibility to both heuristics and lower numeracy associated with more prevalent use of heuristics to make decisions, especially the representativeness heuristic. Performance on the Lipkus Numeracy Task did not predict decision accuracy in the belief updating task, replicating findings reported by Hügelschäfer (2011). These results could be due to a ceiling effect and low variability in the scores on the Lipkus Numeracy Task (see Table 4). Future research should further examine why the Berlin Numeracy Test but not the Lipkus Numeracy Task predicted belief updating, and whether there are differences between these two measures of numeracy.

Concerning thinking dispositions, results presented here converge with literature reporting that the Faith in Intuition and Need for Cognition scales do not predict performance on probability judgment tasks (e.g., Alós-Ferrer & Hügelschäfer, 2012; Alós-Ferrer et al., 2016; Lu, 2015). One reason for this could be due to unshared method variance between the two tasks such that the Faith in Intuition and Need for Cognition scales are self-report measures of preferences for using intuition or effortful cognition when making decisions. In contrast, the belief updating task measures cognitive reflection and the ability to avoid decision biases. Tasks that measure intuitive-analytical thinking dispositions through performance-based ability rather than self-report measures may provide a better one-to-one comparison for examining whether individual differences in thinking dispositions associate with performance on the belief updating task. Last, results showed that scores on the DSST correlated with accuracy in the CconfB situation only, and that the Mill Hill Vocabulary Test and years of education did not predict decision accuracy. These results suggest that perceptual-motor speed (i.e., an index of fluid intelligence; Lindenberger, Mayr, & Kliegl, 1993) may associate with updating beliefs and that crystallized

cognitive abilities do not underlie the use of heuristics when updating beliefs. However, these results should be replicated to bolster the generalizability of these findings.

In summary, younger and older adults use heuristics that lead to decision errors to the same degree when updating beliefs, with committing the representativeness error more often than the conservatism error, without learning over time in the absence of feedback. Results from Experiment 1 replicate previous literature, as well as provide novel findings that extend the literature on aging and cognitive biases when updating beliefs about probability information. First, the results replicate evidence that younger adults use representativeness and conservatism heuristics to make decisions (Achtziger et al., 2014; Kahneman & Tversky, 1972) and provides the novel contribution of extending these results to the older population. Second, the results reveal the prevalence of base-rate neglect in both young and old age, supporting prior literature suggesting that diagnostic (or new) information is more intuitive than base-rate (or prior) information, and therefore is more compelling to use when updating beliefs (Kahneman & Frederick, 2002). Finally, these results show that higher numeracy levels may protect individuals from being susceptible to error-prone heuristics when updating beliefs. Although the results from Experiment 1 shed new light on how younger and older adults use heuristics when updating beliefs, the processes that underlie belief updating and how they may change with age remains unclear. Building on the results of Experiment 1, Experiment 2 of this dissertation examines whether feedback influences reflective processing and the use of heuristics in the belief updating task, while also using ERPs to elucidate the processes that underpin belief updating.

Chapter 6: Experiment 2

Experiment 1 of this dissertation showed that both younger and older adults used heuristics to the same degree when updating beliefs with no feedback. Further, the representativeness error was committed more than the conservatism error, with numeracy moderating this outcome, such that higher numeracy predicted less susceptibility to using error-prone heuristics, and lower numeracy predicted more susceptibility to using error-prone heuristics, particularly the representativeness heuristic. However, it remains unknown whether age differences would surface when updating beliefs with feedback. In addition, although age differences in decision accuracy were not observed in Experiment 1, no study has examined whether younger and older adults differ in the neural correlates that underlie belief updating. Moreover, to my knowledge no study has investigated whether individual differences in numeracy, thinking disposition or cognitive reflection associate with differences in the neural correlates that underlie belief updating with feedback. To fill this gap in the literature, the purpose of Experiment 2 was to investigate whether there are age differences in belief updating when feedback is provided, and whether the neural processes that underlie belief updating are modulated by age or individual differences. Thus, Experiment 2 builds on Experiment 1 by a) including feedback in the belief updating task, b) capturing the neural correlates that underlie belief updating, and c) including the Cognitive Reflection Task, an index of reflective processing, to examine whether Cognitive Reflection Task scores associate with performance.

Investigating how age impacts neural activity is important as age differences in task-related brain activity are well documented even when no age differences in behavioural performance emerge (Cabeza & Dennis, 2012; Friedman, 2003, 2012; Grady, 2012). Neuroimaging research shows that in addition to age-related declines in whole-brain volume, reduced connectivity of white matter tracts, and changes in neurotransmitter systems (e.g.,

dopamine; Erixon-Lindroth et al., 2005; Hedden & Gabrieli, 2004), there are marked declines in the structure and function of the PFC, a cortical region that supports executive functions (Band, Ridderinkhof, & Segalowitz, 2002; Raz, 2005; West, 1996). Aging also changes the operation of large-scale functional brain networks. Specifically, while some research shows age-related reductions in brain activity due to declines in cognitive processing (see Grady, 2012), other research shows increases in brain activity with age (Davis, Dennis, Daselaar, Fleck, & Cabeza, 2008, Grady et al. 1994; Rizio & Dennis, 2014) that have been attributed to compensatory (Reuter-Lorenz, 2002; Vallesi, 2011), as well as dedifferentiation mechanisms (Dennis & Cabeza, 2008; Grady, 2012; Park & Reuter-Lorenz, 2009). In particular, a posterior-anterior shift has been observed with age (PASA; Davis et al., 2008), posited to reflect functional compensation of higher-order cognitive resources in the PFC to offset age-related deficits in other brain regions (Cabeza, et al., 2018; Dennis & Cabeza, 2008; Grady, 2012).

Existing studies on belief updating have employed functional magnetic resonance imaging (fMRI), as well as ERPs to investigate the neural correlates that underpin belief updating. The following section provides a brief overview of this literature.

fMRI Literature

Within the fMRI literature, a variety of paradigms have been used to examine the neural correlates of belief updating. Studies using versions of the urn-ball paradigm have found that the striatum, amygdala, insula and orbitofrontal cortices activated when prior information was presented, whereas the accumulation of new information has been observed in the anterior cingulate cortex (ACC) and parietal regions (Furl & Averbach, 2011), as well as regions comprising the default mode network such as the angular gyri, posterior cingulate, and medial frontal cortex (d'Acremont, Schultz, & Bossaerts, 2013; Vilares, Howard, Fernandes, Gottfried,

& Kording, 2012). The ACC, frontal, and occipital cortices have been shown to activate in response to the presentation of new information, and the insula, striatum, ACC, frontal and parietal regions activate during Bayesian updating (i.e., when an observer detects a change in the environment and updates prior with new information; d'Acremont et al., 2013; Furl & Averbeck, 2011; Poudel, Bhattarai, Dickinson, & Drummond, 2017; Stern, Gonzalez, Welsh, & Taylor, 2010). Together, research shows that prior information activates insular, striatal and frontal regions that have been associated with cognitive control processes (Furl & Averbeck, 2011; Jiang, Beck, Heller, & Egner, 2015), while the presentation of new information has been associated with activation in the ACC, frontal, and parietal cortices (Poudel et al., 2017), regions that associate with conflict monitoring processes and the executive control of information such as response inhibition and attentional control (Braver, Barch, Gray, Molfese, & Snyder, 2001). In addition, belief updating associates with activation of the insula and striatum, in addition to the ACC, frontal and parietal cortices, with the insula-frontostriatal network activated during tasks of cognitive control. Specifically, this network has been shown to mediate flexible cognitive control by adapting to predictions of changing control demands (D'Acremont et al., 2013; Jiang et al., 2015; Uddin, Nomi, Hebert-Seropian, Ghaziri, & Boucher, 2017).

Belief updating tasks that diverge from the commonly used urn-ball paradigm have used probability tasks that present sensory stimuli that require participants to make inferences about probabilistic perceptual or auditory stimuli, or the visuospatial orientation of stimuli. For example, in a visual-perceptual inference task, Summerfield and Koechlin (2008) showed that expectations for prior information were associated with activity in the extrastriate and anterior temporal lobe regions. In addition, the posterior parietal cortex and prefrontal regions (e.g., presupplementary motor area and the dorsal cingulate cortex) activated when new information

was presented, whereas the ACC, putamen, the temporoparietal junction and dopaminergic-rich midbrain regions activated during Bayesian updating (O'Reilly et al., 2013; Schwartenbeck, FitzGerald, & Dolan, 2016; Vossel, Mathys, Stephan, & Friston, 2015). Last, Trapp, Lepsien, Kotz, and Bar (2016) conducted a study using face and object (e.g., house) stimuli, and showed that whereas the fusiform face area activated more for faces that were associated with higher prior probabilities, the parahippocampal place area did not show modulated activation associated with prior probability, suggesting sensitivity to stimulus material.

Localization of regions activated during different stages of belief updating within a Bayesian inference context is inconsistent in the fMRI research reviewed above, likely due to the wide variety of tasks used, with some tasks only indirectly related to Bayesian updating. However, the data suggests that prior information may be represented in the insula-frontostriatal network, and the processing of new information and belief updating may be represented in the ACC and parietal regions, in addition to the insula-frontostriatal network.

ERP Literature

The ERP technique has a high temporal resolution that can track neural activity on a millisecond time scale (Albert, Lopez-Martin, Hinojosa, & Carretie, 2013), representing a continuous measure of processing (Luck, 2014). ERPs are derived from scalp-recorded electroencephalography (EEG) and take form in voltage fluctuations that arise from sensory, motor or cognitive processes that are time-locked to specific events or stimuli (Pires, Leitão, Guerrini, & Simões, 2014). These voltage fluctuations represent a summation of a large number of primarily cortical postsynaptic potential activity (Luck, 2014) that are represented by positive and negative peaks and troughs that vary in amplitude (i.e., magnitude of activity) and latency (i.e., timing of activity; Luck, 2014). An ERP component is a scalp-recorded voltage change

reflecting specific neural and psychological processes (Luck, 2014). For example, an ERP component important for the current dissertation is the N2, with amplitude modulations of the N2 thought to index conflict monitoring processes (Donkers & van Boxtel, 2004; Grützmann, Riesel, Klawohn, Kathmann, & Endrass, 2014). ERP components have traditionally been classified as either exogenous sensory components that are automatically reactive to external factors (e.g., presentation of stimuli) or endogenous components that represent internal factors (e.g., task-related brain processes; Luck, 2014). Analyzing the amplitudes and latencies of ERP components help shed light on the higher-order cognitive processes that underlie behaviour in real time.

Several studies have examined the neural underpinnings of belief updating using ERPs. For example, in a belief updating task, Achtziger et al. (2014) investigated deviations from Bayes' theorem by overweighting prior information (i.e., committing a conservatism error). Results showed that committing the conservatism error was associated with a bias toward one of two options before the presentation of new information, indexed by the lateralized readiness potential (LRP; Leuthold, Sommer, & Ulrich, 1996; Steinhauser, Hübner, & Druey, 2009). The LRP represents a lateralization or shifting of activation in motor cortices to a greater degree on the contralateral side of the response hand (De Jong, Coles, Logan, & Gratton, 1990; Eimer, 1998; Kutas & Donchin, 1980). For example, when the LRP amplitude is stronger in the right hemisphere this suggests increased motor preparedness to choose the left option.

The N2, a fronto-central negative deflection peaking between 200-400 ms after stimulus onset, has been associated with response inhibition (Falkenstein, Hoormann, & Hohnsbein, 1999; Groom & Cragg, 2015), and is considered an index of conflict detection (Donkers & van Boxtel, 2004; Grützmann et al., 2014) on trials that trigger a tendency to make incorrect prepotent

responses (e.g., No-Go trials in the GO/No-Go task). Individual differences in N2 amplitude correspond to the degree of sensitivity toward a response conflict, with larger amplitudes corresponding to higher sensitivity (Amodio, Devine, & Harmon-Jones, 2008; Kirmizi-Alsan et al., 2006). The extent to which the N2 reflects inhibitory control (i.e., regulation of prepotent responses; Bruin, & Wijers, 2002) or conflict monitoring and detection (i.e., contention for the output of two or more response representations; Botvinick, Braver, Barch, Carter, & Cohen, 2001), or both, on conflict trials is still debated (see Smith, Johnstone, & Barry, 2007).

In Achtziger et al.'s (2014) study, committing a representativeness error was associated with less sensitivity to conflict detection indexed by a smaller N2, whereas avoiding the representativeness error was associated with greater sensitivity of detecting conflict, indexed by a larger N2 deflection. Achtziger and colleagues (2014) noted that the results likely do not reflect an explicit detection of conflict between a heuristic and Bayes' theorem because it is unlikely participants were able to make the mental calculations using Bayes' theorem within the ~300 ms time interval of the N2. Rather, the authors speculate that the enlarged N2 amplitude may reflect a distinction between relying on a representativeness heuristic and engaging inhibitory processes when faced with conflict (Achtziger et al., 2014). In another study using a belief updating task, results showed that N2 amplitude increased when prior information was ambiguous (Seer et al., 2016). These results support the notion that the N2 is a marker of different types of conflict such as when information is incompatible or when information is ambiguous.

In the context of belief updating, the role of the P300 component has also been investigated. The P300 encompasses two subcomponents, the P3a and P3b (O'Connell et al., 2012; Polich, 2007). The P3a is observed around 250 ms following stimulus onset and has a fronto-central distribution, whereas the P3b is evoked around 300-600 ms following stimulus

onset and has a more parietal distribution (Fjell, & Walhovd, 2001). The P3a component has been associated with involuntary attention and the processing of novelty, whereas the P3b has been linked to the need for more attentional resources after a conflict has been detected on conflict trials (Larsen, Clayson, Primosch, Leyton, & Steffensen, 2015; Ostwald et al., 2012; Polich, 2007).

Within the context of belief updating, the P3b component has commonly been associated with the unexpectedness of new information (Kolossa, Kopp, & Fingscheidt, 2015; Mars et al., 2008; Seer et al., 2016), as well as stimulus evaluation in order to select appropriate responses (Huster, Enriquez-Geppert, Lavalley, Falkenstein, & Herrmann, 2013; Polich, 2007). Relative to the N2, the functional meaning of the P3b during conflict trials has more consistently been related to inhibition and goal-directed attention (Enriquez-Geppert, Konrad, Pantev, & Huster, 2010; Fallgatter & Strik, 1999; Huster et al., 2013; Smith, Johnstone, & Barry, 2008). Achtziger et al. (2014) observed a significant P300 deflection following the N2. Specifically, in situations in which the representativeness heuristic conflicted with Bayes' theorem, a *large* N2 amplitude was associated with a *small* P300 amplitude. Conversely, in situations in which the conservatism heuristic conflicted with Bayes' theorem, a *small* N2 amplitude was associated with a *large* P300 amplitude. These findings may suggest that a conflict between the representativeness heuristic and Bayes' theorem is detected at any earlier time point, compared with inhibitory processes engaged at a slightly later time point in situations in which the conservatism heuristic conflicts with Bayes' theorem.

Additionally, Kolossa et al. (2015) using a belief updating task, and Bennett, Murawski, and Bode (2015) using a perceptual-learning task, showed that the enhancement of the P3a was associated with Bayesian updating, referring to changes in probability distributions given new

observations (i.e., the belief update size calculated as the difference between prior and posterior beliefs). Last, higher prior probabilities have been related to larger P3a amplitudes compared with lower prior probabilities, and more ambiguous likelihoods have been related to larger P3b amplitudes (Kopp et al., 2016).

In comparison to the results of the existing fMRI research, results from the ERP research on belief updating are more consistent. Based on the research reviewed, belief updating tasks are commonly used, with results showing that the P3a associates with belief updating and the P3b associates with inhibitory control for further information processing. More research is required to interpret the role of the LRP and N2 during belief updating.

The fMRI and ERP literature converge by both reporting neural responses that associate with conflict monitoring (e.g., N2 and ACC; Botvinick et al., 2011; Braver et al., 2001) and response inhibition (P3b, frontoparietal regions; Poudel et al., 2017; Smith et al., 2008). Importantly, the fMRI and ERP research summarized above has been exclusively conducted within the younger adult population. It is unknown whether older adults would show similar ERP modulations or whether there are important age-related differences in the neural correlates that underlie performance during belief updating.

Age differences in ERPs. Previous research is mixed regarding age differences in LRP amplitude, with some studies showing smaller LRP amplitudes in older compared with younger adults (Sterr & Dean, 2008), and others showing larger LRP amplitudes with age (Roggeveen, Prime, & Ward, 2007). Importantly however, most of these studies employed RT tasks such as visual motion processing or modified versions of the Go/No-Go task that require an immediate response following stimulus presentation (Roggeveen, Prime, & Ward, 2007; Vallesi & Stuss, 2010). To my knowledge, no study has investigated the LRP in older adults during a belief

updating task. In light of this, one of the goals of Experiment 2 was to elucidate whether there are age differences in LRP amplitude following the onset of prior information, and to examine whether LRP amplitudes correlate with conservatism errors.

Within the conflict monitoring literature, studies examining age-related differences in ERP components are limited to paradigms that engage conflict using perceptual stimuli such as the Go/No-Go task (e.g., Go: X's; No-Go: O's; Vallesi, Stuss, McIntosh, & Picton, 2009), the Stroop task (congruent: the word BLUE coloured in blue; incongruent: the word BLUE coloured in purple; MacLeod, 1991), and the Flanker task (congruent: middle arrow pointing in the same direction as peripheral arrows; incongruent: middle arrow pointing in the opposite direction as peripheral arrows; Eriksen & Eriksen, 1974). Typically, these paradigms do not evoke age-related differences in accuracy, however older adults tend to respond slower than younger adults.

Concerning age differences in the N2 component, older adults show a smaller amplitude than younger adults (Bokura, Yamaguchi, Matsubara, & Kobayashi, 2002; Hämmerer, Müller, & Lindenberger, 2010; Hsieh, Wu, & Tang, 2016; Korsch, Frühholz, & Herrmann, 2016; Lucci, Berchicci, Spinelli, Taddei, & Di Russo, 2013). Lucci et al. (2013) employed a Go/No-Go paradigm and observed no age differences in accuracy but showed that amongst older adults the frontal No-Go N2 was nearly undetectable compared with younger adults. However, within the same time window, scalp topographies showed significant activity in posterior and prefrontal regions, suggesting that frontal involvement during response inhibition becomes deficient with age and other brain regions are required to compensate for this degradation in order to reach the accuracy levels of younger adults. In light of this, another goal of Experiment 2 was to examine whether there are age differences in the N2 amplitude following the onset of new information in

the belief updating task, and whether N2 amplitude correlates with avoiding the representativeness error.

In addition, most studies report smaller P3b amplitudes on conflict trials in older compared with younger adults (Bokura et al., 2002; Falkenstein, Hoorman, & Hohnsbein, 2002; Korsch et al., 2016; Lucci et al., 2013; Schmeidt-Fehr, & Basar-Eroglu, 2011), however some studies have reported larger P3b amplitudes in older than younger adults (Hämmerer et al., 2010; Vallesi, 2011), and other studies report no age-related changes in P3b amplitude (Enriquez-Geppert et al., 2010; Fjell, & Walhovd, 2001). Based on the studies reviewed, it is unclear whether the P3b is affected by age. Given that the P3b relates to response inhibition and attentional control (Smith et al., 2008), a component of executive control that declines with age (Hasher & Zacks, 1988), it was of interest to examine whether there are age differences in the amplitude of the P3b component following the onset of new information in conflict situations, and whether P3b amplitude correlates with decision accuracy.

In line with research showing the PASA effect (Lucci et al., 2013; Vallesi, McIntosh, & Stuss, 2011), it was also of interest to examine prefrontal electrode sites in the same time window during which the N2-P3 was evaluated to examine whether older adults recruit the PFC more than younger adults when updating beliefs.

In contrast to Achtziger et al., (2014) who showed that decision accuracy increased over time without feedback in younger adults, Experiment 1 of this dissertation showed no increase in decision accuracy over time without feedback in younger and older adults. Although these two studies report conflicting results, at present no study has examined whether belief updating is supported by feedback. It was of interest in the current experiment to investigate whether

accuracy would increase in the presence of feedback, and whether age-related differences would emerge under these circumstances.

Reinforcement learning theory posits that actions followed by positive outcomes are more likely to be repeated in the future, whereas actions followed by negative outcomes are less likely to reoccur (Sutton, & Barto, 1998). Dopamine plays a critical role in reinforcement learning. Dopaminergic neurons from the ventral tegmental area signal how much an outcome deviates from a predicted outcome during learning (Schultz et al., 1997). Learning occurs when an outcome is better than predicted (positive prediction error), reflected in a phasic increase in dopamine neuron activity, whereas an outcome that is worse than expected (negative prediction error) leads to the extinction of the learned behaviour, reflected in a phasic decrease of activity in these neurons (Eppinger et al., 2008). Mesencephalic dopamine neurons are important for the function of the ACC, a brain region involved in conflict monitoring (Botvinick et al., 2001), error (Scheffers, & Coles, 2000), and reward processing (Matsumoto, Suzuki, & Tanaka, 2003), that contain neurons that respond to expected and unexpected rewards (i.e., deviations from predicted outcomes; Ito, Stuphorn, Brown, & Schall, 2003).

Aging is associated with a marked decline in mesencephalic dopamine neurons, which may contribute to the impairments in cognitive control observed in old age, such as conflict monitoring and error processing (Braver & Barch, 2002; Eppinger et al., 2008; Paus, 2001). Feedback-related negativity (FRN) is an ERP component maximal at central and fronto-central electrodes produced by the dorsal ACC (Hauser et al., 2014; Holroyd, & Krigolson, 2007) and is observed between 200 and 300 ms after the onset of feedback (Miltner, Braun, & Coles, 1997). The FRN typically shows a decrease with learning, due to the decreased value of feedback (Holroyd & Coles, 2002). An outcome worse than expected (i.e., an error) evokes a phasic

decrease in the mesencephalic dopaminergic system, and the FRN is generated when this decrease in dopamine disinhibits neurons in the ACC (Eppinger et al., 2008). Nieuwenhuis and colleagues (2002) suggest that older adults show deficits in processing errors and reinforcement learning due to attenuated signaling in the mesencephalic dopamine system (de Boer et al., 2017; Mathewson et al., 2005; West, 2004). Supporting this, Chowdhury and colleagues (2013) showed that dopaminergic drugs are able to restore prediction-error signals in older adults, boosting performance to levels comparable to younger adults. However, others attribute deficits in learning from feedback to an age-related asymmetry in the processing of feedback valence such that positive feedback is over-valued and negative feedback is under-valued¹¹ (i.e., evidence of a positivity bias in older adults; Eppinger et al., 2008; Mather & Carstensen, 2005).

In light of the age differences observed in reinforcement learning, a goal of Experiment 2 was to investigate the potential age differences in decision accuracy during belief updating with feedback. I also sought to investigate age differences in feedback-related neural activity for positive and negative feedback to determine whether older adults are deficient at processing feedback in general regardless of valence (Mathewson et al., 2005), or whether older adults show a deficiency in processing negative feedback specifically (Mather & Carstensen, 2005).

Cognitive Reflection Test

In addition to the assessments described in the General Methods section (pages 56-61), the extended version of the Cognitive Reflection Test (Toplak et al., 2014a) was also used in Experiment 2. The extended Cognitive Reflection Test consists of a set of seven math problems that lure misleading but intuitive answers (Frederick, 2005; Toplak et al., 2014a). As the name

¹¹ However, some studies show no age differences in the FRN between positive and negative feedback (see Di Rosa et al., 2017; Pietschmann et al., 2011).

suggests, correct answers are thought to require reflective thought in order to overcome an initial intuitive response. The Cognitive Reflection Test correlates with a wide variety of tasks within the heuristics and biases literature, as these tasks also require overcoming of an intuitive response (Baron et al., 2015), and has been shown to be a stronger predictor of rational thinking relative to measures of intelligence and executive functioning (Toplak et al., 2011). The inclusion of the Cognitive Reflection Test in Experiment 2 was motivated by the literature showing that cognitive reflection underlies susceptibility to heuristics and decision errors (Baron et al., 2015). Specifically, the Cognitive Reflection Test served to index individual differences in cognitive reflection, with scores expected to predict performance on the belief updating task.

Altogether, the literature suggests that the neural processes that underlie belief updating may change with age, however this has never been empirically examined. In addition, there is no research on whether age differences emerge during belief updating with feedback or whether individual differences in numeracy, thinking dispositions or cognitive reflection underlie performance on a belief updating task. To fill this gap in the literature, the purpose of Experiment 2 was to investigate whether there are age differences in belief updating with feedback, and whether decision accuracy or the neural correlates that underlie belief updating are modulated by age or individual differences.

Hypotheses

Behavioural data. Age differences in accuracy were expected in light of research showing that younger adults are able to learn from feedback more successfully than older adults (Eppinger et al., 2008; Fernandes et al., 2018; Nieuwenhuis et al., 2002). I predicted an Age by Block by Decision Situation interaction such that younger adults would show an increase in accuracy across blocks, and older adults would show no increase in accuracy across blocks.

Further, I hypothesized that the representativeness error would be committed more than the conservatism error, regardless of age or block. Similar to Experiment 1, RTs were expected to decrease across blocks as the task becomes more familiar, but with older adults producing longer RTs compared with younger adults overall.

Additionally, higher scores on the Cognitive Reflection Test - Reflection subscale, Lipkus, BNT, Need for Cognition, DSST and lower scores on the Cognitive Reflection Test - Intuition subscale and Faith in Intuition assessments were hypothesized to predict higher decision accuracy. Conversely, lower scores on the Cognitive Reflection Test - Reflection subscale, Lipkus, BNT, Need for Cognition, DSST and higher scores on the Cognitive Reflection Test - Intuition subscale, and Faith in Intuition assessments were hypothesized to predict lower decision accuracy. Although the Lipkus, Need for Cognition and Faith in Intuition assessments were nonsignificant predictors of decision accuracy in Experiment 1, they were hypothesized to predict decision accuracy in Experiment 2 given that the belief updating task included feedback, affording the opportunity to learn from past decision errors – with learning potentially related to individual differences in these assessments.

ERP data. The ERP technique provided the opportunity to investigate neural activity evoked by the prior and sample before a behavioural response, and neural activity evoked by feedback after a behavioural response. Analysis of ERPs elicited by prior information provided insight into what extent participants were biased toward an option before new information was presented; ERPs elicited by new information elucidated the extent to which conflict detection or inhibitory control mechanisms were engaged before a decision was made; and ERPs evoked following feedback provided insight into the extent to which participants were sensitive to

positive and negative feedback. In light of this, the ERP effects described below were organized according to whether they were time-locked to the prior, new information, or feedback.

ERPs time-locked to prior information. Despite mixed findings concerning age differences in the LRP, I hypothesized that older adults would show larger LRP amplitudes compared with younger adults in the CconfB situation given the research that shows an increase in automatic processing with age (Peters et al., 2007). I also predicted that LRP amplitude would attenuate over time across both age groups. In line with Achtziger et al. (2014), greater LRP amplitudes were also predicted to associate with committing the conservatism error.

ERPs time-locked to sample information. Given that older adults show smaller N2 components than younger adults (Hsieh et al., 2016; Lucci et al., 2013) and research reports mixed results concerning how age affects the P3b (Fjell, & Walhovd, 2001; Korsch et al., 2016; Vallesi, 2011), it was hypothesized that older adults would evoke smaller N2 and P3b amplitudes compared with younger adults in conflict situations. In line with results reported by Achtziger et al. (2014), decision accuracy in conflict situations was predicted to correlate with the amplitude of the N2 and P3b, with avoidance of the representativeness and conservatism errors correlating with larger N2 and P3b amplitudes and committing the representativeness and conservatism errors correlating with smaller N2 and P3b amplitudes. Further, in light of literature showing the PASA effect (Davis et al., 2008; Lucci et al. 2013), it was predicted that older adults would show more prefrontal positivity (pP) in the same time window as the N2-P3 compared with younger adults, which would provide evidence of an age-related compensatory mechanism used to make decisions in the belief updating task. It was also hypothesized that the N2, P3b, and pP components would attenuate over time due to familiarization of the task.

ERPs time-locked to feedback. In line with reinforcement learning theory, it was hypothesized that feedback-related activity in response to positive and negative feedback would decrease from the first to the second block of the task suggesting that participants learned to make accurate decisions and relied less on feedback over time (Holroyd & Coles, 2002). It was also hypothesized that younger adults would elicit a more pronounced FRN than older adults due to a deficit in error processing with age (Mathewson, et al., 2005; Nieuwenhuis et al., 2002; see Appendix XV for the MATLAB script used for signal processing and ICA).

See Table 9 for a summary of the hypotheses for Experiment 2.

Table 9. *Hypotheses for Experiment 2*

Behavioural Hypotheses		
	1	Younger adults were expected to show higher decision accuracy than older adults across decision situations.
	2	Younger, but not older adults, were expected to show an increase in decision accuracy across blocks.
	3	The RconfB situation was expected to produce more decision errors than the CconfB situation.
	4	Older adults were expected to show longer RTs than younger adults overall.
	5	RTs were expected to decrease across blocks regardless of decision situation.
Assessments	6	Higher (lower) scores on the CRT-R, Lipkus, BNT, NFC, DSST and lower (higher) scores on the CRT-I, and FI scales were expected to associate with higher (lower) decision accuracy.
ERP and Behavioural Hypotheses		
LRP	7	Older adults were expected to associate with larger LRP than younger adults in the CconfB situation, with LRP amplitude attenuating across blocks regardless of age, and correlating with decision accuracy in the CconfB situation.
N2, P3b	8	Older adults were expected to associate with smaller N2 and P3b amplitudes than younger adults in conflict situations, with amplitudes of both components attenuating across blocks, regardless of age. Also, N2 and P3b amplitude were expected to correlate with decision accuracy in conflict situations.
pP	9	Older adults were expected to associate with greater pP compared with younger adults, with pP expected to decrease across blocks for older but not younger adults.
FRN	10	Younger adults were expected to associate with a more pronounced FRN than older adults, with FRN amplitude expected to attenuate over time regardless of age group.
Assessments	11	Higher (lower) scores on the CRT-R, Lipkus, BNT, NFC, DSST and lower (higher) scores on the CRT-I, and FI scales were expected to associate with larger (smaller) N2, P3b, and FRN amplitudes, and smaller (larger) LRP amplitudes.

Note. RconfB = Representativeness heuristic conflicts with Bayes' theorem; CconfB = Conservatism heuristic conflicts with Bayes' theorem; RT = reaction time; CRT-R = Cognitive Reflection Test – Reflection subscale; CRT-I = Cognitive Reflection Test – Intuition subscale; BNT = Berlin Numeracy Test; Lipkus = Lipkus Numeracy Task; Digit Symbol Substitution Test;

NFC = Need for Cognition; FI = Faith in Intuition; LRP = lateralized readiness potential; pP = prefrontal positivity; FRN = feedback-related negativity.

Method

Participants

The final sample included 25 younger adults (ages 18-34) and 25 older adults (ages 65-87). All participants completed the battery of assessments described in the General Methods section (pages 56-61), with the addition of the extended Cognitive Reflection Test (Toplak et al., 2014a). In line with previous literature, as well as Experiment 1, several age differences were observed (see Table 10). Older adults were more educated, had higher vocabulary scores, and lower self-reported symptoms of anxiety and negative affect, as well as slower processing speed as indexed by the DSST, compared with younger adults.

Table 10. *Characteristics of the Final Sample in Experiment 2*

		Younger Adults (n = 25)	Older Adults (n = 25)		
		<i>M (SD)</i>	<i>M (SD)</i>	<i>p</i>	<i>d</i>
Demographics					
	Age in Years	22.92 (4.10)	74.28 (5.92)	<.001	-10.11
	Age Range	18-34	65-87	--	--
	Male/Female Ratio	13/12	12/13	.50	--
	Education in Years	14.96 (1.99)	18.00 (2.92)	<.001	-1.22
	Right/Left Handed	20/5	21/4	.50	--
Emotion					
Assessments	Depression ^a	5.04 (5.42)	4.40 (5.13)	.67	0.12
	Anxiety ^a	4.88 (6.30)	2.08 (3.34)	.06	0.56
	Stress ^a	7.20 (7.30)	6.88 (4.69)	.85	0.05
	Positive Affect ^b	30.16 (9.06)	33.20 (7.23)	.19	0.37
	Negative Affect ^b	12.20 (2.71)	10.88 (1.72)	.04	0.58
Cognitive					
Assessments	MMSE ^c	29.40 (.816)	--	--	--
	DSST ^c	89.64 (16.13)	69.16 (11.56)	<.001	1.46
	Mill Hill ^c	17.6 (2.96)	24.72 (3.27)	<.001	-2.28
	BNT ^c	2.16 (1.46)	1.88 (1.33)	.48	0.20
	Lipkus Numeracy ^c	9.56 (2.12)	9.72 (1.99)	.78	0.08
	Faith in Intuition	6.09 (1.52)	5.75 (1.93)	.49	0.20
	Need for Cognition	53.88 (5.95)	53.56 (5.03)	.84	0.06
	CRT-R	2.44 (1.94)	1.80 (1.82)	.24	0.34
	CRT-I	3.28 (1.72)	3.68 (1.75)	.42	0.23

Note. Between-group comparisons were made using independent-sample *t* tests, aside from gender and handedness which were examined using Pearson's chi square. Standard deviations are in parentheses. ^aSubscales of the 21-item Depression Anxiety and Stress Scale; ^bSubscales of the Positive and Negative Affect Schedule; ^cScores reflect number of correct solutions; MMSE = Mini-Mental State Examination; DSST = Digit Symbol Substitution Test; Mill Hill = Mill Hill Vocabulary Test; BNT = Berlin Numeracy Test; Lipkus Numeracy = Lipkus Numeracy Task; CRT-R = Cognitive Reflection Test – Reflection; CRT-I = Cognitive Reflection Test – Intuition; *d* = Cohen's *d* for standardized effect size estimates of the mean difference between groups.

A total of four participants were excluded and later replaced. Two younger adults were excluded because they scored above the cut-off on the depression and anxiety subscales of the DASS-21 (see General Methods section pages 47-48 for a description of the inclusion and exclusion criteria). One older adult was excluded due to technical issues with the stimulus delivery program, and another older adult was excluded due to issues with EEG equipment.

EEG Recording and Processing

The EEG was recorded from 64 Ag/AgCl electrodes arranged in accordance with the International 10/20 configuration (see Figure 6), with data referenced to the posterior Common Mode Source (CMS) and Driven Right Leg (DRL) sites.

Figure 6. Schematic of the 64-channel International 10/20 system.

Horizontal electrooculograms (EOGs) were recorded by electrodes attached to the outer canthi of each eye, and vertical EOG was recorded via electrodes placed at the inferior orbits to measure horizontal and vertical eye movements respectively, as well as eye blinks. The EEG and EOG were digitized continuously at a 512 Hz sampling rate using a Biosemi ActiveTwo AD-Box (Bio-Semi; Wilmington, NC). Data were downsampled to 256 Hz and bandpass filtered

using a 0.1 Hz high-pass cut-off (12db/octave) to reduce low-frequency noise such as drifting, and a 30 Hz low-pass cut-off (24db/octave) to reduce high-frequency noise such as muscle tension. Following this step, data were then re-referenced to the average of the left and right mastoids, -100 ms baseline corrected, and segmented into 1000 ms epochs that were separately time-locked to the onset of the prior, sample, and feedback stimuli. Signal processing was performed using the EEGLAB/ERPLAB toolbox (Delorme & Makeig, 2004) via MATLAB software.

To correct for data contamination, artifacts were removed from the segmented epochs using independent components analysis (ICA). The purpose of ICA is to separate multivariate signals into additive, independent subcomponents (Luck, 2014). Specifically, the raw EEG and EOG data are fed to a neural network in which a learning algorithm generates a matrix of independent components that are equal to the total number of channels of EEG/EOG data and correspond to the time course of the original data, such that for each time point in the EEG/EOG data, each component has a magnitude that indicates the strength of the component at that time point. These components may reflect true neural components or reflect contaminated data (i.e., artifacts) such as eye blinks, eye movements, or muscle activity. Artifacts are removed by setting its contribution to zero and recalculating the voltage at each electrode site. Critically, ICA is a good method for rejecting artifacts due to the independence of the decomposed components. That is, rejected components are more likely to reflect artifacts rather than neural activity (see Luck 2014, online supplementary material for a review of ICA).

In light of the above, the artifact rejection procedure in the current experiment involved identifying and rejecting components that showed horizontal (e.g., saccades) or vertical eye movements (e.g., blinks) via visual inspection of the scalp topographies for each participant.

Once the identified artifacts were removed, the matrix of independent components was multiplied by the original matrix, which recovers the original artifact-corrected EEG/EOG data (Luck, 2014). The artifact-corrected data were then averaged across participants and mean amplitudes (μV) were extracted for analysis.

Materials

Stimuli. The stimuli and number of trials for the combinations of prior and sample information described in the General Methods section (pages 51-56) were also employed in the current experiment, with the addition of a feedback window that indicated whether the participant made a “correct” or “incorrect” decision on each trial. Feedback was presented in the center of the screen (i.e., within a space of 220×220 pixels) in 24-point Arial font against a grey (RGB value: 100-100-100) background, with the word “correct” in green font (RGB value: 0, 255, 0) and the word “incorrect” in red font (RGB value: 255, 0, 0).

Procedure

Belief updating task. The experimenter proceeded to describe the belief updating task using the same method as outlined in the General Methods section (pages 48-51), with the added explanation that feedback would be presented after every decision made. Subsequent to the EEG capping procedure, participants were seated in front of a PC computer inside a quiet room. The single trial sequence presenting urn distributions, prior, and sample information in the belief updating task was identical to that of Experiment 1, with the exception of a feedback window¹² that replaced the blank screen after a decision was made in Experiment 1. Including task instructions, ERP capping, the practice and experimental task, the belief updating task took

¹² Practice trials did not include feedback in order to avoid learning prior to beginning the experiment.

approximately 1.5 hours to complete.

Stimuli were presented in the sequence with the same presentation timing as outlined in the General Methods section (pages 51-53). Different from Experiment 1 however, following a response, a feedback window was presented for 1000 ms indicating whether the participant made a “correct” or incorrect” decision. Figure 7 provides a sample of a single-trial sequence for Experiment 2.

Figure 7. Sample trial of the representativeness heuristic conflicting with Bayes’ theorem in Experiment 2

Cognitive and affective assessments. After the task, the EEG cap was removed and the battery of assessments outlined in the General Methods section (pages 56-61) was administered, with the addition of the extended Cognitive Reflection Test.

Data Analysis

Behavioural data. Analysis of behavioural data used the same independent and dependent variables as Experiment 1 and followed the same analysis procedures outlined in the General Methods section (pages 61-64), with the addition of the data from the extended Cognitive Reflection Test. Although the standard way to score the Cognitive Reflection Test is to add correct responses, responses can also be organized into three categories – correct, “intuitive” incorrect and “other” incorrect. The “intuitive” incorrect category, which tends to make up most incorrect answers (see Campitelli & Gerrans, 2014; Frederick, 2005), represent errors made as if using a heuristic to make decisions. The “other” incorrect category represents errors that differ from the cued intuitive answer, such as a mathematical error or no answer. Therefore, in the current experiment, correct answers are referred to as Cognitive Reflection Test - Reflection, as this subscale describes the ability to detect and override an intuitive but incorrect response. In line with recent research (Brosnan, Hollinworth, Antoniadou, & Lewton, 2014; Piazza & Sousa, 2014; Pennycook et al., 2016; Shenhav, Rand, & Greene, 2012), “intuitive” incorrect answers are referred to as the Cognitive Reflection Test - Intuition subscale representing the inverse of the typical use of the Cognitive Reflection Test by using it to index the failure to detect and override an intuitive, incorrect response.

ERP data.

LRP. In line with prior research, to make inferences about conservatism the LRP was analyzed (Achtziger, et al., 2014; Eimer, 1998). Specifically, disproportionate priors (e.g., Urn A: 25%; Urn B: 75%) were compared independent of sample information. Mirroring Achtziger et al. (2014), stimulus-locked data were segmented into epochs 3100 ms before to 200 ms after presentation of the sample. A 100 ms time interval before prior information was presented was

used for baseline correction. Epochs for the three prior probabilities (25, 50, 75) were averaged separately producing three average waveforms per participant. In line with prior research, the LRPs were determined in two steps following a double-subtraction method (Achtziger et al., 2014; Eimer, 1998). In a first step, difference waveforms for electrode sites C3-C4 were computed for each situation with disproportionate priors (i.e., Urn A: 25%; Urn B: 75%, and vice versa). In a second step, LRPs were computed by subtracting the waveforms for the prior: Urn A: 25%; Urn B: 75%, from the waveform for the prior: Urn A: 75%; Urn B: 25%. The resulting LRP provides an index of conservatism such that positive values indicate stronger motor preparedness for the option of the urn with the higher prior probability (De Jong et al., 1990). Younger and older adult waveforms were averaged, respectively, to produce grand average waveforms for each age group. To quantify the LRP for each participant, the mean amplitude during the 100 ms time interval preceding presentation of the sample was calculated. This time window was chosen because it reflects participants' left-to-right bias for an option immediately preceding sample presentation and was used in previous literature employing the same paradigm (Achtziger et al., 2014).

N2, pP, and P3b. Analysis of the N2 and pP was modeled after previous literature (Achtziger et al., 2014; Lucci et al., 2013), as well as guided by visual inspection of the data. Stimulus-locked data were segmented into epochs from 100 ms before to 1000 ms after sample onset, with 100 ms used for baseline correction. In line with previous studies (Achtziger et al., 2014; Bartholow et al., 2005), only trials with correct responses were analyzed to observe conflict monitoring. Epochs locked to the RconfB situation and CconfB situation were averaged separately, producing two average waveforms per participant. The difference waveform was computed by subtracting RconfB situation waveforms from CconfB situation waveforms. Grand

averages were derived by averaging these ERPs across younger and older adults.

To quantify the N2 in the averaged ERP waveforms for each participant within each age group, the mean amplitude in the 50 ms interval centered around the peak latency between 220-270 ms after sample onset was calculated. This time window was chosen based on visual inspection of the data, previous research that has found the N2 to be maximal within this interval (i.e., typically between 200-400 ms; Bartholow & Dickter, 2008; Nieuwenhuis et al., 2002), and Achtziger et al. (2014) used a comparable time window (i.e., 235-285 ms) to capture the N2 in a similar paradigm. Similar to previous research, the N2 mean amplitude was evaluated by averaging frontocentral channels (i.e., FCz and Fz), where it has been found to be maximal in both younger and older adults (Achtziger et al., 2014; Bartholow, et al., 2005; Bartholow, Riordan, Sauls, & Lust, 2009; Lucci et al., 2003; Nieuwenhuis et al., 2003).

Centroparietal sites (i.e., Cz, Pz) were averaged in order to quantify the P3b in the averaged ERP waveforms for each participant within each age group and block for trials with correct responses only. Following previous aging literature, peak amplitudes were determined between 300-600 ms (Falkenstein et al., 1999; Hsieh et al., 2016) following sample presentation. Mean amplitudes were then obtained using 100 ms windows centered around the peak latencies (405-505 ms for younger adults; 425-525 ms for older adults). Mean amplitudes for each age group and block were then analyzed.

Scalp topographies also showed prefrontal positivity (pP) amongst older adults in the same time windows in which the N2 (200-400 ms) and P3b (300-600 ms) are typically observed. Therefore, to examine pP in addition to the N2 and P3b, mean amplitudes from averaged anterior recordings (i.e., AF3, AFz, AF4) were also analyzed within a 50 ms epoch centered around the peak latency (i.e., 375-425 ms) following the onset of the sample in order to quantify pP in the

averaged ERP waveforms for each participant within each age group and block for trials with correct responses only. The epochs and electrode sites selected for analysis of pP in younger and older adults were based on prior research (Lucci et al., 2013), as well as visual inspection of the data.

FRN. Analysis of the FRN was modeled after previous literature (Eppinger et al., 2008; Forder & Dyson, 2016), as well as visual inspection of the data. Epochs were defined relative to the onset of the feedback, were baseline corrected according to a 100 ms pre-stimulus interval, with neural activity examined 1000 ms post-feedback. The FRN was measured as the mean amplitude within a 50 ms window centering the peak latency of the FRN at the electrode Cz for younger (between 240-290 ms) and older adults (between 325-375 ms) for correct and incorrect responses separately.

Of note, peak latencies for each ERP component (except the LRP) were analyzed using mixed ANOVAs. However, the latency results did not contribute to the main results and therefore were placed in the appendix (see Appendix XVI for latency results).

Brain-behaviour correlations. If between- or within-subject differences were observed in ERP activity (e.g., age or block differences), correlational analyses were used to investigate the association between ERP activity and decision accuracy in the belief updating task, and between ERP activity and individual differences measured by the assessments included in the experiment.

Results

Decision Accuracy for All Decision Situations

Decision accuracy for each of the 15 prior and sample combinations are described in Table 11 for younger and older adults. As described in the General Methods section (pages 53-55), the six bolded situations are of interest in the current dissertation. Independent sample *t* tests

were conducted across counterbalancing versions on decision accuracy for each decision situation of interest. No differences in counterbalancing version were observed in the decision situations or conditions ($ps > .05$).

Table 11. *Means and Standard Deviations for Accuracy in Experiment 2*

Prior	Sample				
	0	1	2	3	4
75	98.66 (4.62)	92.47 (10.15)	84.67 (13.61)	99.33 (3.34)	99.33 (3.34)
	94.00 (15.13)	87.20 (10.97)	74.27 (19.37)	94.66 (9.29)	96.66 (6.82)
50	8	60	60	6	6
	99.33 (3.34)	97.33 (7.88)	96.20 (7.49)	97.26 (4.41)	98.00 (5.54)
25	98.00 (5.54)	98.00 (7.33)	88.46 (16.15)	92.33 (13.86)	95.33 (10.23)
	8	8	60	60	8
25	96.00 (8.72)	98.00 (7.33)	96.00 (11.04)	74.13 (21.92)	87.66 (11.43)
	96.66 (6.82)	95.33 (14.85)	95.34 (13.19)	63.81 (19.44)	80.73 (16.88)
	6	8	6	60	60

Note. Values represent proportion (%) of correct decisions; Prior = prior probability for the left urn; Sample = number of majority colour balls of sample; Younger adults' accuracy is on the top line, older adults' accuracy is on the bottom line. Total number of trials are indicated on the third row. Standard deviations are in parentheses. Bold indicates decision situations of interest.

Combining like-situations. To examine whether there were differences in decision accuracy across like-decision situations, paired-samples t tests were conducted across age group. Results showed no difference in decision accuracy within RalignB situations $t(50) = -1.27, p = .21, d = -0.18$. However, differences in decision accuracy were observed in the RconfB situations, $t(50) = -4.84, p < .001, d = -0.68$, and CconfB situations, $t(50) = -2.34, p = .02, d = -0.33$. Following Achtziger and colleagues (2014), like-situations were combined regardless of these differences as it was important to maximize the total number of trials for each like-situation to increase the signal-to-noise ratio for ERP analysis.

Decision Accuracy

Concerning whether participants' decision accuracy increased over time, trials were first segmented into two blocks (i.e., 207 trials in the first and second block). Then, a 2 (age: young,

old) $\times$ 3 (decision situation: RconfB, CconfB, RalignB) $\times$ 2 (trial block: Block 1, Block 2) mixed ANOVA on decision accuracy was conducted and results showed a main effect of decision situation, $F(1.68, 80.87) = 50.29, p < .001, \eta_p^2 = .51$. Pairwise comparisons showed that the RconfB situation ($M = .74, SD = .18$) promoted less accurate decisions than the CconfB ($M = .87, SD = .09$), $t(49) = -5.92, p < .001, d = -0.84$, and RalignB situations ($M = .94, SD = .09$), $t(49) = -8.99, p < .001, d = -1.28$, and CconfB situations promoted less accurate decisions than RalignB situations, $t(49) = -4.49, p < .001, d = -0.64$. A main effect of trial block was also observed, $F(1, 48) = 66.74, p < .001, \eta_p^2 = .58$, with less accurate decisions made in the first block ($M = .81, SD = .11$) compared with the second block ($M = .89, SD = .10$). Additionally, a main effect of age was observed, $F(1, 48) = 8.00, p = .01, \eta_p^2 = .14$, with younger adults ($M = .89, SD = .09$) making more accurate decisions than older adults ($M = .81, SD = .10$).

Finally, a Decision Situation by Trial Block interaction was observed, $F(1.70, 81.80) = 21.22, p < .001, \eta_p^2 = .31$. To probe the interaction, paired-samples t tests and repeated-measures ANOVAs with pairwise comparisons were conducted to examine potential block effects within each decision situation, as well as potential differences in accuracy for each decision situation within each block. Specifically, accuracy improved from the first ($M = .67, SD = .19$) to the second block ($M = .81, SD = .19$), $t(49) = -8.41, p < .001, d = -1.2$, in the RconfB situation. Similarly, within the CconfB situation, accuracy improved from the first ($M = .84, SD = .12$) to the second block ($M = .90, SD = .09$), $t(49) = -4.04, p < .001, d = -0.57$. However, accuracy did not improve from the first ($M = .93, SD = .10$) to the second block ($M = .95, SD = .11$), $t(49) = -1.82, p = .07, d = -0.26$, in the RalignB situation. Within the first trial block, decision situation was significant, $F(1.70, 83.48) = 75.14, p < .001, \eta_p^2 = .61$. The RconfB situation resulted in less accurate decisions than the CconfB situation, $t(49) = -7.29, p < .001, d = -1.03$, and the RalignB

situation, $t(49) = -11.03, p < .001, d = -1.56$, and the CconfB situation resulted in less accurate decisions than the RalignB situation, $t(49) = -5.16, p < .001, d = -0.73$. Within the second trial block, decision situation was significant, $F(1.66, 81.55) = 17.67, p < .001, \eta_p^2 = .27$. The RconfB situation resulted in less accurate decisions than the CconfB situation, $t(49) = -3.49, p = .003, d = -0.50$, and the RalignB situation, $t(49) = -5.16, p < .001, d = -0.67$. The CconfB situation resulted in less accurate decisions than the RalignB situation, $t(49) = -2.85, p = .02, d = -0.41$. No Decision Situation by Age interaction, $F(1.68, 80.83) = .75, p = .10, \eta_p^2 = .02$, and no Block by Age interaction was observed, $F(1, 48) = .58, p = .45, \eta_p^2 = .01$. See Figure 8 for an illustration of these effects.

Figure 8. Accuracy proportion for each decision situation across trial block in Experiment 2.

Decision RT

In addition to examining decision accuracy over time, RT across blocks was also examined. A 2 (age: young, old) $\times$ 3 (decision situation: RconfB, CconfB, RalignB) $\times$ 2 (trial block: Block 1, Block 2) mixed ANOVA on RT was conducted. Results showed a main effect of decision situation, $F(1.58, 75.62) = 52.73, p < .001, \eta_p^2 = .52$. Planned comparisons showed that RconfB situations ($M = 2425.67, SD = 1067.70$) produced longer RTs than the CconfB ($M = 2262.93, SD = 1101.50$), $t(49) = 3.18, p = .01, d = 0.45$, and RalignB situations ($M = 1639.77, SD = 612.58$), $t(49) = 8.84, p < .001, d = 1.25$, with CconfB situations producing longer RTs than the RalignB situation, $t(49) = 6.62, p < .001, d = 0.95$. Additionally, a main effect of age was observed, $F(1, 48) = 6.49, p = .01, \eta_p^2 = .12$, such that younger adults ($M = 1748.25, SD = 842.00$) produced faster RTs than older adults ($M = 2346.79, SD = 818.66$). A main effect of trial block was also observed, $F(1, 48) = 8.34, p = .01, \eta_p^2 = .15$, such that the first block ($M = 2133.51, SD = 968.82$) produced longer RTs than the second block ($M = 1961.53, SD = 826.38$).

Finally, an Age by Decision Situation by Block three-way interaction was observed, $F(1.65, 79.15) = 13.30, p < .001, \eta_p^2 = .22$. In order to unpack this interaction, stratifying by age, paired-samples t tests and repeated-measures ANOVAs with pairwise comparisons were conducted to examine potential block effects within each decision situation on RT, as well as potential differences in RT for each decision situation within each block.

Within the younger adult group, RTs were slower in the first block ($M = 2448.74, SD = 1190.48$) compared with the second block ($M = 1930.08, SD = 1085.12$) in the RconfB situation, $t(24) = 4.28, p < .001, d = 0.86$. There was no difference in RT between the first ($M = 1486.56, SD = 995.68$) and second block ($M = 1720.99, SD = 860.35$) in the CconfB situation, $t(24) = -1.99, p = .06, d = 0.28$. In the RalignB situation, RTs were slower in the first ($M = 1603.58, SD =$

717.58) compared with the second block ($M = 1272.56$, $SD = 513.38$), $t(24) = 4.83$, $p < .001$, $d = 0.96$. Further, within the first block, decision situation was significant, $F(2, 48) = 39.59$, $p < .001$, $\eta_p^2 = .62$. The RconfB situation resulted in longer RTs compared with the CconfB situation, $t(24) = 8.16$, $p < .001$, $d = 1.63$, and the RalignB situation, $t(24) = 6.64$, $p < .001$, $d = 1.33$. However, the CconfB situation did not differ in RT compared with the RalignB situation, $t(24) = -1.33$, $p = .66$, $d = -0.19$. Similar to Block 1, decision situation was significant in Block 2, $F(2, 48) = 19.43$, $p < .001$, $\eta_p^2 = .45$. Specifically, the RconfB situation did not differ from the CconfB situation, $t(24) = 2.12$, $p = .23$, $d = 0.30$, however the RconfB situation showed longer RTs than the RalignB situation, $t(24) = 5.28$, $p < .001$, $d = 1.06$. The CconfB situation also produced longer RTs than the RalignB situation, $t(24) = 4.58$, $p < .001$, $d = 0.92$.

Within the older adults, RT did not differ between the first ($M = 2713.80$, $SD = 1137.06$) and second block ($M = 2610.07$, $SD = 912.19$) of the RconfB situation, $t(24) = 1.09$, $p = .16$, $d = 0.15$, the first ($M = 2637.06$, $SD = 1239.83$) or second block ($M = 2463.85$, $SD = 908.98$) of the CconfB situation, $t(24) = 1.16$, $p = .35$, $d = 0.16$, nor the first ($M = 1884.30$, $SD = 627.83$) or second block ($M = 1771.64$, $SD = 591.78$) of the RalignB situation, $t(24) = 1.51$, $p = .98$, $d = 0.21$. Within the first block, decision situation was significant, $F(1.51, 36.14) = 19.37$, $p < .001$, $\eta_p^2 = .45$. The RconfB and CconfB situation did not differ in RT, $t(24) = 0.75$, $p = .65$, $d = 0.11$, however the RconfB situation produced longer RTs than the RalignB situation, $t(24) = 5.51$, $p < .001$, $d = 1.10$, and the CconfB situation produced longer RTs than the RalignB situation, $t(24) = 4.21$, $p < .001$, $d = 0.84$. Decision situation was also significant in Block 2, $F(2, 48) = 28.92$, $p < .001$, $\eta_p^2 = .54$. Specifically, no differences were observed in RT between the RconfB and CconfB situation, $t(24) = 1.64$, $p = .22$, $d = 0.23$. However, longer RTs were produced in the RconfB compared with the RalignB situation, $t(24) = 6.15$, $p < .001$, $d = 1.23$, and longer RTs

were produced in the CconfB compared with the RalignB situation, $t(24) = 5.39, p < .001, d = 1.08$. See Figure 9 for an illustration of these effects.

Figure 9. Reaction time for each decision situation across trial block in Experiment 2.

To summarize, the RconfB situation produced the lowest level of accuracy, the CconfB situation produced an intermediate level of accuracy, and the RalignB situation produced the highest level of accuracy, with younger adults making more accurate decisions than older adults overall, and with accuracy increasing across blocks, regardless of age or decision situation. Younger adults showed the longest RTs in the RconfB situation, intermediate RTs in the CconfB situation and the shortest RTs in the RalignB situations, with RT becoming faster from the first to the second block. In contrast, older adults did not show differences in RT across conflict situations or blocks.

Relationship Between Representativeness and Conservatism

Separate Pearson bivariate correlations were conducted in the first and second block to examine associations between accuracy in the RconfB and CconfB situations for each age group

(i.e., whether an increase in accuracy in one decision situation was associated with an increase in accuracy in the other decision situation). For younger adults, decision accuracy in the RconfB and CconfB situations were significantly correlated in the first ($r = .77, p < .001$) and second block ($r = .69, p < .001$). However, for older adults, decision accuracy was not correlated between the RconfB and CconfB situations in the first block ($r = .08, p = .70$), nor in the second block ($r = .25, p = .23$). These results show that as accuracy increases in one decision situation, accuracy also increases in the other decision situation for the younger adults but not the older adults (see Figure 10).

Figure 10. Association of accuracy in conflict situations in Experiment 2.

Self-Assessment

Chi-square tests were conducted on categorical items (i.e., items 1 and 2) and independent-samples t tests were conducted on scaled items (i.e., items 3-5) of the self-assessment for each age group. Results showed that younger adults ($M = 4.44$, $SD = 0.71$) were more likely to report that their decisions were accurate compared with older adults ($M = 3.48$, $SD = 0.87$), $t(48) = 4.26$, $p < .001$, $d = 1.2$, and older adults ($M = 2.56$, $SD = 0.92$) were more likely to report that they felt the task was difficult relative to younger adults ($M = 3.96$, $SD = 0.93$), $t(48) = 5.35$, $p < .001$, $d = 1.5$. No other age-related differences were observed (see Table 12).

Table 12. *Self-Assessment for Experiment 2*

Item	Younger Adults (n = 25)	Older Adults (n = 25)	<i>p</i> -value	<i>d</i>
Influence Decision			.86	--
Likelihood	2 (8%)	2 (8%)		
Sample	0 (0%)	2 (8%)		
Both	23 (92%)	21 (84%)		
Decision Strategy			.19	--
Consistent	11 (44%)	8 (32%)		
Switched	13 (52%)	16 (64%)		
Neither	1 (4%)	1 (4%)		
Accuracy			< .001	1.2
Very inaccurate	0 (0%)	0 (0%)		
Somewhat inaccurate	1 (4%)	6 (24%)		
Neither inaccurate nor accurate	0 (0%)	1 (4%)		
Somewhat accurate	11 (44%)	18 (72%)		
Very accurate	13 (52%)	0 (0%)		
Task Difficulty			< .001	1.5
Very difficult	0 (0%)	2 (8%)		
Somewhat difficult	2 (8%)	11 (44%)		
Neither difficult nor easy	5 (20%)	9 (36%)		
Somewhat easy	10 (40%)	2 (8%)		
Very easy	8 (32%)	1 (4%)		
Confidence with Numbers			.08	0.34
Not at all confident	0 (0%)	1 (4%)		
Somewhat confident	9 (36%)	15 (60%)		
Confident	8 (32%)	3 (12%)		
Very confident	4 (16%)	5 (20%)		
Extremely confident	4 (16%)	1 (4%)		

Individual Differences on Accuracy

Multi-level modeling was used to examine whether individual differences in numeracy, cognitive reflection, thinking disposition, processing speed, verbal intelligence, and years of education predict decision accuracy in the RconfB or CconfB situations. Only significant interactions that include the decision situation factor are further interpreted.

In a first step, the ICC was calculated and showed that approximately 9% of the variance in decision accuracy was due to clustering. In light of the ICC value suggesting at least a portion of the variability of accuracy was due to clustering, a multi-level modeling approach was taken to examine predictors of decision accuracy. Similar to Experiment 1 analysis, random-intercepts and random-slopes models on accuracy in RconfB and CconfB decision situations, with age included as a Level 2 predictor, were conducted to observe which model best fit the data. Based on the AIC, BIC, and deviance values, the random-slopes model improved the model fit over the random-intercepts model (see Table 13). In light of this, assessment scores were included in the random-slope models as predictor variables separately, along with age and decision situation on decision accuracy.

Table 13. *Multi-Level Model Comparison on Accuracy in Experiment 2*

	Parameters	Model	
		Random-intercepts	Random-slopes
Fixed effects			
	Intercept	2.54 (0.17)	2.58 (0.19)
	OR	12.74	13.16
	Younger adults	-0.74 (0.24)*	-0.72 (0.24)
	OR	0.48	0.49
	RconfB	-0.93 (0.05)*	-0.93 (0.14)
	OR	0.39	0.39
Fit statistics			
	deviance	10,510	10,329
	AIC	10,518	10,341
	BIC	10,548	10,386
	df	4	6
Model comparison			$\chi^2 = 181.14^*$

Note. Standard errors and deviations in fixed and random effects parentheses, respectively; RconfB = Representativeness heuristic conflicts with Bayes' theorem; Younger adults and RconfB were used as the reference group; OR = odds ratios; AIC = Akaike's Information Criterion; BIC = Bayesian Information Criterion; df = degrees of freedom; Subscript χ^2 denotes the two models being compared. * $p < .05$.

Berlin Numeracy Task. Results showed a significant Berlin Numeracy Task (BNT) by Decision Situation interaction ($\beta = 0.15$, $SE = 0.13$; $p = .02$), as well as a significant BNT by Age interaction ($\beta = -0.35$, $SE = 0.17$; $p = .01$). However, there was no BNT by Decision Situation by Age interaction ($\beta = 0.11$, $SE = 0.18$; $p = .54$). In line with Experiment 1, data were centered at the 25th percentile, the median, and the 75th percentile, and multi-level models were rerun on these data to probe this interaction. As presented in Table 14, beta estimates for the effect of decision situation increase as scores on the BNT increase. This suggests that the effect of decision situation is strongest among those with lower BNT scores. The odds ratios suggest that among those with low BNT scores, the odds of a correct response in the RconfB situation were 33% of the odds of a correct response in the CconfB situation. However, among those with higher BNT scores, the odds of a correct response in the RconfB situation were 45% of the odds of a correct response in the CconfB situation. These results suggest that the effect of decision situation becomes greater with lower scores on the BNT.

Table 14. *Berlin Numeracy Test as a Predictor of Accuracy in Experiment 2*

	BNT centered at 25 th percentile					BNT centered at median					BNT centered at 75 th percentile				
	β	SE	Z	<i>p</i>	OR	β	SE	Z	<i>p</i>	OR	β	SE	Z	<i>p</i>	OR
Intercept	2.00	0.21	9.51	-	7.41	2.51	0.17	14.48	-	12.29	3.02	0.21	14.16	-	20.39
BNT	0.51	0.12	4.17	<.001	1.66	0.51	0.12	4.18	<.001	1.66	0.51	0.12	4.18	<.001	1.66
RconfB	-1.11	0.22	-4.97	<.001	0.33	-0.96	0.19	-5.15	<.001	0.38	-0.81	0.23	-3.49	<.001	0.45
age	-0.28	0.29	-0.99	<.001	0.75	-0.64	0.24	-2.68	<.001	0.53	-0.99	0.30	-3.28	<.001	0.37
BNT:RconfB	0.15	0.15	1.16	.03	1.16	0.15	0.13	1.16	.02	1.16	0.15	0.13	1.16	.03	1.16
BNT:age	-0.35	-0.35	-2.03	.01	0.70	-0.35	0.17	-2.04	.14	0.70	-0.25	0.17	-2.04	.10	0.70
RconfB:age	-0.05	-0.05	-0.18	.85	0.95	0.06	0.25	0.23	.85	1.06	0.17	0.33	0.53	.85	1.19
BNT:RconfB:age	0.11	0.11	0.61	.54	1.12	0.11	0.19	0.61	.54	1.12	0.11	0.19	0.61	.54	1.12

Note. BNT = Berlin Numeracy Test; RconfB = Representativeness heuristic conflicts with Bayes' theorem (reference group); β = beta estimate; SE = standard error; Z = z-score; OR = odds ratio.

Lipkus Numeracy Task. Results showed a significant Lipkus Numeracy Task (Lipkus) by Decision Situation interaction ($\beta = 0.15$, $SE = 0.13$; $p = .02$), as well as a significant Lipkus by Age interaction ($\beta = -0.35$, $SE = 0.17$; $p = .01$). However, there was no Lipkus by Decision Situation by Age interaction ($\beta = 0.11$, $SE = 0.18$; $p = .06$). In line with analyses conducted on the BNT interaction, Lipkus data were centered at the 25th percentile, the median, and the 75th percentile, and multi-level models were rerun on these data to probe this interaction. As shown in Table 15, beta estimates for the effect of decision situation increase as scores on the Lipkus increase. This suggests that the effect of decision situation is strongest among those with lower Lipkus scores. The odds ratios suggest that among those with lower Lipkus scores, the odds of a correct response in the RconfB situation were 38% of the odds of a correct response in the CconfB situation. However, among those with higher Lipkus scores, the odds of a correct response in the RconfB situation were 40% of the odds of a correct response in the CconfB situation. These results suggest that the effect of decision situation becomes greater with lower scores on the Lipkus.

Table 15. *Lipkus Numeracy as a Predictor of Accuracy in Experiment 2*

	Lipkus centered at 25 th percentile					Lipkus centered at median					Lipkus centered at 75 th percentile				
	β	SE	Z	p	OR	β	SE	Z	p	OR	β	SE	Z	p	OR
Intercept	2.41	0.18	13.22	-	11.16	2.72	0.19	14.47	-	15.20	3.03	0.22	13.35	-	20.69
Lipkus	0.31	0.08	3.66	<.001	1.36	0.31	0.08	3.66	<.001	1.36	0.31	0.08	3.67	<.001	1.36
RconfB	-0.95	0.18	-5.21	<.001	0.38	-0.93	0.19	-4.87	<.001	0.39	-0.91	0.23	-3.93	<.001	0.40
age	-0.49	0.26	-1.94	<.001	0.61	-0.89	0.25	-3.51	<.001	0.41	-1.28	0.30	-4.24	<.001	0.28
Lipkus:RconfB	0.02	0.08	0.29	.02	1.02	0.02	0.08	0.29	.02	1.02	0.02	0.08	0.29	.02	1.02
Lipkus:age	-0.19	0.12	-3.21	.10	0.68	-0.19	0.12	-3.22	.10	0.68	-0.19	0.12	-3.22	.10	0.68
RconfB:age	-0.16	0.26	-0.64	.87	0.85	0.07	0.25	0.28	.87	1.07	0.31	0.30	1.00	.88	1.36
Lipkus:RconfB:age	0.23	0.12	1.94	.06	1.26	0.23	0.12	1.94	.06	1.26	0.23	0.12	1.94	.06	1.26

Note. Lipkus = Lipkus Numeracy Task; RconfB = Representativeness heuristic conflicts with Bayes' theorem (reference group); β = beta estimate; SE = standard error; Z = z-score; OR = odds ratio.

Cognitive Reflection Test – Reflection subscale. Results showed a significant Cognitive Reflection Test – Reflection (CRT-R) by Decision Situation interaction ($\beta = 0.08$, $SE = 0.09$; $p = .03$). However, there was no CRT-R by Age interaction ($\beta = -0.27$, $SE = 0.15$, $p = .13$), nor a CRT-R by Decision Situation by Age interaction ($\beta = 0.15$, $SE = 0.14$; $p = .28$). To probe the CRT-R by Decision Situation interaction, data were centered at the 25th percentile, the median, and the 75th percentile and multi-level models were rerun on these data. As presented in Table 16, beta estimates for the effect of decision situation increase as scores on the CRT-R increase. This suggests that the effect of decision situation is strongest among those with lower CRT-R scores. The odds ratios suggest that among those with lower CRT-R scores, the odds of a correct response in the RconfB situation were 36% of the odds of a correct response in the CconfB situation. However, among those scoring higher on the CRT-R, the odds of a correct response in the RconfB situation were 45% of the odds of a correct response in the CconfB situation. These results suggest that the effect of decision situation becomes greater with lower CRT-R scores.

Table 16. *Cognitive Reflection Test-Reflection as a Predictor of Accuracy in Experiment 2*

	CRT-R centered at 25 th percentile					CRT-R centered at median					CRT-R centered at 75 th percentile				
	β	SE	Z	p	OR	β	SE	Z	p	OR	β	SE	Z	p	OR
Intercept	2.38	0.24	9.88	-	10.77	2.38	0.24	9.88	-	10.77	2.78	0.25	10.90	-	16.12
CRT-R	0.13	0.10	1.32	.20	1.14	0.13	0.10	1.32	.20	1.14	0.13	0.10	1.32	.20	1.14
RconfB	-1.03	0.23	-4.55	<.001	0.36	-1.03	0.23	-4.55	<.001	0.36	-0.81	0.24	-3.31	<.001	0.45
age	-0.41	0.32	-1.29	.01	0.66	-0.41	0.32	-1.29	.01	0.66	-1.22	0.38	-3.15	.01	0.29
CRT-R:RconfB	0.07	0.09	0.78	.03	1.09	0.07	0.09	0.78	.03	1.09	0.08	0.09	0.78	.03	1.08
CRT-R:age	-0.27	0.14	-1.85	.13	0.76	-0.27	0.14	-1.85	.13	0.76	-0.27	0.14	-1.86	.13	0.76
RconfB:age	-0.08	0.29	-0.29	.76	0.92	-0.08	0.29	-0.29	.76	0.92	0.36	0.36	0.99	.76	1.44
CRT-R:RconfB:age	0.15	0.14	1.08	.28	1.16	0.15	0.14	1.08	.28	1.16	0.15	0.14	1.09	.28	1.16

Note. CRT-R = Cognitive Reflection Test – Reflection subscale; RconfB = Representativeness heuristic conflicts with Bayes’ theorem (reference group); β = beta estimate; SE = standard error; Z = z-score; OR = odds ratio

The Need for Cognition, Faith in Intuition, Mill Hill Vocabulary Test, Cognitive Reflection Test – Intuition subscale and years of education did not predict decision accuracy. In addition, the multi-level model with DSST as a predictor of decision accuracy failed to converge. However, to examine whether DSST scores were related to decision accuracy, correlational analyses were conducted. Results showed that scores on the DSST were positively correlated with decision accuracy in the RconfB situation, $r = .47, p = .001$, and the CconfB situation, $r = .45, p = .001$. Additionally, analyses were conducted for each age group separately. For younger adults, correlations between DSST scores and decision accuracy did not reach significance in the RconfB situation, $r = .37, p = .07$, or the CconfB situation, $r = .37, p = .07$. However, for older adults, DSST scores and decision accuracy were correlated in the RconfB situation, $r = .42, p = .04$, but were not correlated in the CconfB situation, $r = .33, p = .10$.

To summarize, the multi-level modeling results show that those who scored lower on the Berlin Numeracy Test, Lipkus Numeracy Task, and the Cognitive Reflection Test – Reflection subscale made less accurate decisions in the RconfB situation compared with the CconfB situation. Those who scored higher on the Berlin Numeracy Test, Lipkus Numeracy Task, and Cognitive Reflection Test – Reflection subscale were less affected by decision situation, scoring higher in both the RconfB and CconfB situations. These results suggest that the Berlin Numeracy Test, Lipkus Numeracy Task and the extended Cognitive Reflection Test were good predictors of individual differences in decision accuracy in the belief updating task. Last, DSST scores in older adults correlated with decision accuracy in the RconfB situation, with correlations between DSST scores and decision accuracy in conflict situations not reaching significance in younger adults.

ERP Data: Prior Information Effects

LRP. The LRP was used to indicate lateralization of ERP activity over motor cortices in order to examine bias toward attending to the higher prior probability in the CconfB situation (Achtziger et al., 2014; Leuthold et al., 1996; Steinhauser et al., 2009). It was hypothesized that committing the conservatism error would evoke a larger LRP, relative to not committing the conservatism error. Grand-average LRP waveforms were examined, with positive values reflecting attending to the prior with the higher probability. Results showed that mean LRP amplitudes for younger adults were not significantly different from zero in the first block ($M = 0.74 \mu\text{V}$, $SD = 2.27$), $t(24) = 1.64$, $p = .18$, $d = 0.23$, but were significantly different from zero in the second block ($M = 0.98 \mu\text{V}$, $SD = 2.10$), $t(24) = 2.34$, $p = .03$, $d = 0.47$. This suggests that younger adults were biased toward attending to the higher prior probability in the second but not the first block. Within the older adults, mean LRP amplitudes were significantly different from zero in both the first ($M = 1.4 \mu\text{V}$, $SD = 2.29$), $t(24) = 3.01$, $p = .01$, $d = 0.60$, and the second block ($M = 0.91 \mu\text{V}$, $SD = 1.85$), $t(24) = 2.46$, $p = .02$, $d = 0.49$, suggesting that older adults were biased toward attending to the higher prior probability throughout the belief updating task (see Figure 11).

To examine whether younger and older adults differed in mean LRP amplitude, a 2 (age: young, old) $\times$ 2 (trial block: Block 1, Block 2) repeated-measures ANOVA was conducted¹³. Results showed no main effect of block, $F(1, 48) = .12$, $p > .05$, $\eta_p^2 = .003$, no main effect of age, $F(1, 48) = .31$, $p > .05$, $\eta_p^2 = .01$, and no Block by Age interaction, $F(1, 48) = 1.16$, $p > .05$, $\eta_p^2 = .02$.

¹³ Of note, decision situation was not included as a factor in the ANOVA because LRP amplitude is only relevant to the CconfB situation.

Further, Pearson bivariate correlations were conducted to examine whether mean LRP amplitude associated with decision accuracy in CconfB situations for each age group. Younger adults showed no correlation between mean LRP amplitude and decision accuracy in CconfB situations in the first ($r = -.16, p = .36$) or second block ($r = .30, p = .58$). Similar results were observed for the older adults ($r = .22, p = .10$; $r = -.09, p = .12$, respectively).

Exploratory analysis showed that Berlin Numeracy Test scores correlated with mean LRP amplitude in the first block ($r = -.41, p = .04$), but not in the second block ($r = .03, p = .89$) for younger adults. No other significant correlations were observed.

To summarize, the results showed that younger adults were biased toward attending to the higher prior probability in the second half of the task, whereas older adults were biased toward attending to the higher prior probability throughout the task. Further, higher numeracy scores on the Berlin Numeracy Test were associated with smaller LRP amplitudes in the first block amongst younger adults.

Figure 11. LRP: Grand averaged ERPs from C3-C4, time-locked to urn distribution onset. Time is on the x-axis and bias-related amplitude are highlighted in a 100 ms epoch before presentation of the sample.

ERP Data: Sample Information Effects

N2, pP. In a first step, one-sample t tests were conducted to examine whether N2 and pP mean amplitudes of the difference wave for RconfB vs. CconfB situations were significantly different from zero in the first and second blocks for each age group. Results showed that for the younger adults, N2 mean amplitude was not significantly different from zero in the first block ($M = 0.00 \mu\text{V}$, $SD = 2.81$), $t(24) = 0.10$, $p = .26$, $d = 0.01$, but was significantly different from zero in the second block ($M = 1.28 \mu\text{V}$, $SD = 2.77$), $t(24) = 2.31$, $p = .03$, $d = 0.46$. pP mean amplitude was not significantly different from zero in the first ($M = -0.06 \mu\text{V}$, $SD = 4.57$), $t(24) = -0.07$, $p = .24$, $d = -0.01$, or the second block ($M = 0.84 \mu\text{V}$, $SD = 3.70$), $t(24) = 1.13$, $p = .35$, $d = 0.16$. Concerning older adults, N2 mean amplitude was not different from zero in the first ($M = -0.04 \mu\text{V}$, $SD = 2.56$), $t(24) = -0.07$, $p = .99$, $d = -0.01$, or second block ($M = 0.31 \mu\text{V}$, $SD = 2.79$), $t(24) = 0.55$, $p = .33$, $d = 0.08$. pP mean amplitude was not different from zero in the first block ($M = 0.45 \mu\text{V}$, $SD = 3.62$), $t(24) = 0.62$, $p = .49$, $d = 0.09$, but was significantly different from zero in the second block ($M = 0.95 \mu\text{V}$, $SD = 2.35$), $t(24) = 2.02$, $p = .05$, $d = 0.29$. These results suggest that the RconfB situation produced a larger N2 mean amplitude compared with the CconfB situation for younger adults in the second block, and the CconfB situation produced a larger pP mean amplitude for older adults compared with the RconfB situation in the second block.

The 2 (age: young, old) $\times 2$ (decision situation: RconfB, CconfB) $\times 2$ (trial block: Block 1, Block 2) $\times 2$ (region: N2, pP) mixed ANOVA on N2 and pP mean amplitude revealed a main effect of region, $F(1, 48) = 11.69$, $p = .001$, $\eta_p^2 = .19$, such that the N2 ($M = 4.27 \mu\text{V}$, $SD = 2.90$) showed a more positive amplitude than the pP ($M = 2.95 \mu\text{V}$, $SD = 4.61$). Further, a main effect of age was observed, $F(1, 48) = 8.70$, $p = .01$, $\eta_p^2 = .15$, such that older adults ($M = 4.99 \mu\text{V}$, SD

= 2.54) produced more positive amplitudes than younger adults ($M = 2.30 \mu\text{V}$, $SD = 3.76$).

Additionally, an Age by Region interaction was observed, $F(1, 48) = 21.61$, $p < .001$, $\eta_p^2 = .31$.

To probe this interaction, independent- and paired-samples t tests were conducted to examine potential differences in region within each age group, as well as the potential age differences for each region. For younger adults, the N2 ($M = 3.80 \mu\text{V}$, $SD = 3.07$) showed a more positive mean amplitude than pP ($M = 0.66 \mu\text{V}$, $SD = 4.94$), $t(24) = 4.73$, $p < .001$, $d = -0.95$. For older adults, no difference was observed between the N2 ($M = 4.75 \mu\text{V}$, $SD = 2.70$), and pP ($M = 5.25 \mu\text{V}$, $SD = 2.84$), $t(24) = -1.13$, $p = .27$, $d = -0.16$. Further, younger and older adults showed no difference in N2 amplitude, $t(24) = -1.15$, $p = .25$, $d = -0.16$, but did show a difference in pP amplitude, $t(24) = -4.03$, $p < .001$, $d = 1.14$, with older adults showing a more positive pP amplitude than younger adults.

Pearson bivariate correlations were conducted to examine potential associations between N2 and pP mean amplitudes and decision accuracy for each age group and block. For younger adults within the first block, N2 mean amplitude was significantly correlated with accuracy in the RconfB situation ($r = -.39$, $p = .05$), such that higher accuracy was associated with larger N2 amplitude. pP amplitude showed a correlation with accuracy in the RconfB situation in the first block ($r = -.41$, $p = .04$), such that higher accuracy was associated with smaller pP amplitude. However, within the second block, younger adults did not show a correlation between accuracy in the RconfB situation and N2 ($r = -.19$, $p = .34$) or pP amplitude ($r = -.17$, $p = .43$).

Additionally, for younger adults within the first block, N2 mean amplitude was significantly correlated with accuracy in the CconfB situation ($r = -.40$, $p = .04$), such that higher accuracy was associated with larger N2 amplitudes. pP amplitude also showed a correlation with accuracy in the CconfB situation ($r = -.51$, $p = .01$), such that higher accuracy was associated

with smaller pP amplitudes. However, within the second block, younger adults did not show a correlation between accuracy in the CconfB situation and N2 ($r = -.13, p = .54$) or pP amplitude ($r = -.06, p = .78$).

Concerning older adults within the first block, no correlation between accuracy in the RconfB situation and N2 ($r = -.05, p = .62$) or pP amplitude ($r = -.05, p = .88$) was observed. Similar results were observed in the second block ($r = -.05, p = .23$; $r = -.12, p = .93$, respectively). Further, within the first block, no correlation was observed between accuracy in the CconfB situation and N2 ($r = .01, p = .45$). No correlations were observed between pP amplitude and decision accuracy in the CconfB situation in the first ($r = -.03, p = .88$) or second block ($r = .13, p = .55$; Figures 12 and 13).

Exploratory correlations between the Lipkus Numeracy Test and N2 mean amplitude in the RconfB situation in the first block did not reach significance ($r = -.34, p = .09$), but a significant correlation was observed in the second block ($r = -.39, p = .05$) for younger adults. pP amplitude also correlated with the Lipkus Numeracy Test in the first block ($r = -.43, p = .03$) in the RconfB situation, but not the second block ($r = -.44, p = .19$). Scores on the Lipkus Numeracy Test were also correlated with N2 ($r = -.45, p = .02$), and pP ($r = -.41, p = .04$) amplitude in the CconfB situation in the first, but not in the second block ($r = -.27, p = .53$; $r = -.29, p = .64$, respectively). Last, younger adults showed a significant correlation between N2 mean amplitude and the Cognitive Reflection Test - Intuition subscale in the first ($r = -.43, p = .03$) and second block ($r = -.41, p = .04$) in the CconfB situation. However, the correlation between pP amplitude and the Cognitive Reflection Test - Intuition subscale did not reach significance ($r = -.35, p = .08$) in the second block. No other correlations were observed between N2 or pP amplitudes and assessment scores in younger adults.

Concerning older adults, N2 mean amplitude was correlated with Lipkus scores in the first block in the CconfB situation ($r = -.41, p = .04$), and N2 amplitude was correlated with Lipkus scores in the second block in the RconfB situation ($r = -.41, p = .04$). Similarly, pP amplitude was correlated with Lipkus scores in the first block in the RconfB situation ($r = .43, p = .03$), and the second block ($r = .49, p = .01$). pP amplitude was correlated with Lipkus scores in the first block in the CconfB situation ($r = .39, p = .05$), and the second block ($r = .54, p = .01$). Additionally, within the second block in the RconfB situation, older adults showed a significant correlation between N2 amplitude and the Cognitive Reflection Test -Reflection subscale ($r = .42, p = .04$) and the Cognitive Reflection Test - Intuition subscale ($r = -.45, p = .02$). A similar result was observed between pP amplitude and the Cognitive Reflection Test - Intuition subscale ($r = -.41, p = .04$) in the second block for the RconfB situation. No other significant correlations were observed.

In summary, a larger N2 was observed in the second block in the RconfB situation amongst younger adults, with larger N2 and smaller pP amplitudes correlating with decision accuracy in the RconfB and CconfB situations in the first block. For older adults, a larger pP was observed in the second block in the CconfB situation. Additionally, N2 amplitude was more positive than pP amplitude in younger adults demonstrating the low neural activity observed in anterior regions, with older adults showing greater pP compared with younger adults. Further, in younger adults, higher Lipkus Numeracy Test scores correlated with a) larger N2 amplitudes in the second block in the RconfB situation, b) smaller pP amplitudes in the first block in the RconfB situation, and c) larger N2 and smaller pP amplitudes in the first block in the CconfB situation. Lower Cognitive Reflection Test – Intuition subscale scores were correlated with larger N2 amplitudes in the first and second block in the CconfB situation. In older adults, higher

scores on the Lipkus Numeracy Task were correlated with a) larger N2 amplitudes in the first block in the CconfB situation, b) larger N2 amplitudes in the second block in the RconfB situation, and c) larger pP amplitudes in the first and second blocks in both the RconfB and CconfB situations. In addition, smaller N2 amplitudes were correlated with higher Cognitive Reflection Test – Reflection subscale scores, larger N2 and smaller pP amplitudes were correlated with higher Cognitive Reflection Test – Intuition subscale scores in the second block in the RconfB situation.

Figure 12. N2: Grand averaged ERPs from FCz and Fz, time-locked to sample onset. Time is on the x-axis and sample-related amplitude differences are highlighted in the 220-270 ms epoch.

Figure 13. pP: Grand averaged ERPs from AF3, AFz, and AF4, time-locked to sample onset. Time is on the x-axis and sample-related amplitude differences are highlighted in the 375-425 ms epoch.

P3b. In a first step, one-sample t tests were conducted to examine whether P3b mean amplitudes of the difference wave for RconfB vs CconfB situations were significantly different from zero in the first and second blocks. Concerning younger adults, results showed that P3b mean amplitude of the difference wave was not significantly different from zero for the first block ($M = 0.33 \mu\text{V}$, $SD = 3.93$), $t(24) = 0.41$, $p = .28$, $d = 0.06$, but was significantly different from zero in the second block ($M = 2.10 \mu\text{V}$, $SD = 3.61$), $t(24) = 2.10$, $p = .01$, $d = 0.58$. Concerning older adults, the difference wave in the first ($M = -0.29 \mu\text{V}$, $SD = 3.12$), $t(24) = -0.47$, $p = .29$, $d = -0.07$, or second block ($M = 0.47 \mu\text{V}$, $SD = 2.47$), $t(24) = 0.95$, $p = .38$, $d = 0.13$, was not different from zero, suggesting no difference in P3b mean amplitude between conflict situations in the first or second block.

The 2 (age: young, old) $\times 2$ (decision situation: RconfB, CconfB) $\times 2$ (trial block: Block 1, Block 2) mixed ANOVA on P3b mean amplitude showed no main effect of decision situation, $F(1, 48) = 0.92$, $p = .34$, $\eta_p^2 = .02$, no main effect of block, $F(1, 48) = 2.97$, $p = .09$, $\eta_p^2 = .06$, and no main effect of age $F(1, 48) = 0.84$, $p = .36$, $\eta_p^2 = .02$. Additionally, no Decision Situation by Age interaction was observed, $F(1, 48) = 1.95$, $p = .17$, $\eta_p^2 = .04$, and no Block by Age interaction was observed, $F(1, 48) = 2.21$, $p = .14$, $\eta_p^2 = .04$. However, a significant Block by Decision Situation interaction was observed, $F(1, 48) = 5.02$, $p = .03$, $\eta_p^2 = .09$. To probe this interaction, paired-sample t tests were conducted. Results showed a significant difference in P3b mean amplitude in the RconfB situation across blocks, such that a more positive P3b was observed in the first block ($M = 4.79 \mu\text{V}$, $SD = 4.28$) compared with the second block ($M = 3.84 \mu\text{V}$, $SD = 4.40$), $t(49) = 2.40$, $p = .02$, $d = 0.34$. No mean amplitude change was observed across blocks for the CconfB situation, $t(49) = -0.69$, $p = .67$, $d = -0.10$. In addition, no difference in P3b amplitude between decision situations was observed in the first block, $t(49) = -0.03$, $p = .32$,

$d = -0.00$. However, the CconfB situation elicited a more positive P3b mean amplitude ($M = 5.12 \mu V$, $SD = 4.99$) than the RconfB situation ($M = 3.84 \mu V$, $SD = 4.39$), $t(49) = -2.86$, $p = .01$, $d = -0.40$, in the second block, supporting the results from the one-sample t -test on the difference wave (see Figure 14).

It was also of interest to examine whether the N2 and P3b components were associated. Within the younger adults, N2 amplitude was significantly correlated with P3b amplitude in the first block ($r = .45$, $p = .02$) but did not reach significance in the second block ($r = .38$, $p = .06$) in the RconfB situation. N2 amplitude was also correlated with P3b amplitude in the first block ($r = .40$, $p = .04$) and second block ($r = .39$, $p = .05$) in the CconfB situation. Concerning older adults, N2 amplitude was correlated with P3b amplitude in the first block ($r = .46$, $p = .02$) and in the second block ($r = .51$, $p = .01$) in the RconfB situation. N2 amplitude was not correlated with P3b amplitude in the first block ($r = .30$, $p = .19$), but was correlated in the second block ($r = .44$, $p = .03$) in the CconfB situation. These results suggest a functional relationship between N2 and P3b components for younger and older adults, such that smaller N2 amplitudes were associated with larger P3b amplitudes.

Correlation analyses showed that P3b mean amplitude was associated with decision accuracy in the RconfB situation in the first block for younger adults ($r = -.48$, $p = .02$), suggesting that higher accuracy in the RconfB situation was associated with smaller P3b mean amplitude within the first block. However, this effect did not hold in the second block ($r = -.04$, $p = .98$). Further, decision accuracy in the CconfB situation within the first block did not correlate with P3b mean amplitude ($r = -.24$, $p = .25$), with similar results concerning the second block ($r = .05$, $p = .80$). Concerning older adults, the correlation between P3b mean amplitude and decision accuracy in the RconfB situation did not reach significance in the first block ($r = -.35$, p

= .09) and was not correlated in the second block ($r = -.31, p = .13$). Similarly, P3b mean amplitude did not correlate with decision accuracy in the CconfB situation in the first ($r = .21, p = .30$) or second block ($r = .01, p = .96$).

Exploratory analysis showed a significant correlation between Berlin Numeracy Task scores and P3b mean amplitude in the RconfB situation within the first block for younger adults ($r = -.39, p = .05$), such that those with higher Berlin Numeracy Task scores produced smaller P3b mean amplitudes. No other significant correlations were observed.

In summary, younger adults showed a greater P3b amplitude in the second block in the CconfB relative to the RconfB situation. The Block by Decision Situation interaction showed that the P3b mean amplitude was smaller in the second compared with the first block for the RconfB situation, and with the CconfB situation producing a larger P3b amplitude than the RconfB situation in the second block. This result may be driven by younger adults, as the difference wave of the conflict situations was only significantly different from zero in the second block for the younger adults only. Further, results showed that the P3b and N2 were correlated, suggesting a functional relationship between these components, such that a smaller N2 was correlated with larger P3b amplitudes in both younger and older adults. Also, analysis showed that smaller P3b mean amplitudes correlated with higher decision accuracy in the belief updating task and higher scores on the Berlin Numeracy Test in the first block in the RconfB situation for younger adults.

Figure 14. P3b: Grand averaged ERPs from Pz and Cz, time-locked to sample onset. Time is on the x-axis and sample-related amplitude differences are highlighted in the 405-505 ms epoch for younger adults, and 425-525 ms epoch for older adults.

ERP Data: Feedback Effects

FRN. The 2 (age: young, old) $\times$ 2 (decision situation: RconfB, CconfB) $\times$ 2 (trial block: Block 1, Block 2) $\times$ 2 (valence: positive, negative) mixed ANOVA on FRN mean amplitude revealed a main effect of block, $F(1, 48) = 54.78, p < .001, \eta_p^2 = .53$, such that the second block ($M = 4.99 \mu V, SD = 7.07$) evoked more negative mean amplitudes than the first block ($M = 8.75 \mu V, SD = 7.40$), a main effect of decision situation, $F(1, 48) = 7.20, p = .01, \eta_p^2 = .13$, such that the CconfB situation ($M = 6.18 \mu V, SD = 6.96$) produced more negative mean amplitudes than the RconfB situation ($M = 7.57 \mu V, SD = 7.52$), a main effect of valence, $F(1, 48) = 9.86, p = .003, \eta_p^2 = .17$, such that negative feedback ($M = 5.67 \mu V, SD = 7.28$) produced more negative mean amplitudes than positive feedback ($M = 8.08 \mu V, SD = 7.73$), and a main effect of age, $F(1, 48) = 15.19, p < .001, \eta_p^2 = .24$, such that younger adults ($M = 3.47 \mu V, SD = 5.37$) produced more negative mean amplitudes than older adults ($M = 10.28 \mu V, SD = 6.89$).

Results revealed an Age by Block by Valence three-way interaction, $F(1, 48) = 13.57, p = .001, \eta_p^2 = .22$. In order to unpack the three-way interaction, stratifying by age, paired-sample t tests were conducted to examine mean amplitude in response to positive and negative feedback within each block, as well as across blocks. Concerning younger adults, negative feedback ($M = 3.09 \mu V, SD = 5.61$) produced larger FRN mean amplitudes than positive feedback ($M = 8.01 \mu V, SD = 8.41$), $t(24) = 3.56, p = .002, d = 0.71$, in the first block. However, no differences in FRN mean amplitude were observed between negative ($M = 1.49 \mu V, SD = 6.22$) and positive feedback ($M = 1.29 \mu V, SD = 7.01$) in the second block, $t(24) = -0.12, p = .23, d = -0.02$. Further, positive feedback elicited more negative mean amplitudes in the second block compared with the first block, $t(24) = 7.09, p < .001, d = 1.42$, however no differences were observed in FRN for negative feedback between the first or second block, $t(24) = 1.47, p = .34, d = 0.21$.

Concerning older adults, negative feedback ($M = 10.90 \mu\text{V}$, $SD = 7.71$) produced larger FRN than positive feedback ($M = 13.00 \mu\text{V}$, $SD = 7.09$) in the first block, $t(24) = 2.75$, $p = .01$, $d = 0.55$. Similarly, negative feedback ($M = 7.19 \mu\text{V}$, $SD = 7.93$) evoked larger FRN than positive feedback ($M = 10.01 \mu\text{V}$, $SD = 6.56$) in the second block, $t(24) = 3.61$, $p = .001$, $d = 0.71$. Further, positive feedback elicited larger FRN in the second block compared with the first block, $t(24) = 3.99$, $p = .001$, $d = 0.80$. Similarly, negative feedback evoked larger FRN in the second block compared with the first block, $t(24) = 4.44$, $p < .001$, $d = 0.89$.

For younger adults within the first block for negative feedback, decision accuracy in the RconfB situation did not correlate with the FRN ($r = .04$, $p = .54$) but did correlate with the CconfB situation ($r = -.42$, $p = .04$). In the second block, younger adults did not show a correlation with FRN mean amplitude in the RconfB situation ($r = .32$, $p = .85$) or the CconfB situation ($r = .15$, $p = .19$). Concerning older adults in the first block for negative feedback, no correlation was observed between the FRN and the RconfB situation ($r = -.33$, $p = .32$) or the CconfB situation ($r = -.12$, $p = .14$). However, a significant correlation was observed in the second block between FRN amplitude and decision accuracy in the RconfB situation ($r = -.40$, $p = .04$) but the correlation did not reach significance between FRN amplitude and accuracy in the CconfB situation ($r = -.37$, $p = .06$). The significant negative correlations suggest that higher accuracy was associated with larger FRN mean amplitudes, with younger adults showing an association in the first block between FRN and accuracy in the CconfB situation, and with older adults showing an association between FRN and decision accuracy in the RconfB situation in the second block.

Further, for younger adults in the first block, FRN in response to positive feedback correlated with accuracy in the RconfB situation ($r = -.48$, $p = .01$), and the CconfB situation ($r =$

-.45, $p = .03$). However, FRN in response to positive feedback did not correlate with accuracy in the RconfB situation ($r = -.37$, $p = .19$), or in the CconfB situation ($r = -.26$, $p = .10$) for younger adults in the second block. Concerning older adults in the first block, FRN in response to positive feedback correlated with accuracy in the RconfB situation ($r = -.41$, $p = .04$), but did not correlate with accuracy in the CconfB situation ($r = -.11$, $p = .37$). In the second block, older adults showed a correlation between FRN in response to positive feedback and accuracy in the RconfB situation ($r = -.53$, $p = .01$), but did not show this correlation in the CconfB situation ($r = -.24$, $p = .56$). The significant negative correlations suggest that higher accuracy is associated with larger negative mean amplitudes in response to positive outcomes, with younger adults showing this pattern in the first block for both conflict situations, and older adults showing this pattern of results for the RconfB situation in the first and second block.

To summarize, larger FRN mean amplitudes were observed in the second compared with the first block, in younger compared with older adults, in the CconfB situation compared with the RconfB situation, and in response to negative compared with positive feedback. Interestingly, positive feedback elicited larger FRN in the second compared with the first block in younger adults with no difference in FRN amplitude across blocks in response to negative feedback – results that converge with past research (Forder & Dyson, 2016). A larger FRN was also observed in response to negative relative to positive feedback in the first block with no difference in the second block for younger adults. Additionally, a larger FRN was observed in response to negative feedback in the first and second block for older adults relative to positive feedback, with a larger FRN in response to both negative and positive feedback in the second than the first block. Further, FRN amplitude in response to negative feedback was correlated with accuracy in the CconfB situation in the first block for younger adults and was correlated

with accuracy in the RconfB situation in the second block for older adults, such that higher accuracy was associated with larger FRN. Last, FRN in response to positive feedback was correlated with accuracy in both conflict situations in the first block for younger adults and was correlated with accuracy in the RconfB situation in the first and second block for older adults (see Figure 15).

Figure 15. FRN: Grand averaged ERPs from Cz, time-locked to sample onset. Time is on the x-axis and sample-related amplitude differences are highlighted in the 240-290 ms epoch for younger adults, and 325-375 ms epoch for older adults.

Summary of Results

The RconfB situation produced the lowest level of accuracy, the CconfB situation produced intermediate decision accuracy, and the RalignB situation produced the highest level of accuracy. Younger adults made more accurate decisions than older adults overall, and accuracy increased across blocks, in both age groups and in all decision situations. Younger adults showed the longest RTs in the RconfB situation, intermediate RTs in the CconfB situation and the shortest RTs in the RalignB situations, with RT becoming faster from the first to the second block. In contrast, older adults did not show differences in RT across decision situations or blocks. Decision accuracy in conflict situations was positively correlated in the first and second block for younger, but not older adults. Data from the self-assessment showed that younger adults were more likely to report that their decisions were “accurate” compared with older adults, and older adults were more likely to report that they felt the task was “difficult” relative to younger adults. Concerning individual differences, results showed that higher scores on the Berlin Numeracy Test, the Lipkus Numeracy Task and the Cognitive Reflection Test – Reflection subscale predicted higher decision accuracy in both conflict situations, whereas lower scores on these assessments predicted a larger difference in accuracy between conflict situations (with lower scores produced more by the RconfB situation compared with the CconfB situation) on the belief updating task.

Concerning the results from ERP analysis, younger adults were biased toward attending to the higher prior probability in the second block, whereas older adults were biased toward attending to the higher prior in both blocks, as indexed by LRP mean amplitude. Further, a larger N2 was observed in Block 2 in the RconfB situation in younger adults, with larger N2 and smaller pP amplitudes correlating with accuracy in the RconfB and CconfB situations in Block 1. For older adults, a larger pP was observed in Block 2 in the CconfB situation. N2 amplitude was

more positive than pP amplitude in younger adults, reflecting low neural activity recruited from anterior regions, with older adults showing greater pP than younger adults. Concerning individual differences, higher scores on the Lipkus Numeracy Task and lower scores on the Cognitive Reflection Test – Intuition subscale showed a general pattern of correlating with larger N2 and smaller pP amplitudes in younger adults. In older adults, higher scores on the Lipkus Numeracy Task were correlated with larger N2 and pP amplitudes. Concerning the P3b component, younger adults showed a larger P3b in Block 2 in the CconfB compared with the RconfB situation. In both younger and older adults, smaller N2 amplitudes were correlated with larger P3b components. Further, smaller P3b amplitudes were associated with higher decision accuracy and Berlin Numeracy Test scores in the RconfB situation in younger adults.

Last, larger FRN amplitudes were observed in Block 2 compared with Block 1, in younger compared with older adults, in the CconfB compared with the RconfB situation, and in response to negative compared with positive feedback. In younger adults, a larger FRN was produced in response to positive feedback in Block 2 compared with Block 1, with no FRN amplitude change across blocks for negative feedback. Additionally, a larger FRN was observed in response to negative feedback in Block 1 and 2 for older adults relative to positive feedback, with a larger FRN in response to both negative and positive feedback in Block 2 than 1. Further, larger FRN amplitude in response to negative feedback was correlated with higher accuracy in the CconfB situation in Block 1 for younger adults and was correlated with higher accuracy in the RconfB situation in Block 2 for older adults. Last, larger FRN amplitude in response to positive feedback was correlated with higher accuracy in both conflict situations in Block 1 for younger adults and was correlated with higher accuracy in the RconfB situation in both blocks for older adults.

Discussion

The goal of Experiment 2 was to build on the results of Experiment 1 by examining the potential age differences in the susceptibility to heuristics that lead to error in a belief updating task, with the addition of examining the effect of trial-by-trial feedback. In addition, I sought to investigate the neural activity that underlies belief updating in younger and older adults. The ERP technique was used to index cognitive control and response to feedback at the brain level given that it records event-related processing in real time. Moreover, a battery of assessments was used to measure individual differences in numeracy, thinking disposition, and cognitive reflection in order to examine how they relate to decision accuracy in the belief updating task. The main findings are discussed in the following sections and are organized according to behavioural and ERP results.

Behavioural Results

In line with Hypothesis 1, results showed that younger adults made more accurate decisions than older adults in conflict situations. In contrast to Hypothesis 2 (i.e., accuracy was expected to increase across trials in younger adults only), decision accuracy improved from the first to the second block in both age groups and in both decision situation. This suggests that younger and older adults can learn to update their beliefs from feedback, but with younger adults learning from feedback to a greater degree than older adults. These results support prior research showing that younger adults learn from feedback more successfully than older adults and that older adults require more trials to learn from feedback relative to younger adults (Eppinger et al., 2008; Mather & Carstensen, 2005). Replicating Experiment 1, as well as prior literature (Achtziger et al., 2014; Dave & Wolfe, 2003; Grether, 1980), the representativeness error was committed more than the conservatism error, demonstrating the prevalence of base-rate neglect (Kahneman & Frederick, 2002; Kahneman & Tversky, 1973), a result that supports Hypothesis

3. In line with Hypothesis 4, younger adults made faster decisions than older adults. However, younger adults showed a decrease in RT from the first to the second block, whereas older adults did not show a difference in RT across blocks, a finding in contrast to Hypothesis 5.

In line with Hypothesis 6, the Berlin Numeracy Test, Lipkus Numeracy Task, and the Cognitive Reflection Test - Reflection subscale predicted decision accuracy such that higher numeracy, and cognitive reflection scores predicted a smaller gap in decision accuracy between conflict situations (i.e., less likely to commit the representativeness error). Moreover, processing speed scores correlated with decision accuracy in conflict situations. These results are consistent with Experiment 1, as well as prior literature showing that numeracy and measures of fluid intelligence modulate susceptibility to decision biases (Cokely et al., in press), and for the first time shows that cognitive reflection, as indexed by the extended Cognitive Reflection Test, is a good predictor of belief updating in younger and older adults. Replicating results from Experiment 1, the Need for Cognition and Faith in Intuition scales did not predict decision accuracy. Together these results suggest that individual difference measures of ability rather than self-report measures were stronger predictors of performance on the belief updating task. Measures of ability such as numeracy level may share more method variance with the belief updating task compared with self-report measures of thinking dispositions.

Taken together, given the increase in accuracy across blocks, the results suggest that feedback may have cued reflective processing in younger and older adults. That is, older adults can learn to avoid committing heuristic-based errors when updating beliefs with feedback, but to a lesser extent than younger adults. It is also possible that an increase in decision accuracy could be due to instrumental conditioning, such that decision situations were paired with correct outcomes in memory. If instrumental conditioning underlies learning from feedback one may

expect approximately equal decision accuracy across decision situations. However, results show evidence that the representativeness error was more difficult to avoid than the conservatism error. This suggests that the increase in accuracy over time may not be due to instrumental conditioning, but rather learning to suppress and override automatic responses in certain decision situations. In addition, these results further support literature showing that numeracy underlies cognitive reflection and decision biases (Klaczynski, 2014; Pennycook et al., 2013; Peters, Diefenbach, Hess, & Västfjäll, 2008; Peters et al., 2006), and for the first time demonstrates that the extended Cognitive Reflection Test (Toplak, et al., 2014a) is a good predictor of decision accuracy in a belief updating task in younger and older adults.

ERP Results

Younger adults showed a bias towards attending to the higher prior probability in the second, but not the first block, as indexed by the LRP. This may be evidence of younger adults inhibiting a preference towards prior probabilities and waiting for sample information before contemplating a decision in the first block, whereas in the second block younger adults may have learned to attend to the higher prior. These results are in line with the notion that younger adults are adaptive decision makers that change strategies in order to improve decisions (Gigerenzer & Gaissmaier, 2011). In contrast, older adults showed a bias toward attending to the higher prior throughout the task. These results support literature showing an age-related decrease in inhibitory control (Hasher & Zacks, 1988), as well as an age-related increase in automatic processing (Peters, Finucane, MacGregor, & Slovic, 2000). Thus, in contrast to Hypothesis 7, LRP amplitude did not attenuate across blocks.

Further, in contrast to Achtziger et al. (2014) who showed that larger LRP amplitudes were correlated with committing a conservatism error, in the current experiment no correlation

was observed. Given that decision accuracy increased across blocks in the CconfB situation, and LRP amplitude either increased (younger adults) or larger LRP amplitudes were sustained (older adults) across blocks, attending to the higher prior may be an adaptive reasoning strategy, rather than evidence of conservatism as argued by Achtziger et al. (2014). Critically, Achtziger et al. (2014) did not present feedback, whereas feedback was presented in this experiment. This indicates that without feedback biasing towards the higher prior associates with committing the conservatism error. However, with feedback, younger and older adults learn to attend to the higher prior and use this information in such a way as to avoid committing this error. Critically, these findings are the first to show that the conservatism error can be avoided by younger and older adults when given the opportunity to learn, via feedback, to orient to the higher prior before, but not in lieu of, considering the sample.

Consistent with findings reported by Achtziger et al. (2014) and partially consistent with Hypothesis 8, younger adults showed a larger N2 for the RconfB compared with the CconfB situation in the second relative to the first block, with larger N2 amplitudes associating with higher accuracy in conflict situations in the first block. These results fall in line with research showing that younger adults with a higher sensitivity to detecting a conflict between using a heuristic and engaging inhibitory processes for reflection, as indexed by the N2, associate with higher decision accuracy in conflict situations (Donkers & van Boxtel, 2004; Grützmann et al., 2014). An association between N2 amplitude and decision accuracy was not observed in the second block possibly due to lack of variance in decision accuracy in conflict situations in the second block. In line with Hypothesis 9, older adults showed larger positivity in prefrontal regions (i.e., pP) for the CconfB compared with the RconfB situation in the second block. Consistent with the compensatory hypothesis (Cabeza et al., 2018; Kropotov, Ponomarev,

Tereshchenko, Müller, & Jäncke, 2016; Morcom & Johnson, 2015) and the PASA effect (Davis et al., 2008), older adults showed more prefrontal activity than younger adults possibly to compensate for deficits in other brain regions in order to engage cognitive control (Braver & Barch, 2002) and sustain good performance, albeit not to the same degree as younger adults. However, pP did not correlate with decision accuracy in older adults.

There may have been no difference in N2 amplitude for younger adults or pP for older adults between conflict situations in the first block because they were in the process of learning which situations comprise conflicting information (i.e., reflective processing was not as engaged when confronted with conflict situations in Block 1 as it was in Block 2). These results go beyond the results reported by Achtziger et al. (2014) by showing that conflict detection or engaged cognitive control, as indexed by the N2 and pP for younger and older adults respectively, also occurs when the conservatism heuristic conflicts with Bayes' theorem. Although speculative, younger adults may have shown larger N2 amplitudes for the RconfB situation, and older adults may have shown larger pP for the CconfB situation in the second block due to an age difference in perceived problem solving ability. That is, younger adults may have registered the RconfB situation more within their means of solving (i.e., a challenge compared with the CconfB situation), whereas older adults may have registered the CconfB situation as more within their means of solving (with the RconfB situation more difficult for older adults to solve).

The larger P3b amplitude observed in the CconfB situation in the second block may reflect top-down allocation of attention, given that the P3b has been associated with attentional control (Friedman, 2003, 2012; O'Connell et al., 2012). A larger P3b may have been observed in the CconfB relative to the RconfB situation in Block 2 because younger adults may have

perceived this situation as attainable to solve (i.e., a conflict situation they have learned to respond accurately to over time). Attention may be more engaged in the CconfB relative to the RconfB situation to ensure a correct decision. No difference in P3b amplitude was observed between conflict situations for older adults perhaps because older adults engaged goal-directed attention for both conflict situations. No age differences were observed in P3b amplitude, however the results suggest that younger adults' P3b activity was associated with task performance, whereas older adults' P3b activity was not. Interestingly, smaller N2 amplitudes were associated with larger P3b amplitudes for the RconfB situation in the first block and the CconfB situation in the second block for younger and older adults. For the RconfB situation, these results may suggest that when a conflict goes undetected and inhibitory processes are not engaged, attentional control is engaged in a later processing stage. However, given that decision accuracy is almost at ceiling in the CconfB situation, inhibitory control is less necessary. Rather, attentional control may be engaged in order to ensure a correct decision is made in the CconfB situation (Friedman, 2003, 2012; O'Connell et al., 2012).

Concerning Hypothesis 10, although older adults showed less sensitivity to feedback, as indexed by the decrement in FRN in response to positive and negative feedback in older relative to younger adults (Eppinger et al., 2008; Hämmerer et al., 2010; Mathewson et al., 2008; Nieuwenhuis et al., 2002; Pietschmann, Simon, Endrass, & Kathman, 2008), older adults still showed evidence of learning from Block 1 to Block 2. Specifically, both younger and older adults showed larger FRNs in the second compared with the first block, suggesting stronger representations of expected outcomes as the task advanced, results in line with the increase in decision accuracy in the second block.

Further, younger adults' decision accuracy in conflict situations correlated with FRN amplitude on correct and incorrect trials in the first block, such that larger FRN amplitudes were associated with higher decision accuracy, with no correlation observed in the latter half of the task perhaps due to a ceiling effect of decision accuracy. Older adults' accuracy correlated with FRN amplitude on correct trials across blocks but correlated with FRN amplitude on incorrect trials only in the second block in the RconfB situation. Although this could be due to a positivity bias in which older adults show a delay when learning from negative feedback (Mather & Carstensen, 2005), the data are more in line with a general decrement in learning from feedback with age. These results support research showing that older adults have difficulty learning from both positive and negative feedback due to ontogenetic changes in dopaminergic neuromodulation with age (Hämmerer et al., 2010; Nieuwenhuis et al., 2002). The results are also in line with research showing that older adults need more trials than younger adults to learn from feedback (Eppinger et al., 2008, 2009; Marschner et al., 2005).

Neural Correlates of Individual Differences

Supporting Hypothesis 11, the Lipkus Numeracy Task correlated with a number of ERP components. In younger adults, Lipkus Numeracy Task scores correlated with larger N2, and smaller pP amplitudes in both conflict situations in the first block. In older adults, Lipkus Numeracy Task scores were correlated with larger N2 and pP amplitudes in the first and second block. These results suggest that higher numeracy, as indexed by the Lipkus Numeracy Task, associates with engaging cognitive control during belief updating in both younger and older adults, with numeracy associated with neural activity in the beginning of the task for younger adults, and throughout the task for older adults. Further, as the Lipkus Numeracy Task was a

predictor of decision accuracy, together these results support the notion that numeracy underlies reflective processing (Klaczynski, 2014; Peters et al., 2006) in both younger and older adults.

The Cognitive Reflection Test – Reflection and - Intuition subscales also correlated with neural activity in both age groups, however the results are counterintuitive. For younger adults, higher scores on the Cognitive Reflection Test - Intuition subscale associated with larger N2 amplitudes in the first and second blocks in the CconfB situation. For older adults, higher Cognitive Reflection Test - Intuition (Cognitive Reflection Test - Reflection) subscale scores correlated with larger N2 (smaller pP) amplitudes in the RconfB situation in the second block. These results could suggest that those who fall prey to decision biases, as indexed by the Cognitive Reflection Test - Intuition subscale, may still detect conflict and engage inhibitory processes. Together, the results show that the Reflection subscale of the extended Cognitive Reflection Test predicts decision accuracy of belief updating. Although the Intuition subscale of the assessment did not predict decision biases, the correlates between this subscale and neural activity suggest that even those who engage reflective processes less still detect conflict and engage cognitive control processes.

Taken together, these results fall in line with two theoretical frameworks. First, similar to other conflict monitoring studies (e.g., Lucci et al., 2013), greater prefrontal activity was observed in older compared with younger adults in the same time window as the N2-P3 complex when updating beliefs in conflict situations. However, it is unclear whether this prefrontal activity represents age-related neural compensation or dedifferentiation. Age-related neural compensation is observed when older adults show more activity in a brain region (typically prefrontal areas) than younger adults when they perform a task at the same level as younger adults or when increased prefrontal activity is positively correlated with performance in older

adults but not in younger adults (termed “successful compensation”; Cabeza & Dennis, 2012; Grady, 2008, 2012). However, the results do not suggest that this activity reflects successful compensation because older adults performed worse than younger adults, and because prefrontal positivity was not correlated with task performance amongst older adults. Some researchers suggest that compensatory mechanisms might still be involved, even if performance in older adults is worse than younger adults such that older adults’ performance might be even worse without this over-recruitment (termed “attempted compensation”; Grady, 2012). If greater prefrontal activity with age is not compensatory, then what does this reflect? It could be argued that the prefrontal activity observed amongst older adults represents extraneous noise and the reduced selectivity in the recruitment of specialized neural processes (i.e., age-related dedifferentiation; Grady et al., 2002; Grady, 2012), especially because positive-going slow waves were also observed in older adults in more posterior brain regions (see P3b waveforms in Figure 14), evidence against a posterior-to-anterior shift in neural activity with age (Davis et al., 2008).

To examine whether the prefrontal positivity in older compared with younger adults is due to age-related neural compensation, future research could use transcranial magnetic stimulation (a technique that applies a series of focally directed magnetic pulses to the scalp to stimulate the underlying neural tissue; Grady, 2012) to deactivate PFC processing to observe whether this negatively effects older adults’ and not younger adults’ performance. In addition, to examine whether this activity is due to age-related dedifferentiation, fMRI could be used to show whether older adults show widespread, non-specific activation during belief updating, given that fMRI has high spatial resolution.

Second, the cognitive architecture of information processing framework postulates three main stages of cognitive reflection: Stage 1) detecting a conflict between an intuitive lure and a normative response; Stage 2) engaging Type 2, reflective processes to suppress and override Type 1 processes; Stage 3) simulating hypothetical alternative responses that are superior to the response generated by Type 1 processes (requiring declarative knowledge such as numeracy), while sustaining decoupling operations that keep hypothetical simulations and the real world separate (Stanovich, 2011; Stanovich & Evans, 2013). Given that older adults show a tendency to process information automatically rather than reflectively (Peters et al., 2007; important for Stage 1), show a deficit in inhibitory control (Hasher & Zacks, 1988; important for Stage 2), show lower numeracy skills than younger adults (Peters et al., 2006; important for Stage 3), and diminished executive functions and fluid intelligence (Salthouse, 2019; important for Stage 3) compared with younger adults, age differences in decision accuracy were expected.

Although speculative, younger and older adults may have learned to detect conflict in conflict situations from feedback, given that decision accuracy increased over time (i.e., Stage 1 of cognitive reflection). In addition, younger and older adults showed evidence of cognitive control, indexed by the N2 and pP components, as well as the slow RTs in conflict situations relative to alignment situations (i.e., Stage 2 of cognitive reflection). Reflective processing and engagement of inhibitory control may have been cued by feedback, with younger and older adults showing stronger evidence of cognitive control in the second compared with the first block as making correct decisions was learned over time. As previously described, the Cognitive Reflection Test measures the ability to suppress and override automatic processes (Stanovich et al., 2016), numeracy has been used as a measure of declarative knowledge (Clark, 2001; Perkins, 1995), and processing speed has been used to index fluid intelligence (Kane & Engle, 2002) in

cognitive reflection tasks. Results from this experiment show that cognitive reflection and numeracy predicted decision accuracy, and speed correlated with decision accuracy, regardless of age. These results support prior literature showing that reflective information processing, proficient knowledge (e.g., numeracy) and fluid intelligence underlies performance on cognitive reflection tasks.

In sum, both younger and older adults showed evidence of conflict detection (Stage 1) and cognitive control (Stage 2), with individual differences in numeracy, cognitive reflection and processing speed relating to decision accuracy regardless of age, but with younger adults performing better than older adults overall. At what stage during cognitive reflection does processing differ in younger and older adults? As a reminder, no age differences were observed on the numeracy measures or on the extended Cognitive Reflection Test. However, age differences were observed in processing speed, an index of fluid intelligence, which also correlated with decision accuracy. Age and processing speed have been shown to be highly correlated in the literature (e.g., Bors & Forrin, 1995; Manard, Carabin, Jaspar, & Collette, 2014; Salthouse, 2019). Therefore, the association between speed and accuracy observed in the current experiment may simply reflect the association between age and decision accuracy. Sustaining decoupling operations have previously been shown to be modulated by inhibitory control, working memory capacity, and fluid intelligence (Kane & Engle, 2002), cognitive processes that decline with age (Harada, Natelson Love, & Triebel, 2013; Salthouse, 2012). Therefore, although the processing speed results could reflect the link between age and accuracy, the results may also suggest that older adults have difficulty sustaining decoupling operations (Stage 3) due to diminished fluid abilities compared with younger adults. However, future research is required to

further investigate at which stage in the process of cognitive reflection older adults differ than younger adults, as well as the role of fluid intelligence in cognitive reflection in old age.

In conclusion, the results from Experiment 2 replicate and add several important findings to the literature. In line with previous research, the representativeness error was committed more than the conservatism error (Achtziger et al., 2014; Grether, 1980, 1992). Further, both younger and older adults were able to learn to avoid committing heuristic-based errors with feedback, but to a lesser degree in older adults likely due to a delay in learning from feedback relative to younger adults. Additionally, the neural correlates that underpin belief updating differ as a function of age, with older adults engaging more anterior regions than younger adults and younger adults engaging more frontocentral regions than older adults during belief updating. The behavioural and ERP results also indicate that younger and older adults detected conflict in conflict situations and engaged cognitive control processes. Moreover, individual differences in cognitive reflection and numeracy predicted decision accuracy, with processing speed correlating with decision accuracy. Given that no age differences were observed in cognitive reflection or numeracy, but age differences were observed in processing speed, age differences in belief updating may stem from sustaining decoupling operations, a process heavily dependent on fluid abilities. Taken together, these data suggest that: a) different neural processes support belief updating in younger and older adults; b) individual differences in ability such as numeracy, cognitive reflection and processing speed, rather than self-reports of thinking disposition, modulate performance in belief updating; and c) although speculative, decision biases may stem from diminished fluid abilities in old age.

Chapter 7: Experiment 3

To briefly summarize, Experiment 1 of this dissertation showed no age differences and no increase in accuracy over time in belief updating without feedback, with numeracy predicting decision accuracy in both conflict situations and processing speed associating with the avoidance of committing the conservatism error. Building on Experiment 1, Experiment 2 employed the same belief updating task but with the addition of trial-by-trial feedback, and the recording of neural activity. Results from Experiment 2 showed that accuracy increased in both age groups over time but with younger adults updating beliefs more accurately than older adults overall. ERP and behavioural data showed evidence of conflict detection and cognitive control processes (N2, pP, P3b, and slower RTs) in conflict situations in both age groups, but with older adults responding to feedback less than younger adults (FRN). Numeracy and cognitive reflection predicted decision accuracy, with no age differences on these assessments. Last, processing speed, an index of fluid intelligence, associated with decision accuracy such that higher scores in processing speed associated with higher scores in decision accuracy, with younger adults performing better on this assessment than older adults. Together, these results show that younger and older adults can learn to update beliefs if feedback is presented, with older adults updating beliefs less accurately than younger adults potentially due to a failure in sustaining decoupling operations, which depend on fluid intelligence, or from learning from feedback less successfully than younger adults, or some combination of the two, and with numeracy and cognitive reflection modulating performance, regardless of age.

Experiments 1 and 2 require making decisions under conditions of risk. However, there are two types of decisions that are made under uncertain conditions – decisions under risk and decisions under ambiguity (Knight, 1921). Whereas decisions under risk are situations in which

probabilities of possible outcomes are known or can be estimated, decisions under ambiguity are situations in which the probabilities of possible outcomes are unknown or are not well defined, and are reducible such that estimates become more accurate as the number of observations increase (Ellsberg, 1961; Huettel, Stowe, Gordon, Warner, & Platt, 2006; Knight, 1921; Payzan-LeNestour, & Bossaerts, 2011; Volz & Gigerenzer, 2011). A distinction between making decisions under risk and ambiguity was illustrated by the Ellsberg Paradox (Ellsberg, 1961). The Ellsberg Paradox shows that decision makers violate the postulates of subjective expected utility (Savage, 1954), such that a known probability (i.e., measurable risk) is preferred over an unknown probability (i.e., unmeasurable risk), even when the known probability is low (e.g., a 50% chance of winning \$100) and the unknown probability may be high (e.g., an unknown chance of winning \$100) - a phenomenon in decision theory termed ambiguity aversion (Ellsberg, 1961; Heath & Tversky, 1991; Seer et al., 2016).

Older adults show a deficit in learning tasks that involve ambiguity, a finding that has been associated with a greater aversion to ambiguous information with age (Tymula et al., 2013). Specifically, older adults show deficits in probabilistic compared with deterministic learning tasks (Eppinger et al., 2008; Samanez-Larkin, et al., 2014; Weiler, Bellebaum, & Daum, 2008), in situations where feedback is ambiguous (Eppinger et al., 2011; Herbert et al., 2011), and when changes in learning are required such as reversal learning (Eppinger et al., 2008; Eppinger & Kray, 2011; Hämmerer et al., 2011; Herbert et al., 2011; Pietschmann et al., 2011; Weiler et al., 2008) compared with their younger adult counterparts. However, older adults perform similarly to younger adults when making decisions under risk with no feedback (Armstrong & Spaniol, 2017; Denburg, Tranel, & Bechara, 2005; Mata et al., 2011; Zamarian et al., 2008). Evidence across several cognitive domains point towards a deficit in forming adequate task representations

about ambiguous information in old age (Gazzaley, 2011; Hämmerer et al., 2014; Hämmerer et al., 2010; Störmer, Winther, Li, & Andersen, 2013). Specifically, age-related impairments in representing ambiguous information have been attributed to a diminished ability to infer unknown information (Hämmerer et al., 2019) or incorporate requisite levels of ambiguity to guide learning whereby certain information is overvalued and uncertain information is undervalued (Nassar et al., 2016).

For example, Nassar and colleagues (2016) gave younger and older adults a predictive inference task in which they had to infer the mean of a noisy variable that underwent change points. The task required participants to estimate where a helicopter would drop a bag from the sky by adjusting the location of a bucket on the ground. Participants had to infer the mean location of the helicopter that occasionally changed locations and dropped bags from the sky on each trial. Results of the study revealed that older adults assigned weight to certain information and neglected to assign weight to unknown information, providing evidence of underestimating uncertainty with age. Younger adults, in contrast, showed evidence of representing ambiguous information more accurately, by acknowledging missing information and incorporating its relevance when making decisions. Whereas younger adults showed an adjusted learning rate over time, older adults did not show improvement in learning about ambiguous information. This was also true after accounting for working memory and reasoning abilities. That is, age-related deficits in learning from ambiguity were dissociated from normal decline in fluid abilities with age (Nassar et al., 2016).

Additionally, in a recent study, Hämmerer and colleagues (2019) investigated whether younger and older adults differ in their ability to form adequate task representations to make decisions using a novel probabilistic reversal learning paradigm. The task required participants to

infer the current season (i.e., winter or summer) which was not explicitly stated to participants, and to detect switches between seasons based on sales of winter-specific and summer-specific items. Compared with younger adults, older adults were less efficient at learning from feedback, and were less certain about their inferences. Further, participants' pupil diameter was measured to index increased activity of the noradrenergic system (i.e., increased neuronal firing in the locus coeruleus; Joshi, Li, Kalwani, & Gold, 2016), with increased noradrenergic modulation occurring when events or stimuli are deemed relevant to the task context (Dayan, 2012). For younger adults, larger pupil diameters were associated with informative outcomes relative to uninformative outcomes. However, for older adults, larger pupil diameters were associated with both informative and uninformative feedback, as well as less efficient belief updating and less accurate inferences about the current season relative to the younger adult group.

The reason for the age differences observed in decisions made under ambiguity compared with decisions made under risk has been attributed to the different neural structures and cognitive mechanisms these decisions rely on (Brand, Labudda, & Markowitsch, 2006). Specifically, the “limbic” loop includes the orbitofrontal cortex (OFC), ventromedial PFC (VMPFC), and limbic structures such as the amygdala, ventral striatum, and their associated neurotransmitter systems. These brain structures support decisions under ambiguity, regulating emotion processing, feedback processing, and reward-based learning (Brand et al., 2006). The “cognitive” loop comprises the dorsolateral PFC (DLPFC), lateral OFC, dorsal striatum and supports decisions under risk that rely on executive functions (Brand et al., 2006). Both the PFC, part of the “cognitive” and “limbic” loops, and the neural systems that support feedback processing and reward learning, part of the “limbic” loop, have been shown to decline with age (Eppinger et al., 2008; West, 2000). Older adults may be less efficient at making decisions under ambiguity

compared with younger adults because the systems that support feedback processing, which is required for learning to infer about ambiguous information, declines with age. However, this is speculative and requires empirical testing in future research.

Of importance to this dissertation, it is currently unknown how older adults use heuristics to make decisions when relevant information is ambiguous and needs to be inferred. As previously described, Experiments 1 and 2 of this dissertation employed an urn-ball paradigm in which participants made decisions under risk such that probability distributions (i.e., the prior and sample) were defined and could be used to make normative, Bayesian inferences about the riskiness of a decision (Savage, 1954), such as the probability of the computer drawing a sample from Urn A relative to Urn B. Given the age-related increase in the aversion to ambiguous information (Tymula et al., 2013), it is possible that ambiguous information would modulate the use of heuristics differently in older compared with younger adults (decision under uncertainty), than when information is certain (decision under risk).

Concerning individual differences, although Nassar et al. (2016) reported that age-related deficits in uncertainty-driven learning are dissociated from general decline in fluid abilities, it was of interest to examine whether fluid intelligence would associate with performance in a belief updating task, as found in the previous experiments of this dissertation. Additionally, Experiments 1 and 2 showed that numeracy and cognitive reflection predicted accuracy when making decisions under risk, however it is unknown whether individual differences in these assessments would predict accuracy when making decisions under ambiguity.

To fill these gaps in the literature, the current experiment sought to examine potential age differences concerning how decisions under ambiguity influence the use of heuristics compared with decisions under risk. Experiment 3 employed a modified version of the urn-ball paradigm

used in the previous experiments that comprised three different conditions. One condition presented the probability distributions for both the prior and sample information (i.e., certain prior, certain sample condition), which served as a control condition and was analogous to the urn-ball task used in Experiments 1 and 2. The second condition presented the probability distribution of prior information but presented ambiguous sample information (i.e., certain prior, uncertain sample condition). The third condition presented ambiguous prior information and the probability distribution of sample information (i.e., uncertain prior, certain sample condition). In line with Experiment 2, participants had to avoid committing the representativeness and conservatism errors by learning through feedback in the certain prior, certain sample condition. In the two conditions that presented uncertain information however, participants had to *infer* the unknown information and avoid committing heuristic-based decision errors by learning from feedback.

Hypotheses

In light of the results of Experiment 2 and research showing an aversion to ambiguity with age (Hämmerer et al., 2019; Nassar et al., 2016), I expected to replicate results from Experiment 2, such that younger adults would make more accurate decisions than older adults, and the representativeness error would be committed more than the conservatism error. However, I hypothesized that older adults would use prior information to make decisions in the certain prior, uncertain sample condition leading to higher accuracy in the RconfB situation and lower accuracy in the CconfB situation compared with accuracy in the certain prior, certain sample condition. In contrast, I hypothesized that older adults would use sample information to make decisions in the uncertain prior, certain sample condition leading to lower accuracy in the RconfB situation and higher accuracy in the CconfB situation compared with accuracy in the

certain prior, certain sample condition. I expected no difference in accuracy across conditions in younger adults.

Last, I predicted that higher scores on the Lipkus Numeracy Task, Berlin Numeracy Task, DSST, Cognitive Reflection Test – Reflection subscale, and Need for Cognition assessments and lower scores on the Cognitive Reflection Test – Intuition subscale and Faith in Intuition assessments would associate with more accurate decisions across decision situations, and lower scores on the Lipkus Numeracy Task, Berlin Numeracy Task, DSST, Cognitive Reflection Test – Reflection subscale, and Need for Cognition assessments and higher scores on the Cognitive Reflection Test – Intuition subscale and Faith in Intuition assessments would associate with less accurate decisions across decision situations, regardless of age. Although the thinking disposition assessments did not predict performance in the two previous experiments, it was of interest to examine whether they would predict performance when making decisions under ambiguity. See Table 17 for a summary of the hypotheses set for Experiment 3.

Table 17. *Hypotheses for Experiment 3*

List of Hypotheses	
	1 Younger adults were predicted to make more accurate decisions than older adults.
	2 The representativeness error was expected to be committed more than the conservatism error.
	3 Older adults were expected to make less decision errors in the RconfB situation and more decision errors in the CconfB situation in the certain prior, uncertain sample condition compared with the certain prior, certain sample condition.
	4 Older adults were expected to make more decision errors in the RconfB situation and less decision errors in the CconfB situation in the uncertain prior, certain sample condition compared with the certain prior, certain sample condition.
	5 Younger adults were predicted to show no difference in decision accuracy across conditions.
	6 Younger adults were predicted to produce faster RTs than older adults.
Assessments	7 Higher (lower) scores on the BNT, Lipkus, DSST, CRT-R and NFC, and lower (higher) scores on the CRT-I and FI scales were expected to associate with higher (lower) decision accuracy, regardless of age.

Note. RconfB = Representativeness heuristic conflicts with Bayes' theorem; CconfB = Conservatism heuristic conflicts with Bayes' theorem; RalignB = Representativeness heuristic aligns with Bayes' theorem; CPCS = certain prior, certain sample; CPUS = certain prior, uncertain sample; UPCS = uncertain prior, certain sample; RT = reaction time; BNT = Berlin Numeracy Test; Lipkus = Lipkus Numeracy Task; DSST = Digit Symbol Substitution Test; CRT-R = Cognitive Reflection Test – Reflection subscale; CRT-I = Cognitive Reflection Test – Intuition subscale; NFC = Need for Cognition; FI = Faith in Intuition.

Method

Participants

The final sample included 30 younger adults (ages 18-35) and 30 older adults (ages 65-85). All participants completed the battery of assessments described in the General Methods section (pages 56-61), with the addition of the extended Cognitive Reflection Test. Consistent with Experiments 1 and 2, several age differences emerged. Older adults were more educated, had higher vocabulary scores, lower self-reported symptoms of anxiety and negative affect, and slower processing speed as indexed by the DSST, compared with younger adults. A total of six

participants were excluded and later replaced. Two younger adults were excluded because they scored above the cut-off on the depression and anxiety subscales of the DASS-21, one younger adult and three older adults were excluded because they scored below the cut-off in the RalignB situations (i.e., below 70%), an index of task comprehension. See the General Methods section (pages 47-48) for a description of the inclusion and exclusion criteria. The characteristics of the final sample are presented in Table 18.

Table 18. *Characteristics of the Final Sample in Experiment 3*

		Younger Adults (n = 30)	Older Adults (n = 30)	<i>p</i>	<i>d</i>
		<i>M (SD)</i>	<i>M (SD)</i>		
Demographics					
	Age in Years	24.57 (4.88)	72.63 (5.18)	< .001	-9.55
	Age Range	18-35	65-85	-	-
	Male/Female Ratio	15/15	15/15	1.0	-
	Education in Years	15.63 (2.08)	17.87 (2.81)	.001	-0.91
	Right/Left Handed	27/3	25/5	.71	-
Emotion					
Assessments	Depression ^a	4.93 (5.22)	3.67 (4.90)	.34	0.25
	Anxiety ^a	5.13 (4.22)	1.73 (3.10)	.001	0.92
	Stress ^a	8.73 (6.29)	5.73 (5.84)	.06	0.49
	Positive Affect ^b	31.80 (8.53)	33.70 (8.40)	.39	-0.22
	Negative Affect ^b	13.77 (5.06)	11.20 (1.94)	.01	0.67
Cognitive					
Assessments	MMSE ^c	-	28.73 (1.20)	-	-
	DSST ^c	86.87 (10.42)	60.33 (11.13)	< .001	2.46
	Mill Hill ^c	17.53 (4.28)	21.60 (3.95)	< .001	-0.99
	BNT ^c	1.57 (1.28)	1.20 (1.24)	.27	0.29
	Lipkus Numeracy ^c	10.47 (1.93)	9.63 (2.39)	.14	0.39
	Faith in Intuition	5.84 (1.59)	5.88 (1.98)	.93	-0.02
	Need for Cognition	65.93 (11.08)	62.20 (9.97)	.18	0.35
	CRT-R	2.93 (2.18)	2.30 (1.84)	.23	0.31
	CRT-I	3.00 (1.87)	2.90 (1.58)	.82	0.06

Note. Between-group comparisons were made using independent-sample *t* tests, aside from gender and handedness which were examined using Pearson's chi square. Standard deviations are in parentheses. ^aSubscales of the 21-item Depression Anxiety and Stress Scale; ^bSubscales of the Positive and Negative Affect Schedule; ^cScores reflect number of correct solutions; MMSE = Mini-Mental State Examination; DSST = Digit Symbol Substitution Test; Mill Hill = Mill Hill Vocabulary Test; BNT = Berlin Numeracy Test; Lipkus Numeracy = Lipkus Numeracy Task; Cognitive Reflection Test – Reflection subscale; Cognitive Reflection Test – Intuition subscale. *d* = Cohen's *d* for standardized effect size estimates of the mean difference between groups.

Materials

The stimuli in Experiment 3 were presented differently than the previous experiments in order to vary the certainty of the prior and sample information. In Experiments 1 and 2, the prior was presented in the form of numbers that indicated the probability of the computer choosing an urn to draw a sample (e.g., there is a 75% likelihood the computer will choose the left urn; 25% right urn). In Experiment 3, however, twelve urns were presented, along with a representative proportion of eight blue and green balls within each urn. Of the twelve urns, two types of urns existed – one type of urn contained 75% of one colour and 25% of another colour, the other type of urn contained 50% of one colour and 50% of another colour. Additionally, there were two different types of urn proportions – one type of urn proportion consisted of 75% in one urn (9/12) and 25% in the other urn (3/12), and the other type of urn proportion consisted of 50% in one urn (6/12) and 50% in the other urn (6/12). The urns were presented in two horizontal rows with six urns per row. The majority colour was counterbalanced, and always presented on top, and the minority colour was always presented on the bottom. Sample stimuli consisted of four balls positioned vertically and were presented below the two rows of prior information in the middle of the screen. Below the sample, the two urn types were presented indicating the two decision options, with one positioned on the left and one on the right in the center of the screen.

The certain prior, certain sample condition presented all prior and sample information (see Figure 16A). The certain prior, uncertain sample condition presented all the prior information, but presented only a portion of the sample information, such that one of the four balls were half blue and half green representing a “masked” ball (i.e., the ball was either blue or green), with the majority colour as the top half of the ball and minority colour the bottom half of the ball. The masked ball was always positioned as the first ball of the sample (see Figure 16B).

The actual colour of the masked ball of the sample was always the majority colour. The uncertain prior, certain sample condition presented all the sample information but masked some of the prior information. Specifically, eight of the twelve priors were masked. Eight balls with half blue and half green balls represented the urns that were masked (see Figure 16C). The actual urn proportions of the masked urns were either one of the two types of urns (i.e., 75%-25% or 50%-50%). In line with Experiment 2, feedback was provided indicating whether the participant made a “correct” or “incorrect” decision on each trial.

The colours of the stimuli were the same as the previous experiments (see pages 55-56 in the General Method section), however the size differed such that the blue and green balls were smaller (image size 40×40 pixels) in the current study compared with Experiments 1 and 2 because there were more stimuli to present simultaneously. However, in line with the previous experiments, all stimuli were presented in the center of the computer screen within a space of 220×220 pixels.

In Experiments 1 and 2, participants completed 424 trials total. Experiment 3, in contrast, included 330 trials total, with 110 trials per condition (i.e., 30 trials per decision situation of interest, and 20 filler trials), with two mandatory 1-minute breaks separating each condition, and three 30-second breaks in the middle (i.e., 55th trial) of each condition separating counterbalanced colour (i.e., majority blue [green] and minority green [blue] balls counterbalanced in each condition).

Procedure

Belief updating task. The practice task included a total of 16 trials, with 4 trials per condition, and with conditions counterbalanced¹⁴. The concept of the belief updating task in the

¹⁴ No feedback was included in the practice task in order to avoid learning effects prior to the experiment beginning.

current experiment was the similar to Experiments 1 and 2. That is, participants had to decide from which of two urns a computer-drawn sample had most likely originated, given the prior and sample information. However, two conditions were added that varied the certainty of prior and sample information. In the certain prior, certain sample condition, participants were first presented with prior information (i.e., twelve urns containing eight representative blue and green balls), then the sample of four balls drawn from one of the urns, and then the decision prompt indicating decision options.

In the certain prior, uncertain sample condition the same trial sequence applied, except that the top ball of the sample was masked. In this condition, participants were told that the masked ball drawn by the computer could either be blue or green. Unbeknownst to participants, the masked ball in the sample was the majority colour ball. The participants' goal was to learn through feedback that the masked ball represented the majority colour in order to reduce uncertainty of the sample information (i.e., infer hidden sample information) and increase decision accuracy. In the uncertain prior, certain sample condition the same trial sequence applied except that eight of the twelve urns were masked. In this condition, participants were told that the masked urns contained proportions of blue and green balls that represented one of the two types of urns (i.e., 75%-25% or 50%-50%), however the number of each type of urn remained unknown. Unbeknownst to participants, the four urns with their contents visible reflected the same information as all twelve urns presented in the certain prior, certain sample condition. That is, if the four unmasked urns showed three 75%-25% urn types and one 50%-50% urn type, the masked urns reflected the same proportions, such that there were nine 75%-25% urn types, and three 50%-50% urn types. The participants' goal was to learn through feedback that the eight masked urns represented the same information the four unmasked urns

represented in order to reduce uncertainty of the prior information (i.e., infer the hidden prior information).

Single trial sequence. First, a fixation cross was presented for 1000 ms to cue the beginning of a trial. Next, twelve urns with eight half blue, half green balls were presented for 1000 ms in two rows at the top of the screen¹⁵. Prior information was then presented for 3000 ms. Subsequently, sample information was presented in addition to prior information for 2000 ms, and then the decision prompt was presented at the bottom of the screen (there was no time maximum to make a decision). Sample information was presented for 2000 ms before the participant could make a decision because this experiment presented more stimuli than the previous experiments, and I wanted to ensure participants attended to both the prior and the sample before a decision could be made. Once a decision was made, feedback (“correct” or “incorrect”) was presented for 1000 ms. Figure 16, Panels A-C illustrates a single-trial sequence of the situation in which the representativeness heuristic conflicts with Bayes’ theorem in each condition of Experiment 3.

¹⁵ The half blue-half green balls reflected masked information within the twelve urns and was used to cue the participant that prior information was about to be presented.

Figure 16. Sample trial of the representativeness heuristic conflicting with Bayes' theorem in Experiment 3. Panel A demonstrates this decision situation in the certain prior, certain sample condition; Panel B demonstrates this decision situation in the certain prior, uncertain sample; Panel C demonstrates this decision situation in the uncertain prior, certain sample condition. The labels described in Panel A also apply to Panels B and C.

Cognitive and affective assessments. Upon task completion, the battery of assessments outlined in the General Methods section (pages 56-61) was administered, including the extended Cognitive Reflection Test (Toplak et al., 2014a).

Data Analysis

Data analysis for Experiment 3 followed the procedures outlined in the General Methods section (pages 62-64). In addition to decision situation and age, condition (certain prior certain sample, certain prior uncertain sample, uncertain prior certain sample) was also included as a factor in the omnibus ANOVA.

Results

Decision Accuracy for All Decision Situations

Decision accuracy for each of the 15 prior and sample combinations is provided in Table 19, separately for each age group and condition. Following the previous experiments, the situations of interest are bolded in the table. In line with the previous experiments, independent sample *t* tests were conducted across counterbalancing version on decision accuracy for each decision situation of interest within each condition. No differences in counterbalancing version were observed in the decision situations or conditions ($ps > .05$).

Table 19. Means and Standard Deviations for Accuracy in Experiment 3

Prior		Sample				
		0	1	2	3	4
75						
CPCS		95.00 (15.26)	94.20 (9.47)	71.93 (29.57)	98.33 (9.13)	100.00 (0.00)
		85.00 (32.56)	83.93 (25.46)	50.20 (27.63)	96.67 (18.26)	91.67 (18.95)
CPUS		91.67 (18.95)	92.37 (11.40)	54.50 (26.82)	90.00 (20.34)	98.33 (9.13)
		83.33 (30.32)	78.87 (25.30)	53.07 (22.56)	81.67 (27.80)	90.00 (24.21)
UPCS		95.00 (15.26)	91.10 (13.16)	72.20 (28.25)	100.00 (0.00)	98.33 (9.13)
		85.00 (29.80)	86.77 (19.14)	49.97 (25.51)	91.67 (23.06)	88.33 (21.51)
		2	15	15	2	2
50						
CPCS		100.00 (0.00)	100.00 (0.00)	97.07 (7.23)	97.53 (5.09)	100.00 (0.00)
		96.67 (12.68)	86.67 (29.16)	87.07 (16.56)	91.70 (11.58)	93.33 (22.16)
CPUS		98.33 (9.13)	96.67 (12.68)	90.23 (11.20)	75.73 (25.58)	97.80 (8.37)
		86.67 (29.16)	91.67 (23.06)	87.33 (15.57)	62.13 (23.65)	90.07 (17.76)
UPCS		98.33 (9.13)	100 (0.00)	95.07 (7.43)	97.73 (4.77)	98.90 (6.03)
		98.33 (9.13)	91.67 (23.06)	84.00 (16.50)	84.67 (16.68)	90.00 (25.02)
		2	2	15	15	3
25						
CPCS		98.33 (9.13)	93.40 (13.43)	96.67 (12.68)	58.83 (32.13)	88.40 (14.86)
		93.33 (17.29)	93.37 (16.11)	83.33 (27.33)	37.23 (30.31)	77.93 (26.48)
CPUS		95.00 (20.13)	94.50 (12.51)	90.00 (24.21)	74.57 (21.45)	84.80 (13.90)
		93.33 (17.29)	91.13 (23.04)	86.67 (22.49)	67.57 (19.74)	64.00 (26.89)
UPCS		100.00 (0.00)	97.80 (8.37)	90.00 (20.34)	65.67 (28.67)	87.07 (15.06)
		93.33 (21.71)	87.80 (28.34)	73.33 (31.44)	36.50 (29.16)	78.00 (27.68)
		2	3	2	15	15

Note. CPCS = certain prior, certain sample; CPUS = certain prior, uncertain sample; UPCS = uncertain prior, certain sample. Values represent proportion (%) of correct decisions; Prior = prior probability for the left urn; Sample = number of majority colour balls of sample; Younger adults' accuracy is on the top line, older adults' accuracy is on the bottom line. Standard deviations are in parentheses. Bold indicates decision situations of interest. Total number of trials are indicated on the third row.

Combining like-situations. Collapsing across age, within the certain prior, certain sample condition, paired-sample t -tests showed no difference in the RalignB situation, $t(59) = -1.32, p = .19, d = -0.17$. However, differences between RconfB situations, $t(59) = -4.05, p < .001, d = -0.52$, and CconfB situations, $t(59) = -2.39, p = .02, d = -0.32$, were observed. Within the certain prior, uncertain sample condition, the RalignB situation, $t(59) = 5.96, p < .001, d = 0.77$, the RconfB, $t(59) = 5.12, p < .001, d = 0.65$, and the CconfB situations, $t(59) = -3.62, p < .001, d = -0.46$, showed significant differences. Within the uncertain prior, certain sample condition, the RalignB situation, $t(59) = -0.73, p = .47, d = 0.07$, showed no differences, however the RconfB, $t(59) = -2.89, p = .01, d = -0.37$, and the CconfB situations, $t(59) = -2.20, p = .03, d = -0.26$, showed significant differences. Like-situations were combined regardless of these differences as the primary interest was to examine decision situation type, and not variations within decision situation type.

Decision Accuracy

A 2 (age: young, old) $\times 3$ (decision situation: RconfB, CconfB, RalignB) $\times 3$ (condition: certain prior certain sample, certain prior uncertain sample, uncertain prior certain sample) mixed ANOVA on decision accuracy was conducted. Analyses revealed a main effect of decision situation, $F(1.28, 74.22) = 83.66, p < .001, \eta_p^2 = .59$, with pairwise comparisons showing that the RconfB ($M = .58, SD = .22$) led to less accurate decisions than the CconfB ($M = .84, SD = .15$), $t(59) = -7.96, p < .001, d = -1.03$ and the RalignB situations ($M = .88, SD = .09$), $t(59) = -12.72, p < .001, d = -1.64$, with no differences between the CconfB and RalignB situations, $t(59) = -0.66, p = .12, d = -0.09$, a main effect of age, $F(1, 58) = 25.32, p < .001, \eta_p^2 = .30$, such that younger adults ($M = .83, SD = .10$) made more accurate decisions than older adults ($M = .70, SD = .09$), as well as a main effect of condition, $F(2, 116) = 7.65, p = .001, \eta_p^2 = .12$, such that

the certain prior, certain sample condition ($M = .78, SD = .14$) led to more accurate decisions than the certain prior, uncertain sample condition ($M = .74, SD = .12$), $t(59) = 3.33, p = .01, d = 0.43$, the uncertain prior, certain sample condition ($M = .77, SD = .13$) led to more accurate decisions than the certain prior, uncertain sample condition, $t(59) = -3.01, p = .01, d = -0.38$, and with no difference between the certain prior, certain sample and uncertain prior, certain sample condition, $t(59) = 0.66, p = .90, d = 0.09$.

Additionally, a Decision Situation $\times$ Age $\times$ Condition three-way interaction was observed, $F(3.26, 189.28) = 6.42, p < .001, \eta_p^2 = .10$. In order to unpack this interaction, I conducted follow-up two-way ANOVAs of decision situation and condition on accuracy, separately for younger and older adults. Concerning younger adults, a main effect of decision situation was observed, $F(1.40, 40.61) = 41.42, p < .001, \eta_p^2 = .59$, a main effect of condition was observed, $F(2, 58) = 9.91, p < .001, \eta_p^2 = .26$, and a Decision Situation by Condition interaction was observed, $F(2.34, 67.75) = 4.21, p = .003, \eta_p^2 = .13$.

To probe the interaction, pairwise comparisons showed that decision situation differed within the certain prior, certain sample condition, $F(1.17, 34.01) = 29.87, p < .001, \eta_p^2 = .51$, such that the RconfB situation ($M = .65, SD = .29$) produced less accurate decisions than the CconfB situation ($M = .91, SD = .10$), $t(29) = -4.77, p < .001, d = -0.87$, and the RalignB situation ($M = .97, SD = .05$), $t(29) = -6.39, p < .001, d = -1.17$, and the CconfB situation produced less accurate decisions than the RalignB situation, $t(29) = -3.30, p = .01, d = -0.61$. Decision situation also differed in accuracy in the certain prior, uncertain sample condition, $F(2, 58) = 27.57, p < .001, \eta_p^2 = .49$, such that the RconfB situation ($M = .65, SD = .20$) produced less accurate decisions than the CconfB situation ($M = .89, SD = .09$), $t(29) = -7.58, p < .001, d = -1.38$ and the RalignB situation ($M = .83, SD = .15$), $t(29) = -4.82, p < .001, d = -0.88$, however

the CconfB and RalignB situations did not differ in accuracy, $t(29) = 1.82, p = .22, d = 0.26$.

Similarly, decision situation also differed in accuracy in the uncertain prior, certain sample condition, $F(1.34, 39.99) = 25.58, p < .001, \eta_p^2 = .47$, such that the RconfB situation ($M = .69, SD = .26$) produced less accurate decisions than the CconfB situation ($M = .89, SD = .12$), $t(29) = -4.27, p = .001, d = -0.78$, and the RalignB situation ($M = .97, SD = .04$), $t(29) = -6.14, p < .001, d = -1.12$, and the CconfB situation produced less accurate decisions than the RalignB situation, $t(29) = -3.21, p = .01, d = -0.59$.

Further, the RconfB situation did not differ across conditions, $F(2, 58) = .66, p = .52, \eta_p^2 = .02$, the CconfB situation did not differ across conditions, $F(1.65, 47.72) = 1.25, p = .29, \eta_p^2 = .04$, however differences were observed within the RalignB situation across conditions, $F(1.29, 37.34) = 23.97, p < .001, \eta_p^2 = .45$, such that decision accuracy was higher in the certain prior, certain sample ($M = .97, SD = .05$) compared with the certain prior, uncertain sample condition ($M = .83, SD = .15$), $t(29) = 5.25, p < .001, d = 0.96$, decision accuracy was higher in the uncertain prior, certain sample ($M = .96, SD = .04$) compared with the certain prior, uncertain sample condition, $t(29) = -4.97, p < .001, d = -0.91$, and with no differences in accuracy between certain prior, certain sample and uncertain prior, certain sample conditions, $t(29) = 0.71, p = .48, d = 0.10$.

Concerning older adults, a main effect of decision situation was observed, $F(1.12, 32.56) = 43.16, p < .001, \eta_p^2 = .60$. Although there was no main effect of condition, $F(2, 58) = 1.08, p = .35, \eta_p^2 = .04$, a significant Decision Situation by Condition interaction was observed, $F(4, 116) = 12.75, p < .001, \eta_p^2 = .31$. Pairwise comparisons showed that decision situation differed within the certain prior, certain sample condition, $F(1.36, 39.50) = 42.19, p < .001, \eta_p^2 = .59$, such that the RconfB situation ($M = .44, SD = .25$) produced less accurate decisions than the CconfB

situation ($M = .81$, $SD = .24$), $t(29) = -5.42$, $p < .001$, $d = -0.99$, and the RalignB situation ($M = .90$, $SD = .10$), $t(29) = -10.27$, $p < .001$, $d = -1.87$, however there was no difference in accuracy between the CconfB situation and the RalignB situation, $t(29) = -1.06$, $p = .14$, $d = -0.15$.

Decision situation differed within the certain prior, uncertain sample condition, $F(2, 58) = 5.03$, $p = .01$, $\eta_p^2 = .15$, such that the RconfB situation ($M = .60$, $SD = .17$) did not differ in accuracy from the CconfB situation ($M = .71$, $SD = .22$), $t(29) = -1.96$, $p = .17$, $d = -0.28$, and the CconfB situation did not differ from the RalignB situation ($M = .75$, $SD = .15$), $t(29) = -0.83$, $p = .80$, $d = -0.12$, however the RconfB produced less accurate decisions than the RalignB situation, $t(29) = -3.26$, $p = .01$, $d = -0.59$. Decision situation also differed within the uncertain prior, certain sample condition, $F(1.34, 38.73) = 45.23$, $p < .001$, $\eta_p^2 = .61$, such that the RconfB situation ($M = .43$, $SD = .23$) produced less accurate decisions than the CconfB situation ($M = .83$, $SD = .19$), $t(29) = -6.19$, $p < .001$, $d = -1.13$, and the RalignB situation ($M = .84$, $SD = .12$), $t(29) = -9.57$, $p < .001$, $d = -1.74$, however no difference in accuracy was observed between the CconfB and the RalignB situations, $t(29) = -0.48$, $p = .95$, $d = -0.07$.

Critically, the RconfB situation significantly differed across conditions, $F(2, 58) = 10.70$, $p < .001$, $\eta_p^2 = .27$, such that the certain prior, uncertain sample condition ($M = .60$, $SD = .17$) produced higher accuracy than the certain prior, certain sample ($M = .44$, $SD = .25$), $t(29) = 3.58$, $p = .004$, $d = 0.65$, and uncertain prior, certain sample conditions ($M = .43$, $SD = .23$), $t(29) = 4.38$, $p < .001$, $d = 0.80$, however no differences in accuracy were observed between the certain prior, certain sample and the uncertain prior, certain sample conditions, $t(29) = 0.16$, $p = 1.0$, $d = 0.02$. The CconfB situation also differed across conditions, $F(2, 58) = 5.68$, $p = .01$, $\eta_p^2 = .16$, such that the certain prior, uncertain sample condition ($M = .71$, $SD = .22$) produced less accurate decisions than the uncertain prior, certain sample condition ($M = .83$, $SD = .19$), $t(29) = -2.85$, p

= .02, $d = -0.52$, with no differences observed between the certain prior, uncertain sample condition and the certain prior, certain sample condition ($M = .81$, $SD = .24$), $t(29) = -2.41$, $p = .06$, $d = -0.34$, and with no difference in accuracy was observed between the certain prior, certain sample and uncertain prior, certain sample conditions, $t(29) = -0.59$, $p = .92$, $d = -0.08$. Last, accuracy in the RalignB situation differed across conditions, $F(1.59, 46.23) = 15.30$, $p < .001$, $\eta_p^2 = .35$, such that the certain prior, certain sample condition ($M = .90$, $SD = .10$) produced more accurate decisions than the certain prior, uncertain sample condition ($M = .75$, $SD = .15$), $t(29) = 5.12$, $p < .001$, $d = 0.94$, and the uncertain prior, certain sample condition ($M = .84$, $SD = .12$), $t(29) = 2.71$, $p = .03$, $d = 0.50$, and the uncertain prior, certain sample condition produced more accurate decisions than the certain prior, uncertain sample condition, $t(29) = -3.03$, $p = .02$, $d = 0.56$.

In summary, condition did not affect decision accuracy in conflict situations for younger adults. Critically however, condition influenced decision accuracy amongst older adults, particularly in the RconfB situation, with higher accuracy produced by the certain prior, uncertain sample condition compared with the certain prior, certain sample and uncertain prior, certain sample conditions, and with no difference between the latter conditions (see Figure 17).

Figure 17. Accuracy proportion for each decision situation and condition in Experiment 3. Error bars represent standard error of the mean. RconfB = representativeness heuristic conflicts with Bayes' theorem; CconfB = conservatism heuristic conflicts with Bayes' theorem; RalignB = representativeness heuristic aligns with Bayes' theorem; CPCS = certain prior, certain sample; CPUS = certain prior, uncertain sample; UPCS = uncertain prior, certain sample

Decision RT

A 2 (age: young, old) $\times$ 3 (decision situation: RconfB, CconfB, RalignB) $\times$ 3 (condition: certain prior certain sample, certain prior uncertain sample, uncertain prior certain sample) mixed ANOVA on decision RT was conducted and revealed a main effect of decision situation, $F(1.65, 95.82) = 9.12, p < .001, \eta_p^2 = .14$, a main effect of condition, $F(1.65, 95.82) = 9.12, p < .001, \eta_p^2 = .14$, as well as a main effect of age, $F(1, 58) = 35.08, p < .001, \eta_p^2 = .38$, such that younger adults ($M = 2009.37, SD = 1167.45$) made faster decisions than older adults ($M = 5207.60, SD = 2717.56$).

Further, a significant Decision Situation by Condition interaction on RT emerged, $F(4, 232) = 8.48, p < .001, \eta_p^2 = .13$. To probe this interaction, repeated-measures ANOVAs with

pairwise comparisons were conducted to examine potential decision situation effects within each condition on RT, as well as potential differences in RT for each condition within each decision situation. Within the certain prior, certain sample condition, a main effect of decision situation was observed, $F(1.81, 106.79) = 4.14, p = .02, \eta_p^2 = .07$. Specifically, the RconfB situation ($M = 3646.91, SD = 3112.79$) produced longer RTs than the RalignB situation ($M = 2998.06, SD = 2690.23$), $t(59) = 3.36, p = .004, d = 0.43$, however no differences in RT were observed between the RconfB and the CconfB situations ($M = 3232.17, SD = 3537.36$), $t(59) = 1.81, p = .21, d = 0.26$, or between the CconfB and RalignB situations, $t(59) = 0.91, p = .75, d = 0.12$. Within the certain prior, uncertain sample condition, a main effect of decision situation was observed, $F(1.77, 104.68) = 16.10, p < .001, \eta_p^2 = .21$. Pairwise comparisons showed that the RconfB situation ($M = 4873.45, SD = 3803.95$) produced longer RTs than the CconfB situation ($M = 3492.29, SD = 2743.82$), $t(59) = 4.76, p < .001, d = 0.61$, the RalignB situation ($M = 4615.49, SD = 3605.82$) produced longer RTs than the CconfB situation, $t(59) = -4.16, p < .001, d = -0.54$, however no differences were observed between the RconfB and RalignB situations, $t(59) = 1.23, p = .53, d = 0.16$. Within the uncertain prior, certain sample condition, a main effect of decision situation was observed, $F(2, 118) = 9.47, p < .001, \eta_p^2 = .14$. Pairwise comparisons showed that the RconfB situation ($M = 3719.63, SD = 3209.15$) produced longer RTs than the CconfB situation ($M = 2917.50, SD = 2462.93$), $t(59) = 3.59, p = .002, d = 0.46$, and the RalignB situation ($M = 2980.86, SD = 2466.99$), $t(59) = 3.91, p = .001, d = 0.50$, and with no differences in RT between the CconfB and RalignB situations, $t(59) = -3.20, p = .98, d = -0.41$.

Further, within the RconfB situation, a main effect of condition was observed, $F(1.69, 99.56) = 7.08, p = .001, \eta_p^2 = .12$. Specifically, the certain prior, certain sample condition produced faster RTs than the certain prior, uncertain sample condition, $t(59) = -2.85, p = .02, d$

= -0.37, the uncertain prior, certain sample condition produced faster RTs than the certain prior, uncertain sample condition, $t(59) = 3.19, p = .01, d = 0.41$, however no RT differences emerged between certain prior, certain sample and uncertain prior, certain sample conditions, $t(59) = -0.25, p = .99, d = -0.32$. Within the CconfB situation, no main effect of condition was observed, $F(1.69, 99.46) = 1.66, p = .20, \eta_p^2 = .03$. However, within the RalignB situation, a main effect of condition was observed, $F(1.50, 88.75) = 21.49, p < .001, \eta_p^2 = .27$. Pairwise comparisons showed that the certain prior, certain sample condition produced faster RTs than the certain prior, uncertain sample condition, $t(59) = -4.61, p < .001, d = -0.59$, but did not differ in RT from the uncertain prior, certain sample condition, $t(59) = 0.08, p = 1.0, d = 0.01$, and the uncertain prior, certain sample condition produced faster RTs than the certain prior, uncertain sample condition, $t(59) = 5.74, p < .001, d = 0.74$.

In summary, the results show that the RconfB situation produced longer RTs than the CconfB and RalignB situations, and the certain prior, uncertain sample condition produced longer RTs than the certain prior, certain sample and uncertain prior, certain sample conditions (see Figure 18).

Figure 18. Reaction time for each decision situation and condition in Experiment 3. Error bars represent standard error of the mean. RconfB = representativeness heuristic conflicts with Bayes' theorem; CconfB = conservatism heuristic conflicts with Bayes' theorem; RalignB = representativeness heuristic aligns with Bayes' theorem; CPCS = certain prior, certain sample; CPUS = certain prior, uncertain sample; UPCS = uncertain prior, certain sample.

Relationship between Representativeness and Conservatism

Separate Pearson bivariate correlational analyses were conducted for younger and older adults. No significant correlation between RconfB and CconfB decision accuracy was observed for younger adults ($r = .13, p = .49$) or older adults ($r = -.21, p = .27$) within the certain prior, certain sample condition. Within the certain prior, uncertain sample condition, younger adults showed a significant positive correlation ($r = .49, p = .01$) between RconfB and CconfB decision accuracy (see Figure 19), however older adults did not show a significant correlation ($r = -.29, p = .12$). Last, within the uncertain prior, certain sample condition, younger adults did not show a significant correlation ($r = .26, p = .17$), however older adults showed a negative correlation between RconfB and CconfB decision accuracy ($r = -.38, p = .04$; see Figure 20).

Figure 19. Association of accuracy in Experiment 3 in CPUS condition for younger adults.

Figure 20. Association of accuracy in Experiment 3 in UPCS condition for older adults.

Self-Assessment

Chi-square tests were conducted on categorical items (i.e., items 1-6) of the self-assessment, and independent-sample t tests were conducted on scaled items (i.e., items 7-9) of the self-assessment for each age group. Within the uncertain prior, certain sample condition, younger adults reported that “both the likelihood and sample information” influenced their

decisions the most, whereas older adults reported that the “sample” influenced their decisions the most, $\chi^2(2) = 14.22, p = .001, d = 0.49$. Further, younger adults ($M = 4.13, SD = .86$) were more likely to report that their decisions were “accurate” compared with older adults ($M = 3.57, SD = .97$), $t(58) = 2.39, p = .02, d = 0.61$, and older adults ($M = 2.73, SD = .99$) were more likely to report that they felt the task to be “difficult” relative to younger adults ($M = 3.30, SD = .99$), $t(58) = 2.23, p = .03, d = 0.58$. Last, younger adults ($M = 3.47, SD = 1.01$) were more likely to report that they felt more “confident working with numbers in general” compared with older adults ($M = 2.73, SD = 1.17$), $t(58) = 2.60, p = .01, d = 0.68$. No other age-related differences were observed (see Table 20).

The self-assessment also included an item where participants could provide any other comments about the task (e.g., detailing how decisions were made). In the participant comments, 11 (36.7%) younger and 2 (6.7%) older adults explicitly stated that they inferred the masked ball in the sample to be the majority colour ball in the certain prior, uncertain sample condition, and 19 (63.3%) younger and 4 (13.3%) older adults explicitly stated that they inferred the masked urns (i.e., unknown prior information) in the uncertain prior, certain sample condition to be the same as the four unmasked urns.

In summary, older adults were more likely to report that they used the sample to make decisions in the uncertain prior, certain sample condition compared with younger adults who reported using both the prior and sample to make decisions. Older adults were also more likely to report that they felt they made less accurate decisions, felt the task to be more difficult, and were less confident working with numbers in general compared with younger adults. Further, participants’ comments about their task performance revealed that more younger adults than

older adults explicitly stated that they were able to make inferences about prior and sample information that was masked.

Table 20. *Self-Assessment for Experiment 3*

Item	Younger Adults (n = 30)	Older Adults (n = 30)	<i>p</i> -value	<i>d</i>
CPCS - Influence Decision				
Likelihood	3 (10%)	2 (7%)	.64	0.12
Sample	2 (7%)	4 (13%)		
Both	25 (83%)	24 (80%)		
CPCS - Decision Strategy				
Consistent	23 (77%)	25 (83%)	.52	-0.08
Switched	7 (23%)	5 (17%)		
Neither	0 (0%)	0 (0%)		
CPUS - Influence Decision				
Likelihood	13 (43%)	16 (53%)	.08	0.29
Sample	1 (3%)	5 (17%)		
Both	16 (53%)	9 (30%)		
CPUS - Decision Strategy				
Consistent	21 (70%)	27 (90%)	.11	0.27
Switched	7 (23%)	3 (10%)		
Neither	2 (7%)	0 (0%)		
UPCS - Influence Decision				
Likelihood	9 (30%)	5 (17%)	.001	0.49
Sample	4 (13%)	18 (60%)		
Both	17 (57%)	7 (23%)		
UPCS - Decision Strategy				
Consistent	20 (67%)	26 (87%)	.07	-0.24
Switched	10 (33%)	4 (13%)		
Neither	0 (0%)	0 (0%)		
Accuracy				
Very inaccurate	1 (3%)	1 (3%)	.02	0.61
Somewhat inaccurate	1 (3%)	5 (17%)		
Neither inaccurate nor accurate	0 (0%)	2 (7%)		
Somewhat accurate	19% (63%)	20 (67%)		
Very accurate	9 (30%)	2 (7%)		
Task Difficulty				
Very difficult	0 (0%)	2 (7%)	.03	0.58
Somewhat difficult	9 (30%)	12 (40%)		
Neither difficult nor easy	5 (17%)	9 (30%)		
Somewhat easy	14 (47%)	6 (20%)		
Very easy	2 (7%)	1 (3%)		
Confidence with Numbers				
Not at all confident	0 (0%)	6 (20%)	.01	0.68
Somewhat confident	5 (17%)	6 (20%)		
Confident	12 (40%)	9 (30%)		
Very confident	7 (23%)	8 (27%)		
Extremely confident	6 (20%)	1 (3%)		

Note. CPCS = certain prior, certain sample; CPUS = certain prior, uncertain sample; UPCS = uncertain prior, certain sample.

Individual Differences on Accuracy

In line with Experiments 1 and 2, multi-level modeling was used to examine whether individual differences in numeracy, cognitive reflection, thinking dispositions, processing speed, verbal intelligence, and years of education predicted decision accuracy in the RconfB and CconfB situations. The ICC was calculated and showed that approximately 9% of the variance in decision accuracy was due to clustering. In light of the ICC value suggesting at least a portion of the variability of accuracy was due to clustering, a multi-level modeling approach was taken to examine predictors of decision accuracy. Random-intercepts and random-slopes models on accuracy in RconfB and CconfB decision situations, with age included as a Level 2 predictor, were conducted to observe which model best fit the data. Based on the AIC, BIC, and deviance values, the random-slopes model improved the model fit over the random-intercepts model (see Table 21). In light of this, predictor variables were included in separate random-slopes models.

Table 21. *Multi-Level Model Comparison on Accuracy in Experiment 3*

Parameters		Model	
		Random-intercepts	Random-slopes
Fixed effects			
	Intercept	2.31 (0.15)	2.45 (0.18)
	OR	10.05	11.55
	Younger adults	-0.93 (0.19)	-0.89 (0.19)
	OR	0.39	0.41
	RconfB	-1.49 (0.05)	-1.64 (0.19)
	OR	0.23	0.19
	CPCS	0.02 (0.03)	0.02 (0.03)
	OR	1.02	1.02
Fit statistics			
	deviance	11,183	10,596
	AIC	11,193	10,610
	BIC	11,229	10,661
	df	5	7
Model comparison			$\chi^2 = 586.97^{***}$

Note. Standard errors and deviations in fixed and random effects parentheses, respectively; RconfB = Representativeness heuristic conflicts with Bayes' theorem; CPCS = certain prior, certain sample. Younger adults, RconfB and CPCS were used as the reference groups; OR = odds ratios; AIC = Akaike's Information Criterion; BIC = Bayesian Information Criterion; df = degrees of freedom; Subscript χ^2 denotes the two models being compared. * $p < .05$, ** $p < .01$, *** $p < .001$.

The purpose of using multi-level modeling in this experiment was to examine whether assessments of individual differences predict decision accuracy in the RconfB and CconfB decision situations, and whether age influences these effects. Only significant interactions that included the decision situation factor were analyzed further, with nonsignificant results not reported. Interactions observed outside of this were further elucidated with Pearson bivariate correlations.

Lipkus Numeracy Task. A significant Lipkus by Decision Situation interaction was observed ($\beta = -0.31$, $SE = 0.21$; $p = .05$). To unpack this interaction, Lipkus data were centered at the 25th percentile, the median, and the 75th percentile. As Table 22 shows, beta estimates for the effect of decision situation increase as scores on the Lipkus increase. This suggests that the effect of decision situation is strongest for those with lower Lipkus scores. The odds ratios suggest that those with lower Lipkus scores, the odds of a correct response in the RconfB situation are 16% of the odds of a correct response in the CconfB situation. However, those with higher Lipkus scores, the odds of a correct response in the RconfB situation are 23% of the odds of a correct response in the CconfB situation (i.e., those with higher Lipkus scores show less of a gap in accuracy between conflict situations). These results suggest that the effect of decision situation becomes greater with lower scores on the Lipkus.

Table 22. *Lipkus Numeracy as a Predictor of Accuracy in Experiment 3*

	Lipkus centered at 25 th percentile					Lipkus centered at median					Lipkus centered at 75 th percentile				
	β	SE	Z	p	OR	β	SE	Z	p	OR	β	SE	Z	p	OR
Intercept	2.31	0.24	9.54		10.05	2.55	0.20	12.53		12.81	2.67	0.25	10.55		14.45
Lipkus	0.12	0.10	1.18	<.001	1.13	0.12	0.10	1.18	<.001	1.13	0.12	0.10	1.18	<.001	1.13
RconfB	-1.83	0.32	-5.68	<.001	0.16	-1.59	0.27	-5.92	<.001	0.20	-1.47	0.33	-4.41	<.001	0.23
age	-0.84	0.31	-2.73	.01	0.43	-0.71	0.30	-2.37	.02	0.49	-0.64	0.37	-1.74	.08	0.53
CPCS	-0.14	0.09	-1.53	.56	0.87	-0.12	0.09	-1.38	.55	0.89	-0.11	0.11	-0.97	.55	0.90
Lipkus:RconfB	0.12	0.14	0.87	.05	1.13	0.12	0.14	0.87	.05	1.13	0.12	0.14	0.87	.05	1.13
Lipkus:age	0.07	0.13	0.53	.59	1.07	0.07	0.13	0.53	.59	1.07	0.07	0.13	0.53	.60	1.07
RconfB:age	0.33	0.41	0.81	.42	1.40	-0.30	0.39	-0.76	.42	0.74	-0.62	0.48	-1.27	.20	0.54
Lipkus:CPCS	0.01	0.04	0.28	.08	1.01	0.01	0.04	0.28	.08	1.01	0.01	0.04	0.28	.08	1.01
RconfB:CPCS	0.19	0.12	1.65	.44	1.21	0.24	0.10	2.30	.44	1.27	0.26	0.13	1.98	.45	1.30
age:CPCS	0.19	0.11	1.71	.85	1.21	0.19	0.12	1.66	.85	1.21	0.19	0.15	1.31	.85	1.21
Lipkus:RconfB:age	-0.32	0.17	-1.84	.85	0.73	-0.32	0.17	-1.84	.85	0.73	-0.32	0.17	-1.84	.85	0.73
Lipkus:RconfB:CPCS	0.02	0.05	0.46	.50	1.02	0.02	0.05	0.46	.50	1.02	0.02	0.05	0.45	.50	1.02
Lipkus:age:CPCS	-0.00	0.05	-0.00	.87	0.99	-0.00	0.05	-0.00	.87	1.00	-0.00	0.05	-0.01	.87	0.99
RconfB:age:CPCS	-0.27	0.14	-1.91	.03	0.76	-0.29	0.14	-2.00	.03	0.75	-0.29	0.18	-1.63	.03	0.75
Lipkus:RconfB:age:CPCS	-0.01	0.06	-0.12	.91	0.99	-0.01	0.06	-0.12	.91	0.99	-0.01	0.06	-0.11	.91	0.99

Note. Lipkus = Lipkus Numeracy Task; RconfB = Representativeness heuristic conflicts with Bayes' theorem (reference group); CPCS = certain prior, certain sample (reference group); β = beta estimate; SE = standard error; Z = z-score; OR = odds ratio.

To further describe the relationships between decision accuracy and condition, Pearson bivariate correlations were conducted to investigate the linear relationship between Lipkus Numeracy Task scores and decision accuracy within each condition for each age group. For younger adults, significant correlations were observed between Lipkus Numeracy Task scores and the RconfB situation in the certain prior, uncertain sample condition, $r = .35, p = .05$, the uncertain prior, certain sample condition, $r = .46, p = .01$, but did not reach significance for the correlation between accuracy and the certain prior, certain sample condition, $r = .33, p = .07$. Further, no correlations were observed between Lipkus scores and the CconfB situation in any condition for younger adults. For older adults, no correlations were observed between Lipkus scores and the RconfB situation in any condition. However, a significant correlation was observed between Lipkus scores and the CconfB situation in the uncertain prior, certain sample condition, $r = .41, p = .02$. The correlation did not reach significance between accuracy in the CconfB situation and Lipkus scores in the certain prior, certain sample condition, $r = .32, p = .09$, and no correlation was observed between these variables in the certain prior, uncertain sample condition, $r = .19, p = .31$.

Berlin Numeracy Task. A significant Berlin Numeracy Task by Age by Condition ($\beta = 0.06, SE = 0.10; p = .01$), and a significant Decision Situation by Age by Condition ($\beta = -0.29, SE = 0.13; p = .01$), was observed. However, the model with BNT as a predictor of decision accuracy failed to converge.

To shed light on the effects of the interactions, Pearson bivariate correlations were conducted on Berlin Numeracy Task scores and decision accuracy within each decision situation, condition and for each age group. For younger adults, significant correlations were observed between Berlin Numeracy Task scores and the RconfB situation in the certain prior, certain

sample condition, $r = .49$, $p = .01$, the certain prior, uncertain sample condition, $r = .40$, $p = .03$, and in the uncertain prior, certain sample condition, $r = .40$, $p = .03$. Correlations were also observed between Berlin Numeracy Task scores and the CconfB situation in the uncertain prior, certain sample, $r = .44$, $p = .01$, certain prior, certain sample condition, $r = .35$, $p = .05$, and no correlation in the certain prior, uncertain sample condition, $r = .28$, $p = .13$. For older adults, no correlations were observed between Berlin Numeracy Task scores and the RconfB situation in any condition. However, significant correlations were observed between Berlin Numeracy Task scores and accuracy in the CconfB situation in the certain prior, uncertain sample condition, $r = .39$, $p = .03$, the uncertain prior, certain sample condition, $r = .37$, $p = .04$. The correlation did not reach significance between these variables in the certain prior, certain sample condition, $r = .39$, $p = .08$.

Cognitive Reflection Test – Reflection subscale. A significant Cognitive Reflection Test – Reflection subscale (CRT-R) by Decision Situation interaction ($\beta = -0.38$, $SE = 0.23$; $p = .05$) emerged. To unpack this interaction, CRT-R data were centered at the 25th percentile, the median, and the 75th percentile. As Table 23 shows, beta estimates for the effect of decision situation increase as scores on the CRT-R increase. This suggests that the effect of decision situation is strongest for those with lower CRT-R scores. The odds ratios suggest that among those with lower CRT-R scores, the odds of a correct response in the RconfB situation are 10% of the odds of a correct response in the CconfB situation. However, for those with higher CRT-R scores, the odds of a correct response in the RconfB situation are 27% of the odds of a correct response in the CconfB situation (i.e., those with higher CRT-R scores show less of a gap in accuracy between conflict situations). These results suggest that the effect of decision situation becomes greater with lower scores on the CRT-R.

Table 23. *Cognitive Reflection Test - Reflection as a Predictor of Accuracy in Experiment 3*

	CRT-R centered at 25 th percentile					CRT-R centered at median					CRT-R centered at 75 th percentile				
	β	SE	Z	p	OR	β	SE	Z	p	OR	β	SE	Z	p	OR
Intercept	2.13	0.30	6.99		8.39	2.49	0.25	9.89		12.05	3.20	3.31	9.70		24.84
CRT-R	0.36	0.13	2.76	<.001	1.44	0.36	0.13	2.76	<.001	1.44	3.60	1.31	2.75	<.001	1.44
RconfB	-2.35	0.42	-5.64	<.001	0.10	-2.25	0.34	-6.61	<.001	0.11	-2.10	4.16	-4.95	<.001	0.27
age	-1.00	0.39	-2.54	<.001	0.37	-1.09	0.33	-3.31	<.001	0.34	-1.30	4.52	-2.78	<.001	0.28
CPCS	-0.10	0.09	-1.05	.56	0.91	-0.13	0.08	-1.60	.56	0.88	-1.80	1.19	-1.56	.56	0.83
CRT-R:RconfB	-0.38	0.17	0.57	.04	1.10	-0.38	0.17	0.57	.04	1.10	-0.38	1.68	0.57	.04	1.10
CRT-R:age	-0.09	0.18	-0.48	.19	0.92	-0.09	0.18	-0.48	.19	0.92	-0.09	0.18	-0.48	.19	0.92
RconfB:age	1.22	0.55	2.23	.88	3.41	0.84	0.45	1.87	.88	2.33	7.90	5.91	0.14	.89	1.08
CRT-R:CPCS	-0.03	0.05	-0.63	.42	0.97	-0.03	0.05	-0.63	.42	0.97	-2.90	4.59	-0.63	.42	0.97
RconfB:CPCS	0.24	0.12	2.05	.34	1.27	0.24	0.10	2.46	.34	1.27	2.40	1.38	1.75	.34	1.27
age:CPCS	0.16	0.12	1.39	.86	1.18	0.18	0.10	1.76	.86	1.20	2.10	1.53	1.34	.86	1.23
CRT-R:RconfB:age	-0.03	0.24	-1.62	.95	0.68	-0.03	0.23	-1.62	.95	0.68	-0.03	2.36	-1.62	.95	0.68
CRT-R:RconfB:CPCS	0.00	0.05	0.00	.70	1.00	0.00	0.05	0.00	.70	1.00	-5.40	5.42	-0.00	.70	0.99
CRT-R:age:CPCS	0.01	0.06	0.24	.37	1.01	0.01	0.06	0.24	.37	1.01	1.40	5.84	0.24	.37	1.01
RconfB:age:CPCS	-0.33	0.15	-2.19	.02	0.72	-0.31	0.13	-2.42	.02	0.74	-2.60	1.82	-1.44	.02	0.77
CRT-R:RconfB:age:CPCS	0.02	0.07	0.32	.75	1.02	0.02	0.07	0.32	.75	1.02	2.30	7.06	0.32	.75	1.02

Note. CRT-R = Cognitive Reflection Test – Reflection subscale; RconfB = Representativeness heuristic conflicts with Bayes’ theorem (reference group); CPCS = certain prior, certain sample (reference group); β = beta estimate; SE = standard error; Z = z-score; OR = odds ratio.

Pearson bivariate correlations were also run to investigate the linear relationship between CRT-R scores and decision accuracy within each condition for each age group. For younger adults, significant correlations were observed between CRT-R scores and the RconfB situation in the certain prior, certain sample condition, $r = .67, p < .001$, the certain prior, uncertain sample condition, $r = .61, p < .001$, and the uncertain prior, certain sample condition, $r = .65, p < .001$, as well as between CRT-R scores and the CconfB situation in the certain prior, certain sample condition, $r = .45, p = .01$, the certain prior, uncertain sample condition, $r = .53, p = .003$, and the uncertain prior, certain sample condition, $r = .39, p = .04$. For older adults, no correlations were observed between CRT-R scores and the RconfB situation in any conditions, $ps > .05$. However, significant correlations were observed between CRT-R scores and the CconfB situation in the certain prior, certain sample condition, $r = .41, p = .04$, the certain prior, uncertain sample condition, $r = .54, p = .03$, and the uncertain prior, certain sample condition, $r = .33, p = .05$.

Faith in Intuition. A Faith in Intuition by Decision Situation interaction ($\beta = 0.41, SE = 0.26; p = .02$) was observed, as well as a Decision Situation by Age by Condition interaction ($\beta = -0.30, SE = 0.13; p = .02$).

To unpack the Faith in Intuition by Decision Situation interaction, the model was re-estimated three times after centering the Faith in Intuition variable at the 25th percentile, the median, and finally the 75th percentile. As Table 24 shows, beta estimates for the effect of decision situation decrease as scores on the Faith in Intuition increase. This suggests that the effect of decision situation is strongest among those with higher Faith in Intuition scores. The odds ratios suggest that among those with higher Faith in Intuition scores, the odds of a correct response in the RconfB situation are 9% of the odds of a correct response in the CconfB situation. However, among those with lower Faith in Intuition scores, the odds of a correct

response in the RconfB situation are 16% of the odds of a correct response in the CconfB situation (i.e., those with lower Faith in Intuition scores show less of a gap in accuracy between conflict situations). These results suggest that the effect of decision situation becomes greater with higher scores on the Faith in Intuition (see Table 24).

Table 24. *Faith in Intuition as a Predictor of Accuracy in Experiment 3*

	FI centered at 25 th percentile					FI centered at median					FI centered at 75 th percentile				
	β	SE	Z	p	OR	β	SE	Z	p	OR	β	SE	Z	p	OR
Intercept	2.92	0.34	8.63		18.58	2.76	0.26	10.66		15.78	2.61	0.32	8.25		13.61
FI	-0.13	0.17	-0.78	<.001	0.88	-0.13	0.17	-0.78	<.001	0.88	-0.13	0.17	-0.78	<.001	0.88
RconfB	-1.83	0.42	-4.34	<.001	0.16	-2.13	0.32	-6.60	<.001	0.12	-2.40	0.40	-6.08	<.001	0.09
age	-1.20	0.44	-2.73	<.001	0.30	-1.27	0.34	-3.69	<.001	0.28	-1.34	0.41	-3.27	<.001	0.26
CPCS	-0.18	0.11	-1.69	.56	0.84	-0.13	0.08	-1.67	.56	0.88	-0.09	0.10	-0.95	.56	0.91
FI:RconfB	-0.24	0.21	-1.15	.83	0.79	-0.24	0.21	-1.15	.83	0.79	-0.24	0.21	-1.15	.83	0.79
FI:age	-0.06	0.20	-0.28	.14	0.94	-0.06	0.20	-0.28	.14	0.94	-0.06	0.20	-0.28	.14	0.94
RconfB:age	0.11	0.55	0.21	.91	1.12	0.63	0.44	1.45	.90	1.89	1.11	0.52	2.13	.92	3.02
FI:CPCS	0.04	0.05	0.72	.94	1.04	0.04	0.05	0.72	.94	1.04	0.04	0.05	0.72	.94	1.04
RconfB:CPCS	0.28	0.13	2.16	.38	1.32	0.23	0.10	2.39	.38	1.26	0.19	0.12	1.63	.38	1.21
age:CPCS	0.23	0.13	1.72	.87	1.26	0.18	0.10	1.81	.87	1.20	0.14	0.12	1.22	.87	1.16
FI:RconfB:age	0.41	0.26	1.61	.02	1.51	0.41	0.26	1.61	.02	1.51	0.41	0.36	1.61	.02	1.51
FI:RconfB:CPCS	-0.04	0.06	-0.60	.62	0.96	-0.04	0.06	-0.60	.62	0.96	-0.04	0.06	-0.60	.62	0.96
FI:age:CPCS	-0.04	0.06	0.58	.66	0.97	-0.04	0.06	-0.58	.66	0.97	-0.04	0.06	-0.58	.66	0.97
RconfB:age:CPCS	-0.33	0.16	-2.04	.02	0.72	-0.30	0.13	-2.35	.02	0.74	-0.26	0.15	-1.79	.02	0.77
FI:RconfB:age:CPCS	0.03	0.08	0.39	.69	1.03	0.03	0.07	0.40	.69	1.03	0.03	0.08	0.40	.69	1.03

Note. FI = Faith in Intuition; RconfB = Representativeness heuristic conflicts with Bayes' theorem (reference group); CPCS = certain prior, certain sample (reference group); β = beta estimate; SE = standard error; Z = z-score; OR = odds ratio.

The multi-level model with Faith in Intuition as a predictor of decision accuracy failed to converge when probing the Decision Situation by Age by Condition interaction. Thus, Pearson bivariate correlations were conducted on Faith in Intuition scores and decision accuracy within each decision situation, condition, and for each age group. For younger adults, correlations were observed between Faith in Intuition scores and decision accuracy in the RconfB situation in the certain prior, certain sample, $r = -.45, p = .01$, certain prior, uncertain sample, $r = -.40, p = .03$, and uncertain prior, certain sample conditions, $r = -.49, p = .01$. However, no correlations were observed between Faith in Intuition scores and the CconfB situation in any condition. For older adults, Faith in Intuition scores did not correlate with decision accuracy in the RconfB situation, nor in the CconfB situation in any condition ($ps > .05$).

Last, the Cognitive Reflection Test – Intuition subscale, Need for Cognition assessment, Mill Hill Vocabulary Test and years of education did not predict performance on the belief updating task. In addition, the multi-level model with DSST as a predictor of decision accuracy failed to converge. However, to examine whether DSST scores were related to decision accuracy, correlational analyses were conducted. Results showed that scores on the DSST were correlated with decision accuracy in the RconfB situation, $r = .39, p = .002$, and in the CconfB situation, $r = .30, p = .02$. Additionally, correlations between DSST scores and decision accuracy were conducted separately for each age group. Younger adults did not show a correlation in the RconfB situation, $r = .02, p = .92$, or in the CconfB situation, $r = .24, p = .19$. Similarly, older adults did not show a correlation in the RconfB situation, $r = .26, p = .16$, or in the CconfB situation, $r = -.08, p = .69$.

In summary, results from the multi-level models showed that the Lipkus Numeracy Task, the Cognitive Reflection Test - Reflection subscale and Faith in Intuition assessments were good

predictors of decision accuracy in the belief updating task such that higher scores on the Lipkus Numeracy Task and Cognitive Reflection Test - Reflection subscale and lower scores on the Faith in Intuition assessment predicted a smaller gap in decision accuracy between conflict situations (i.e., higher scores in both conflict situations). Further, correlational analyses showed that higher scores on the Lipkus Numeracy Task and Berlin Numeracy Test were associated with higher decision accuracy in the RconfB situation for younger adults and the CconfB situation for older adults, and with higher scores on the DSST correlating with higher scores in both conflict situations. In addition, higher scores on the Cognitive Reflection Test – Reflection, and lower scores on the Faith in Intuition assessments were associated with higher decision accuracy in the conflict situations in all conditions for younger adults only (with Faith in Intuition scores only correlating with accuracy in the RconfB situation).

Summary of Results

Younger adults' decision accuracy was unaffected by condition. However, for older adults, the certain prior, uncertain sample condition produced higher accuracy compared with the uncertain prior, certain sample and the certain prior, certain sample conditions in the RconfB situation, with no difference in accuracy between the latter conditions. Additionally, older adults made more accurate decisions in the uncertain prior, certain sample condition compared with the certain prior, uncertain sample condition, and with no difference between uncertain prior, certain sample and certain prior, certain sample conditions in the CconfB situation. Concerning RT, the RconfB situation produced longer RTs than the CconfB and RalignB situations, and the certain prior, uncertain sample condition produced longer RTs than the uncertain prior, certain sample and the certain prior, certain sample conditions. Decision accuracy between conflict situations was positively correlated in the certain prior, uncertain sample condition for younger adults, and

decision accuracy between conflict situations was negatively correlated in the uncertain prior, certain sample condition for older adults. Concerning the self-assessment, older adults were more likely to report using the “sample” to make decisions in the uncertain prior, certain sample condition, whereas younger adults reported using “both the prior and sample” to make decisions in this condition. In contrast to younger adults, older adults were more likely to report that they felt they made less accurate decisions, felt the task to be more difficult, and were less confident working with numbers in general. Further, more younger adults than older adults explicitly stated that they were able to make correct inferences about prior and sample information that was masked. Additionally, higher scores on the Lipkus Numeracy Task, and Cognitive Reflection Task – Reflection subscale, and lower scores on the Faith in Intuition assessment predicted a smaller gap in decision accuracy between conflict situations (i.e., higher scores in both conflict situations). Finally, higher numeracy scores were associated with higher accuracy in the RconfB situation for younger adults, and the CconfB situation for older adults, with higher DSST scores, collapsed across age, being associated with higher accuracy scores in both conflict situations. Higher Cognitive Reflection Test – Reflection scores were associated with higher accuracy in both conflict situations and lower Faith in Intuition scores correlated with higher accuracy in the RconfB situation in younger adults only.

Discussion

The purpose of this experiment was two-fold. First, I sought to test the hypothesis that older adults' decisions would be influenced by ambiguous information more than younger adults' decisions, such that when prior information was ambiguous older adults would be more likely than younger adults to use the sample to make decisions, and when sample information was ambiguous older adults would be more likely than younger adults to use prior information to make decisions. Second, I sought to examine whether individual differences in numeracy, cognitive reflection, thinking disposition and processing speed would predict decision accuracy. These results would build on the previous experiments by providing further evidence that, beyond age disparities, differences in individuals modulate the accuracy of belief updating when making decisions under risk and ambiguity.

In line with Hypotheses 1 and 2, results showed that younger adults made more accurate decisions than older adults, and with the representativeness error committed more than the conservatism error – findings that replicate Experiments 1 and 2, and support prior literature (Achtziger et al., 2014; Grether, 1980; Peters et al., 2007). Supporting Hypothesis 3, older adults showed more accurate decisions in the certain prior, uncertain sample condition compared with the certain prior, certain sample and uncertain prior, certain sample condition for the RconfB situation, with no difference between the latter conditions, and less accurate decisions in the certain prior, uncertain sample condition compared with the certain prior, certain sample and uncertain prior, certain sample conditions for the CconfB situation, with no differences between the latter conditions. These results demonstrate a preference to make decisions based on certain information, such that when sample information was ambiguous, older adults used prior information to make decisions. These results are consistent with findings showing that older adults have an aversion to ambiguity and rely more strongly on certain information to make choices (Hämmerer et al., 2019; Nassar et al., 2016; Tymula et al., 2013). The current findings build on this literature by demonstrating that older adults

use certain information to make decisions even when provided with negative feedback indicating that this strategy leads to error. These results suggest that older adults do not learn how to make inferences about ambiguous information as younger adults do, at least within the context of belief updating.

In contrast to Hypothesis 4, although the uncertain prior, certain sample condition produced less accurate decisions in the RconfB situation and more accurate decisions in the CconfB situation compared with the certain prior, uncertain sample condition, unexpectedly no differences in decision accuracy emerged between the uncertain prior, certain sample and certain prior, certain sample conditions for older adults in the RconfB situation. This critical finding is a testament to the overweighting of sample information regardless of whether prior information is certain or uncertain. These results are further qualified by most older adults reporting to use prior information to make decisions when sample information was ambiguous (i.e., 53%) and using sample information to make decisions when prior information was ambiguous (i.e., 60%), as well as the significant negative correlation between the conflict situations in the uncertain prior, certain sample condition showing that lower accuracy in the RconfB situation associated with higher accuracy in the CconfB situation.

Consistent with Hypothesis 5, as well as previous research, younger adults showed no difference in decision accuracy in the RconfB and CconfB situations in each condition (Hämmerer et al., 2019; Nassar et al., 2016; Tymula et al., 2013), with the RconfB situation producing less accurate decisions and longer decision times than the CconfB situation overall (Achtziger et al., 2014; Grether, 1980). Younger adults' reports in the self-assessment regarding how they represented masked information suggest that at least some younger adults were able to learn how to make inferences about ambiguous information in order to avoid decision biases. This was further qualified by the positive correlation in decision accuracy between conflict situations in the certain prior, uncertain sample condition, demonstrating that higher accuracy in the RconfB situation was associated with higher accuracy in the CconfB situation in a condition with ambiguous sample information. Additionally, in

line with Hypothesis 6, the results showed that conflict situations produce longer RTs than alignment situations, with older adults producing longer RTs than younger adults.

Together these results complement and extend the judgment and decision making literature. First, replicating Experiments 1 and 2, these results show that the representativeness error is committed more than the conservatism error (Achtziger et al., 2014; Grether, 1980). Second, replicating Experiment 2 results, as well as the aging and reinforcement learning literature, younger adults were able to learn to make accurate decisions more successfully than older adults from feedback (Fernandes et al., 2018). To improve performance, perhaps older adults required more trials to use feedback to learn how to represent ambiguous probabilistic information. This is plausible considering research shows that older adults generally require more trials than younger adults to learn from external feedback (Eppinger et al., 2008; Marschner et al., 2005). It is further supported by research showing that ambiguity is reducible such that decisions made under ambiguity become closer to normative decisions with increasing observations (i.e., the decision maker learns to “feel the weight” of the unknown information through sequential iterations of the task; Ellsberg, 1961; Knight, 1921; Payzan-LeNestour, & Bossaerts, 2011; Volz & Gigerenzer, 2012). Third, the results are in line with the notion that, relative to younger adults, older adults have an aversion to ambiguity (Hämmerer et al., 2019; Nassar et al., 2016; Tymula et al., 2013). Nassar et al., (2016) attributed this aversion to a deficit in making inferences about ambiguous information with age. Results presented here support this argument by showing that younger adults accurately inferred ambiguous information by using feedback to learn how to represent information that was masked. In contrast, older adults did not show evidence of learning to make inferences about masked information to reduce uncertainty and instead avoided ambiguous information even when this strategy led to incorrect decisions.

What cannot be disassociated, however, is what strategy was used to make decisions under ambiguity, particularly amongst older adults. That is, it is unclear whether older adults were

dispositionally more averse to ambiguity (Nassar et al., 2016), whether they were less successful at learning from feedback in order to make inferences (Hämmerer et al., 2019), or both. Past research suggests that it may be a combination of older adults overweighting unambiguous information (Tymula et al., 2013) and not learning from feedback as quickly as younger adults (Eppinger et al., 2008), with the latter finding also supported by the results of Experiment 2. Although younger adults show an aversion to ambiguity (albeit to a lesser degree than older adults; Tymula et al., 2013), younger adults are able to improve decision quality with feedback. Without being able to learn from feedback in Experiment 3, making inferences about masked information may not be possible. In line, with prior work (Eppinger et al., 2008; Marschner et al., 2005), evidence from Experiment 2 also showed that older adults required more trials to learn from feedback compared with younger adults. Given that the number of trials per condition in Experiment 3 was less than the number of trials in Experiment 2, it is possible that older adults required more trials per condition to learn from feedback in order to make inferences about ambiguous information. Future research should utilize EEG to measure the FRN component in order to examine whether older adults gradually begin to use feedback over time to make inferences about ambiguity to improve decision quality if a large number of trials are provided. This avenue of research may elucidate whether ambiguity aversion can be mitigated in older adults given the opportunity to learn from feedback over many trials.

Concerning Hypothesis 7, the Lipkus Numeracy Task, Cognitive Reflection Test - Reflection subscale, and Faith in Intuition assessments predicted performance on the belief updating task. Numeracy and cognitive reflection predicting decision accuracy replicate findings from Experiment 2. These findings indicate that numeracy and cognitive reflection may equip younger and older adults to accurately update beliefs and avoid decision biases even when information is ambiguous. Novel to this experiment however, is that the Faith in Intuition assessment predicted decision accuracy, such that lower scores predicted higher decision accuracy in belief updating. Previous literature has shown that

higher scores on the Faith in Intuition assessment associate with failures in belief updating (Alós-Ferrer & Hügelschäfer, 2012), as well as reliance on heuristic reasoning (Epstein et al., 1996; Shiloh et al., 2002). However, other studies have shown that this assessment does not relate to heuristic-based decision making (Alós-Ferrer et al., 2016; Lu, 2015). That is, the literature is mixed concerning whether intuitive thinking dispositions relate to belief updating. In contrast to Experiments 1 and 2, the current study used the Faith in Intuition assessment to predict accuracy when making decisions under ambiguity rather than risk. The results suggest that having a disposition to make decisions based on intuition relates to committing heuristic-based errors particularly when updating beliefs under ambiguous circumstances. Future research should further examine whether intuitive thinking dispositions associate with avoiding ambiguous information and the overweighting of certain information.

Correlational analyses showed that higher scores on the Lipkus Numeracy Task, Berlin Numeracy Test, Cognitive Reflection Test – Reflection subscale, and lower scores on the Faith in Intuition assessment associated with a greater avoidance of committing the representativeness error in younger adults. In contrast, higher scores on the Lipkus Numeracy Task, Berlin Numeracy Test, and Cognitive Reflection Test – Reflection subscale associated with a greater avoidance of committing the conservatism error in older adults. These results indicate that individual differences in younger and older adults protect against committing different decision biases. Whereas numeracy and cognitive reflection protect against committing the representativeness error in younger adults (with younger adults generally avoiding the conservatism error), for older adults although numeracy and cognitive reflection protect against committing the conservatism error, older adults are still susceptible to the representativeness error. Regardless of age, scores on the DSST correlated with decision accuracy in both conflict situations. These results converge with prior literature showing that measures of fluid intelligence such as processing speed, are strongly associated with Type 2 processing (Burgess et al.,

2011), as well as tasks that tap executive functions such as working memory and inhibition (Kane & Engle, 2002; Gray et al., 2003) - processes important for belief updating when confronted with conflicting information (Stanovich, 2018).

In summary, the results from Experiment 3 replicate and extend the heuristics and biases literature and build on the results of the two previous experiments that comprise this dissertation. First, the results show that younger adults made more accurate decisions than older adults regardless of whether decisions were based on risk or ambiguous information (Hämmerer et al., 2019; Nassar et al., 2016; Peters et al., 2007; Tymula et al., 2013). Second, novel to the current experiment the results provide strong evidence that learning to avoid decision biases when making decisions under ambiguity are influenced by age. Specifically, results showed that the representativeness error was committed more than the conservatism error. However, older adults were more susceptible to the representativeness error compared with younger adults, and only avoided committing this error when sample information was ambiguous. These results suggest that younger adults can learn to make inferences about ambiguous information when feedback is presented in order to update beliefs whereas older adults avoid ambiguity and resort to overweighting the available information that is certain. Third, this is the first study to show that individual differences in numeracy, cognitive reflection, thinking dispositions and processing speed associate with updating beliefs when information is ambiguous.

Chapter 8: General Discussion

Prior literature shows that younger adults deviate from the principles of Bayes' theorem during belief updating by using heuristics that lead to decision biases. Two common errors committed include overweighting prior information (i.e., the conservatism error) and overweighting new information (i.e., the representativeness error; Achtziger et al., 2014; Grether, 1980, 1992). In the heuristics and biases literature, the representativeness error is committed more than the conservatism error (Kahneman & Frederick, 2002). However, this literature has been built on younger adult samples, with limited research investigating how older adults use heuristics during belief updating. In the aging literature, the tendency to process information automatically has been associated with age (Mutter & Pliske, 1994; Yates & Patalano, 1999; although not all research supports this; see Delaney et al., 2015). This finding has been attributed to deficits in cognitive control (Kropotov et al., 2016) and a preference to save cognitive resources in late life (Bruine De Bruin, et al., 2015).

Cognitive reflection tasks such as the belief updating task used in the current dissertation, requires the decision maker to detect a conflict between an automatic and a normative response, suppress and override automatic processes, and simulate alternative responses while simultaneously separating simulations from the real world (Stanovich et al., 2016). Given that the processes underlying cognitive reflection heavily depend on cognitive control (Oldrati, Patricelli, Colombo, & Antonietti, 2016), older adults are likely to be at a disadvantage when performing these tasks compared with younger adults. In addition, belief updating tasks typically do not include feedback (e.g., Kolossa et al., 2015; Kopp et al., 2016). Although older adults show deficits learning from feedback compared with younger adults (Nieuwenhuis et al., 2002; Eppinger et al., 2008), it was not clear how older adults would use heuristics during belief

updating compared with younger adults, or whether younger and older adults would learn to avoid committing heuristic-based decision errors with feedback.

Additionally, belief updating tasks require decision makers to make decisions under risk, in which all information relevant to solving a problem is presented (Volz & Gigerenzer, 2012). However, decisions made under ambiguity are more common in the real world. Although no age differences emerge when making decisions under risk (Armstrong & Spaniol, 2017; Mata et al., 2011), older adults perform worse than younger adults when making decisions under ambiguity, due to an aversion to ambiguous information with age (Hämmerer et al., 2019; Nassar et al., 2016). Whether age differences would emerge in the use of heuristics when updating beliefs based on ambiguous information, or whether younger and older adults would learn to make inferences about ambiguous information from feedback to support belief updating was unknown. Last, previous research shows that individual differences in younger adults such as numeracy, fluid intelligence, thinking dispositions and cognitive reflection associate with performance on tasks that examine the use of heuristics and decision biases (e.g., Alós-Ferrer & Hügelschäfer, 2012; Campitelli & Gerrans, 2014; Cokely et al., in press). Thus, in addition to age differences, it was also of interest to examine whether individual differences in younger and older adults would modulate the use of heuristics when updating beliefs.

The purpose of this dissertation was to fill these gaps in the literature using three experiments to examine: a) how age influences the use of heuristics during belief updating, b) whether the neural correlates that underlie belief updating differ as a function of age, c) whether younger and older adults would learn to avoid committing decision biases during belief updating with feedback, d) how age influences the use of heuristics when making decisions under risk

relative to ambiguity, and e) how individual differences influence the use of heuristics during belief updating.

To briefly restate the main results, younger and older adults committed the representativeness error more than the conservatism error in each experiment, highlighting the prevalence of base-rate neglect (Kahneman & Tversky, 1972). No age differences were observed in the use of heuristics during belief updating with no feedback. However, younger adults made more accurate decisions than older adults during belief updating with feedback. This indicates that decision makers can avoid decision biases from learning from feedback when updating beliefs, with older adults learning to a lesser degree than younger adults. The ERP evidence suggests that older adults made less accurate decisions than younger adults possibly due to a delay in learning from feedback as indexed by a less defined FRN in older relative to younger adults in both blocks of the task. ERP data also showed evidence of conflict detection (indexed by the N2) cognitive control (indexed by pP) and inhibitory control (indexed by the P3b) in response to a conflict between an intuitive lure (i.e., heuristic) and Bayes' theorem in younger and older adults, but with older adults showing more prefrontal neural activity (pP) than younger adults, and younger adults showing more frontocentral neural activity (N2) than older adults.

Additionally, younger adults showed no differences in performance when making decisions under risk relative to ambiguity, with self-assessment data indicating that some younger adults learned to make inferences about ambiguous information to support belief updating. In contrast, decision accuracy and self-assessment data showed that older adults were more likely than younger adults to make decisions based on new information (i.e., the sample) regardless of whether prior information was ambiguous or not, and only made decisions based on prior information when new information was ambiguous. Last, no age differences were observed

in numeracy or cognitive reflection (with lower scores observed on the Cognitive Reflection Test – Reflection subscale), however age differences were observed in processing speed. In all experiments, individual differences in numeracy, processing speed and cognitive reflection¹⁶ predicted or correlated with decision accuracy. The following sections highlight the theoretical and empirical implications of the main results of these experiments.

Prevalence of Base-Rate Neglect

It has been postulated that base-rate neglect occurs because individual-case evidence cues an intuitive response that is difficult to override (Kahneman & Frederick, 2002; Kahneman & Tversky, 1973; Tversky & Kahneman, 1974), whereas base-rate information is more likely to cue some degree of slower, deliberative processing (Bonner & Newell, 2010; De Neys & Glumicic, 2008; Ferreira, Garcia-Marques, Sherman, & Sherman, 2006; Kahneman & Frederick, 2002), with a general preference to forego deliberative, Type 2 processing, for more intuitive, Type 1 processing (Stanovich et al., 2016). This is supported by evidence showing that when the problem structure allows for base-rates to be more easily processed, base-rate neglect becomes less prevalent (Barbey & Sloman, 2007; Brainerd, 2007; Reyna & Mills, 2007).

As a brief reminder, the corrected odds show that the level of difficulty for each decision situation is approximately the same (see Appendix X). Thus, the results do not stem from varying difficulty in the decision situations in terms of their calculability, but rather how information is used to make decisions. Committing the representativeness error, a by-product of base-rate neglect (Kahneman & Tversky, 1972), was commonly observed in each experiment of this dissertation. In the belief updating task, whereas prior information presented probabilistic information (e.g., 75% likelihood the computer drew the sample from the left urn; 25% right

¹⁶ Only Experiments 2 and 3 included the extended Cognitive Reflection Test.

urn), new information was presented in absolute numbers (e.g., 3 blue and 1 green ball drawn) which could be perceived as more concrete evidence than prior probabilities (Barbey & Sloman, 2007). That is, prior and new information differed concerning their level of certainty, with prior information less certain than new information. In line with research showing a preference to process certain relative to uncertain information with little effort (Ellsberg, 1961; Kahneman, 2011; Stanovich, 2011), it is possible that the prevalence of committing the representativeness error observed in Experiments 1 to 3, stems from a preference to base decisions on concrete information that is more easily processed and more difficult to override than putting forth the effort to reason about base-rate probabilities (Bonner & Newell, 2010; De Neys & Glumicic, 2008; Kahneman & Frederick, 2002).

In addition, base-rate neglect may also be due to an aversion to numerical information, such that certain information presented symbolically (the new information) is preferred over less certain information presented numerically (prior information), with the former more easily processed than the latter. That is, it cannot be ruled out that a potential source of base-rate neglect observed in these experiments stem from either the certainty or the format (numeric vs. symbolic) in which information was presented. Importantly however, Experiment 3 did not present any numbers (only symbolic stimuli), with base-rate neglect persisting. This suggests that the difference in certainty between prior and new information may have influenced decision accuracy more than format differences (i.e., numbers vs. symbolic stimuli), however this remains an open question for future research.

Building on the heuristics and biases, as well as the aging literature, a novel contribution of this dissertation series provides evidence that younger and older adults can learn to avoid committing the representativeness error by learning from feedback. Further, the results show that

younger and older adults are susceptible to heuristic-based errors to the same degree during belief updating with no feedback, with age differences emerging when updating beliefs with feedback.

Belief updating without feedback. In contrast to my hypothesis, Experiment 1 showed that younger and older adults committed the representativeness error more than the conservatism error to the same degree. Specifically, without feedback, the data suggests that both age groups were more likely to use new information to make decisions and did not learn to avoid committing heuristic-based decision errors over time. However, because the proportion of representativeness errors approximated 50% for both age groups, it is unclear whether a strategy was used (e.g., using new information to make decisions) or if participants were making decisions randomly. Evidence supporting the former stems from the proportion of correct decisions made in the other situations of interest. That is, in situations in which the conservatism heuristic conflicts with Bayes' theorem or the representativeness heuristic aligns with Bayes' theorem, basing decisions on new information leads to the correct response. Given the higher decision accuracy in these decision situations in Experiment 1 (74% in the CconfB situation and 93% in the RalignB situation) and the lower decision accuracy in situations in which the representativeness heuristic conflicts with Bayes' theorem in Experiment 1 (53%), these results imply that there was a preference to base decisions on new information, regardless of the decision situation, in both age groups. That is, the results could be showing use of normative strategies to make decisions in the situations in which the conservatism heuristic conflicts with Bayes' theorem and the representativeness heuristic aligns with Bayes' theorem. On the other hand, the results may not necessarily reflect avoidance of the conservatism error per se, but rather provides evidence of the high prevalence of base-rate neglect.

Together, these results support literature showing that younger adults deviate from Bayes' theorem by using heuristics that lead to decision errors (Achtziger et al., 2014; Grether, 1980, 1992). Additionally, this is the first study to show that younger and older adults commit heuristic-based errors, particularly the representativeness error, to the same extent in a belief updating task with no feedback and with numeracy modulating and processing speed associating with this outcome. In contrast to Achtziger et al. (2014), decision accuracy did not improve over time. This could potentially be due to methodological differences between studies. In Experiment 1 of this dissertation, the representativeness heuristic may have been more compelling to use because relevant information was presented simultaneously in which new information and the urn distributions could easily be matched (Grether, 1980). In contrast, Achtziger et al. (2014) presented the urn distributions and prior information *before* presenting new information, forcing participants to recall the urn distributions and prior information in memory, which may have gradually encouraged reflective processing over time. In addition, Experiments 2 and 3 of this dissertation showed no age differences in the Cognitive Reflection Test and no age differences in belief updating in Experiment 1. The Cognitive Reflection Test requires one to suppress intuitive lures in order to generate better responses without feedback, similar to the version of the belief updating task that does not include feedback in Experiment 1. Together, these results show that younger and older adults gravitate to decision biases to the same degree when feedback is not provided, with committing the representativeness error being the most prevalent decision bias during belief updating.

Belief updating with feedback. In Experiments 2 and 3, results showed that older adults made more decision errors than younger adults, with the representativeness error committed more than the conservatism error in both age groups. In addition, results in Experiment 2 also

showed an increase in accuracy over time in younger and older adults. These results can be attributed to the inclusion of feedback, and support prior research showing that younger adults learn more successfully from feedback compared with older adults (Eppinger et al., 2008; Fernandes et al., 2018; Nieuwenhuis et al., 2002).

It could be argued that the increase in decision accuracy in Experiment 2 was due to instrumental conditioning such that participants memorized the pairing of decision situation and correct outcome, rather than due to feedback cueing reflective processes. However, if the pairing of decision situation and correct outcome were memorized, one would expect decision accuracy for each decision situation to be high and approximately equal. Importantly, even with trial-by-trial feedback, the representativeness error was committed more than the conservatism error (i.e., a difference in accuracy between decision situations). This indicates that younger and older adults found it difficult to suppress and override the representativeness heuristic in order to choose the option that aligns with a normative response, even when feedback was provided. These results are supported by the longer RTs in situations in which the representativeness relative to the conservatism heuristic conflicted with Bayes' theorem.

Although accuracy increased over time in Experiment 2, there was still evidence of a struggle to make correct decisions when the representativeness heuristic conflicted with Bayes' theorem in the latter half of the task. ERP evidence from Experiment 2 showed that the N2 for younger adults and pP for older adults, markers of conflict detection and engagement of cognitive control (Grützmann et al., 2014; Lucci et al., 2013), increased from Block 1 to Block 2 in response to the representativeness heuristic conflicting with Bayes' theorem. These results provide evidence at the brain level that younger and older adults engaged control processes in response to conflicting information. As described above, these results could be due to the

difficulty overriding concrete evidence (i.e., the sample) in order to generate a normative response that reflects the integration of prior and sample information (Bonner & Newell, 2010; De Neys & Glumicic, 2008; Kahneman & Frederick, 2002).

Additionally, in Experiment 2, ERP evidence in response to feedback showed that older adults produced a less defined FRN component in response to positive and negative feedback compared with younger adults, and with older adults showing a larger FRN component in the second relative to the first block. This suggests that both younger and older adults were able to build expectations about their decisions in order to react to unexpected positive and negative feedback (Eppinger et al., 2008), with older adults showing a delay in building expectations about feedback compared with their younger adult counterparts.

In line with previous research, these results suggest that the age differences observed in decision accuracy may stem from younger adults using feedback more successfully than older adults (Eppinger et al., 2008; Nieuwenhuis et al., 2002), and that older adults require more trials than younger adults to learn from feedback to improve performance (Eppinger, Schuck, Nystrom, & Cohen, 2013). Importantly, a larger FRN component was observed in response to negative feedback when a conservatism error was made relative to a representativeness error. These results suggest that there were greater expectations for outcomes on trials where the conservatism relative to the representativeness heuristic conflicted with Bayes' theorem. Given that more accurate decisions were made when the conservatism relative to the representativeness heuristic conflicted with Bayes' theorem, expectations concerning the outcomes for the former situation may be stronger than the latter. That is, committing a conservatism error may evoke greater FRN in response to negative feedback than committing a representativeness error because making an incorrect decision in the former situation is less likely than the latter situation. These

results converge with the behavioural data showing that the representativeness heuristic was more difficult to avoid than the conservatism heuristic.

An interesting finding that emerged from Experiment 3 was the age difference in committing the representativeness error when prior or sample information was ambiguous. As previously described, older adults compared with younger adults show an aversion to ambiguous information (Hammerer et al., 2019; Nassar et al., 2016; Tymula et al., 2013). Results from Experiment 3 showed that whereas younger adults were not influenced by ambiguous information, older adults showed evidence of making decisions based on new information regardless of whether prior information was ambiguous or not, and avoided the representativeness error only when new information was ambiguous. These results demonstrate the reliance on new information to make decisions, particularly in old age.

A common finding in each experiment was that numeracy and cognitive reflection predicted and processing speed associated with decision accuracy in the belief updating task regardless of whether feedback was presented or not. These results indicate that those higher in numeracy and cognitive reflection were less likely to commit the representativeness error than those lower in numeracy and cognitive reflection, evidence that at least some participants were not “guessing” when the representativeness heuristic conflicted with Bayes’ theorem Experiment 1. Additionally, faster processing speed was associated with a greater avoidance of the representativeness error in all three experiments. These results support past research showing that numeracy, cognitive reflection and fluid abilities protect against committing decision biases in younger adults (Cokely et al., in press; McVay & Kane, 2012; Stanovich et al., 2018), and builds on this literature by extending these results to the older population.

What remains unclear however is whether younger and older adults made more accurate decisions in situations in which the conservatism heuristic conflicted with Bayes' theorem because normative principles were used or because participants defaulted to making decisions based on new information. Analysis of the LRP in Experiment 2 showed evidence that younger and older adults gravitated towards the higher prior probability in situations in which the conservatism heuristic conflicted with Bayes' theorem. Although this shows that participants attended to prior information, this is not evidence of using a normative strategy to make decisions. Therefore, it remains unclear whether younger and older adults used normative strategies to make decisions when the conservatism heuristic conflicted with Bayes' theorem or whether the representativeness heuristic was used.

Together, Experiments 1 to 3 of this dissertation provide behavioural and ERP evidence that sheds light on the prevalence of neglecting the base-rate during belief updating. The results show that younger and older adults can learn to avoid committing this error when updating beliefs if feedback is presented, but with older adults showing a delay in learning from feedback compared with younger adults. Additionally, higher numeracy and fluid abilities may protect against committing the representativeness error regardless of whether feedback is presented or not, and cognitive reflection may protect against base-rate neglect when feedback is provided. Last, the prevalence of committing the representativeness over the conservatism error when updating beliefs may stem from a difference in certainty between prior and new information. However, more research is required to empirically test this speculation.

Cognitive Architecture of Information Processing

A framework of the cognitive architecture of information processing in the heuristics and biases literature has been used to examine miserly processing that leads to decision biases

(Stanovich, 2018). Importantly, this literature is based on younger adult data. Although a recent study examined age differences on the Cognitive Reflection Test (Hertzog, Smith, & Ariel, 2017), to my knowledge no study has investigated information processing that underlies performance on a belief updating task that pits an error-prone heuristic against Bayes' theorem in younger and older adults. Of note, the purpose of this dissertation was not to directly test the processes that underlie each stage of the information processing framework in younger and older adults. However, the data acquired across these experiments in combination with what is known about cognitive aging may provide insight into the age differences observed in belief updating. The point of this section was to provide a "first look" at how miserly processing in older adults differs from younger adults by referencing relevant data acquired from these experiments, as well as the cognitive aging literature and to relate these findings to the information processing framework of cognitive reflection.

As a brief reminder, within the heuristics and biases literature, the cognitive architecture of information processing framework consists of three processing stages: 1) detecting the inadequacy of an automatic response; 2) suppressing and overriding automatic processing to engage reflective processing; 3) generating hypothetical alternative responses that are superior to the automatic response, while sustaining decoupling operations (Stanovich et al., 2016). Previous work shows that individual differences on the Cognitive Reflection Test have been associated with Stage 1 (Stanovich et al., 2016), inhibitory control has been associated with Stage 2, fluid intelligence has been associated with Stage 3 (Stanovich, 2011), stored knowledge such as numeracy has been associated with Stages 1 and 3 (Liberali, Reyna, Furlan, Stein, & Pardo, 2012; Stanovich, 2018; Stanovich et al., 2016), and thinking dispositions have been associated with Stages 1 and 2 of the framework (Stanovich & West, 1998). Given that older adults show

deficits in inhibitory control (Hasher & Zacks, 1988), and fluid intelligence (Salthouse, 2019), show lower numeracy and a greater preference for automatic processing than younger adults (Peters et al., 2007; Peters et al., 2006), older adults are likely at a disadvantage when performing tasks requiring cognitive reflection compared with younger adults. However, results from the experiments in this dissertation did not show consistently worse performance in older relative to younger adults. Rather, results showed that younger and older adults deviate from normative decision making to the same degree. Increases in normative responses were only observed when cued by feedback, with older adults deviating from normative responses more than younger adults even with feedback.

Conflict detection, cognitive control, and alternative responses. Previous research has used RTs to gain insight into the difficulty of a task, with longer RTs reflecting greater difficulty (e.g., making decisions based on conflicting information) and shorter RTs reflecting less difficulty (e.g., making decisions based on non-conflicting information; Alós-Ferrer et al., 2016). Although RTs do not map one-to-one on latent processes, they may provide insight into the difficulty of a task. Across all experiments younger and older adults showed longer RTs in conflict situations than alignment situations. This indicates that even without feedback, conflict situations were more difficult than alignment situations, which may suggest that conflicting information engaged more reflective processing than non-conflicting information.

ERP evidence showed that during belief updating with feedback, younger and older adults detected conflict and engaged cognitive control processes, indexed by the N2 component in younger adults and pP in a later time window in older adults. The N2 component has been used as a marker of conflict detection (e.g., Donkers & van Boxtel, 2004) and the over-recruitment of the PFC in older adults has been associated with cognitive control in the conflict

monitoring literature (such as in Go/No-Go tasks; Lucci et al., 2013). Engagement of cognitive control implies that further information processing was necessary because the information presented did not elicit an immediately clear response. Additionally, more defined N2 and pP components were observed in Block 2 relative to Block 1, with larger N2 and pP amplitudes correlating with higher decision accuracy in younger and older adults, respectively. These results suggest that learning from feedback may have cued engagement of reflective processes in both age groups. Thus, concerning Stage 1 and 2 of the information processing framework, the RT and ERP data provide evidence that control processes were engaged in response to a heuristic conflicting with Bayes' theorem, with larger amplitudes associating with normative responses in both age groups.

Generating an alternative response that is superior to an automatic response, while simultaneously sustaining decoupling operations relies heavily on fluid abilities (Stanovich, 2011). Age differences in processing speed, an index of fluid intelligence (Salthouse, 2019), was observed in all three experiments, with older adults performing worse than younger adults. Additionally, in each experiment, correlational analyses showed that higher processing speed associated with higher decision accuracy. However, no associations between fluid intelligence and decision accuracy were generally observed when the correlations were conducted separately within each age group. These results suggest that diminished processing speed in older adults compared with younger adults may have contributed to the age differences observed on the belief updating task, with Stage 3 of the framework reflecting the process that depends most heavily on fluid intelligence. Although age differences observed in Experiments 2 and 3 may be driven by younger adults learning from feedback more successfully than older adults (as indexed by a diminished FRN in older adults; Eppinger et al., 2008), research has shown that fluid intelligence

underlies learning from feedback (Ferdinand et al., 2016). Of note, age and processing speed are highly correlated. Therefore, the relation between processing speed and decision accuracy may reflect the relation between age group and decision accuracy. Future research should further investigate whether diminished fluid abilities with age influence generating normative responses and sustaining decoupling operations during cognitive reflection or learning from feedback when updating beliefs.

Numeracy, cognitive reflection and thinking dispositions. As previously described, individual differences in numeracy, cognitive reflection, and thinking dispositions have been associated with decision biases (Cokely et al., in press; Toplak et al., 2014b; Toplak et al., 2011). In all three experiments of this dissertation, no age differences were observed on the numeracy assessments or the Cognitive Reflection Test¹⁷, Need for Cognition or Faith in Intuition assessments, however numeracy and cognitive reflection predicted decision accuracy, regardless of age. These results build on the current literature by showing that having a proclivity for numeracy and reflective processing underlies performance on heuristics and biases tasks regardless of age (Cokely et al., in press; Toplak et al., 2014b; Toplak et al., 2011).

With the exception of higher faith in intuition predicting lower decision accuracy in Experiment 3, across all experiments thinking dispositions (i.e., Need for Cognition and Faith in Intuition) did not predict decision accuracy. These results could be due to a difference in what is measured by thinking disposition assessments and the belief updating task measure. Whereas thinking disposition assessments reflect one's perception of how one makes decisions (i.e., self-report responses), the belief updating task reflects decision behaviour. Self-reports of how one

¹⁷ The Cognitive Reflection Test was only included in Experiments 2 and 3.

believes they make decisions (i.e., based on analytical vs. intuitive processing) did not associate with how decisions were actually made in the belief updating task.

Together, these results suggest that diminished fluid abilities rather than a deficit in inhibitory control may contribute to the age differences observed in the belief updating task. Additionally, numeracy and cognitive reflection consistently predicted decision accuracy in each experiment in younger and older adults, suggesting that these cognitive domains may protect against decision biases when updating beliefs. These results are the first to shed light on the processes that underlie cognitive reflection in old age and how this data can be situated in the cognitive architecture of information processing framework. Importantly, these data do not provide direct evidence of different cognitive processes or abilities influencing specific stages of the framework. However, the results in combination with well known theories of cognitive aging provide a preliminary examination concerning which stages age differences in belief updating stem from. An avenue for future research is to investigate at which stages in the information processing framework older adults differ from younger adults. Specifically, future research should examine how diminished inhibitory control and fluid abilities in old age effects overriding automatic responses and generating alternative responses while sustaining decoupling operations, respectively. This may further elucidate why older adults use heuristics and gravitate towards decision biases more than younger adults and how this behaviour may implicate decisions in the real world.

Decisions Under Risk vs. Ambiguity

Previous literature shows that older adults perform similarly to younger adults when making decisions based on risk information (Armstrong & Spaniol, 2017; Mata et al., 2011). However, when making decisions based on ambiguous information older adults show an

aversion to ambiguity compared with younger adults (Tymula et al., 2013). Specifically, older adults have difficulty making inferences about ambiguous information (Hämmerer et al., 2019) and tend to devalue information that is uncertain and overvalue information that is certain (Nassar et al., 2016) compared with younger adults.

The reason for this age difference has been attributed to the different neural structures and cognitive mechanisms that underpin decisions under risk and ambiguity (Brand et al., 2006). Specifically, whereas the “cognitive” loop supports holding and manipulating information on-line required for making decisions under risk, the “limbic” loop supports feedback and reward-based processing as information is gradually learned which is required for making decisions under ambiguity. Experiment 1 and 2 of this dissertation required making decisions under risk, with age differences observed in the latter but not the former. Whereas making normative decisions in Experiment 1 likely depended on the cognitive loop such that decisions were made based on explicit risk information, Experiment 2 may have depended more on the limbic loop such that decisions were improved by learning from feedback, with older adults less likely to learn from feedback than younger adults (Eppinger et al., 2008).

Results from Experiment 3 support prior research showing that older adults are less likely to learn to infer about ambiguous information by learning from feedback compared with younger adults (Hämmerer et al., 2019). Research shows that when making decisions under ambiguity, initially decisions are made under complete uncertainty. However, as ambiguous information is gradually figured out through learning from feedback, decisions under ambiguity start to reflect decisions under risk (similar to Experiment 2; Nassar et al., 2016). In both age groups, avoiding the representativeness error was the most difficult regardless of condition. However, younger adults were able to infer ambiguous information by learning from feedback

more than older adults particularly in the situation in which the representativeness heuristic conflicted with Bayes' theorem.

Critically in Experiment 3, older adults showed no differences in decision accuracy between the uncertain prior, certain sample and certain prior, certain sample conditions. In these conditions, accuracy was lower in situations in which the representativeness heuristic conflicted with Bayes' theorem and higher in situations in which the conservatism heuristic conflicted with Bayes' theorem. This suggests that older adults used new information to make decisions in these conditions. Conversely, when new information was ambiguous (i.e., the certain prior, uncertain sample condition), accuracy was higher in situations in which the representativeness heuristic conflicted with Bayes' theorem and lower in situations in which the conservatism heuristic conflicted with Bayes' theorem, compared with the other two conditions. Together, these results suggest that older adults used new information to make decisions regardless of whether prior information was ambiguous, and only used prior information when new information was ambiguous.

These results show that older adults avoid ambiguous information more than younger adults. In addition, the preference to rely more on information of greater relative to less certainty was observed throughout the dissertation, such that most decisions were made based on new information rather than prior probabilities. Although both younger and older adults showed evidence of this, given feedback younger adults learned to incorporate prior information into decisions more than older adults. These results converge with literature showing age differences when making decisions that rely on cognitive mechanisms supported by the limbic loop (Brand et al., 2006), with age differences observed in Experiments 2 and 3 likely due to a diminished ability to learn from feedback with age.

Last, numeracy and cognitive reflection predicted decision accuracy in Experiment 3. Past research has shown that numeracy is a good predictor of decisions under ambiguity in younger adults (Cokely et al., in press). However, to my knowledge this is the first study to show that numeracy underlies decisions under ambiguity in the older population, and that cognitive reflection, as measured by the extended Cognitive Reflection Test (Toplak et al., 2014a), predicts accuracy when making decisions under ambiguity in both younger and older adults. Correlational analyses showed that higher numeracy and cognitive reflection scores in younger adults associated with the situation in which the representativeness heuristic conflicted with Bayes' theorem in conditions where information was ambiguous. Older adults showed a similar pattern when making decisions under ambiguity specifically when the conservatism heuristic conflicted with Bayes' theorem. Individual differences in younger and older adults correlating with different conflict situations has been observed in each experiment, possibly due to a difference in conflict situation difficulty – not in the sense that computation differs in difficulty across decision situations but rather that older adults may have more difficulty avoiding the representativeness error than younger adults. Building on the previous experiments however, these results show that numeracy and cognitive reflection are also good predictors of belief updating about ambiguous information.

In summary, the results from these experiments show that older adults avoid ambiguous information more than younger adults, with younger adults learning to make inferences about ambiguous information and incorporate less certain information to improve decisions to a greater degree than older adults. Given that most decisions made in the real world are based on uncertain information, these results have important implications for how decisions are made in real life. For example, older adults may decide against taking medications or opting for surgery when the

prevalence of a disease is low and the relevant diagnostic test for testing the presence of disease is not 100% accurate. Future studies should further investigate whether age differences in decisions under ambiguity are derived from an aversion to ambiguous information, or a deficit in learning from feedback in old age, or some combination of the two.

Limitations

The experiments that comprise this dissertation are not without limitations. Achtziger et al. (2014) presented prior information separately from new information, with participants making a decision when only the sample was presented on the screen. It is possible that results reported by Achtziger et al. (2014) could have been confounded by a recency effect, such that new information influenced decisions more because it was on the screen at the time participants made decisions, increasing accessibility of new information and only accessing prior information through memory. In order to avoid a memory-related confound and isolate reasoning, the current experiments presented this information in the same order (i.e., urn distributions then prior then new information) but with all information remaining on the screen when decisions were made. Although participants do not need to hold information in memory, new information was presented last which could have made it more salient or accessible (although participants could look up to the prior information making that the most recent information observed). Therefore, although mnemonic demands may have been reduced, whether all information remained on the screen after sequential presentation eliminated the confound of recency is unclear.

A second limitation is that correct decisions in situations where the conservatism heuristic conflicted with Bayes' theorem may reflect base-rate neglect rather than the avoidance of a conservatism error. In light of this, base-rate neglect may have been a strategy applied across all decision situations. That is, although base-rate neglect was discussed as the most common

error committed by younger and older adults, it is possible that it was the only heuristic used across experiments.

A third limitation is that it cannot be argued with certainty that increases in decision accuracy over time are due to learning from feedback. That is, the increase in decision accuracy over time could be due to instrumental conditioning in which outcomes are learned based on the reinforcing and inhibiting effects of feedback. In addition, increases in accuracy could be due to practice effects (i.e., familiarity with the task). However, in Experiments 2 and 3, the experimenter asked participants whether they were able to memorize the correct answers to decision situations, and all participants responded that there were too many decision situations to memorize answers (although participants did express an awareness of decision situations repeating). It is possible that participants could have memorized decision situations and the corresponding answer, such that decision accuracy reflects recognition memory rather than reasoning ability. However, if participants memorized decision situations and their outcomes, one would expect approximately equal accuracy in each decision situation. Given that accuracy fluctuated in each decision situation and held the same pattern across all experiments regardless of whether feedback was presented or not, participants may not have made decisions based on decision situation and outcome pairings in memory.

Last, prior and new information may have been perceived differently in the belief updating tasks. That is, prior information may have been perceived as less certain because it reflects a probability distribution, whereas new information may have been perceived as more certain because it reflects absolute numbers. The difference in “certainty” of the prior and new information may have confounded results, such that this difference contributed to why base-rate neglect was so prevalent across these experiments.

Future Directions

In order to determine whether order of prior and new information influences how heuristics are used, or to rule out a potential recency effect confound, future studies could utilize eye tracking technology or think-aloud procedures to determine what information is attended to before making a decision, or counterbalance the presentation order of prior and new information in way that would make sense in the task. Further, equating the certainty of prior and new information would also be important to investigate in the future to determine whether the difference in certainty influenced the use of heuristics in these experiments, specifically the prevalence of base-rate neglect.

In addition, learning from feedback over time was not examined in Experiment 3 because each condition did not consist of enough trials to create trial blocks. Thus, it remains an open question whether accuracy in decisions under risk compared with decisions under ambiguity increase in younger and older adults during belief updating with feedback. Future research should examine the rate of learning when updating beliefs when relevant information fluctuates in certainty and determine whether older adults can learn to make inferences about ambiguous information given ample opportunity (i.e., more trials than what was provided in Experiment 3).

In each experiment, participants completed assessments subsequent to performing the hour-long belief updating task. It is possible that participants could be cognitively depleted after performing the belief updating task and the scores on the assessments may not reflect younger and older adults' typical performance (although age differences in some of the assessments such as the Mill Hill Vocabulary Test and the DSST are well documented in the aging literature; Salthouse, 2019). Future research should consider splitting experiments like these into two sessions, in which one session participants complete the belief updating task and another session

requires participants to complete the assessments (with sessions counterbalanced). Designing the experiment this way may help avoid depletion effects.

Last, building on the main results of this dissertation, future research should further examine age differences in cognitive reflection tasks in the heuristics and biases literature. Specifically, do older adults perform worse than younger adults during belief updating with feedback because they do not learn from feedback as well as younger adults, or due to diminished fluid abilities with age, important for generating normative responses once non-normative responses have been suppressed, or a combination of both? This would further highlight how age influences the processes that underpin cognitive reflection.

Conclusions

In conclusion, the findings of this dissertation have important implications for theories surrounding the use of heuristics, decision biases, and aging. Specifically, these experiments demonstrated that both younger and older adults commonly use the representativeness heuristic leading to decision errors when updating beliefs. However, given an opportunity to learn from feedback, both age groups learned to avoid making heuristic-based decision errors, with younger adults learning to a greater degree than older adults. These data are the first to show that error-prone heuristics, particularly base-rate neglect, can be avoided with feedback, with older adults showing a delay in learning from feedback compared with younger adults – findings that converge with prior research that show a deficit in learning from feedback with age (Eppinger et al., 2008). In addition, the results further our understanding of how the neural correlates associated with belief updating are influenced by age, such that both younger and older adults showed evidence of conflict detection and engagement of cognitive control in response to conflicting information. However, whereas younger adults showed typical frontocentral activity

in response to conflict, older adults showed less frontocentral activity and more prefrontal activity than younger adults. These results are the first to show that older adults engage more anterior brain regions than younger adults, possibly to compensate for age-related declines in other brain regions, when updating beliefs.

Further, a novel finding observed was that ambiguous information influenced the use of heuristics in older adults, but not younger adults. These results have important real world implications that highlight susceptibility to decision biases in old age, specifically when information is not certain. Additionally, a consistent finding throughout the experiments was that individual differences in numeracy, processing speed, and cognitive reflection were associated with belief updating performance. These results are the first to show that higher numeracy, fluid abilities, and cognitive reflection in both young and old age protect against gravitating toward decision biases when updating beliefs.

These results also build on the heuristics and biases literature by shedding new light on how the processes that underpin cognitive reflection in old age relate to the stages that comprise the information processing framework (Stanovich et al., 2016). Specifically, given that generating normative responses while suppressing automatic non-normative responses, and learning from feedback depend heavily on fluid abilities, age differences observed in belief updating may have been driven by diminished fluid intelligence. An avenue for future research is to pinpoint at which stages in the process of cognitive reflection younger and older adults show differences. These results would further elucidate why automatic processing and the use of heuristics increase with age. Altogether, these findings add to our knowledge of how younger and older adults update beliefs about risk and ambiguous information, and show how avoidance of decision biases is possible given the opportunity to learn from previous decision errors.

Appendices

Appendix I: Consent Form for Experiment 1

CONSENT FORM

DECISION GAME I

You are being invited to participate in a research study. Please read this consent form so that you understand what your participation will involve. Before you agree to participate, please ask any questions you may have.

Investigators:

Principal Investigator: Bonnie Armstrong, Department of Psychology, Ryerson University, bonnie.armstrong@psych.ryerson.ca

Faculty Supervisor and Co-Investigator: Dr. Julia Spaniol, Department of Psychology, Ryerson University
jspaniol@psych.ryerson.ca

This study is being conducted in partial fulfillment of the doctoral degree requirements for Ms. Bonnie Armstrong. The study is funded by a grant from the Natural Sciences and Engineering Research Council.

Purpose of Study: In this study we are investigating how people make decisions based on probability information. Approximately 80 individuals will participate in this study. Results of this experimental study will be examined at the group level. We will not examine individual results. The group results may be published or presented at conferences.

Description of Study: The study involves 1 testing session at Ryerson University. The session will involve a two computer-based tasks, as well as paper-and-pencil questionnaires. The time to complete the session is approximately 1.5 hours, and will not exceed 2 hours. Participants will receive an incentive of \$15 for participating, as well as an opportunity to win an additional \$5 that will be based on your performance on the Decision Game computer task described below. If you choose to participate, your identity will be kept confidential and your responses will be recorded in such a way that you cannot be identified. You will receive a copy of this consent form, which contains the names and contact numbers of the study investigators.

Tasks: The session will involve paper-and-pencil questionnaires and two computer tasks. You will be asked questions pertaining to various personality traits; dispositions and preferences about certain statements and in given situations; and overall mood and feelings. You will also complete some simple tests of thinking and vocabulary. The first computer task will present coloured words to you. Your task is to identify the font colour of the word. In the Decision Game computer task, you will be presented with two jars that each contain blue and green balls. The

computer will draw a sample of balls from one of the jars. You will be given information regarding the likelihood that each jar was chosen. You will then be shown the sample of balls that the computer drew, and you will be asked to guess which jar the balls were drawn from. The goal is to guess correctly as often as possible.

Confidentiality: Each individual's results are confidential. Neither your identity nor any personal information will be available to anyone other than the investigator. Each participant will receive an identification number (ID). Only the ID number will appear on electronic files, data files as well as on hard-copy data forms and questionnaires. Only individuals involved in the research team will have access to a central password-protected electronic file that matches participant IDs with identifying information (name, date of birth, responses to demographic and health-related questions). This information is being kept in the event that a follow-up study is conducted in the future and if you acknowledge and agree to this use of your information, as well as to be contacted in the future. If a follow-up study is conducted in the future and you would like to be contacted and be included in the follow-up study, the investigator will contact you in order to obtain your consent to use this information. No personal information will be disclosed in any resulting publication or presentation. All research records will be stored in a locked file cabinet in a locked laboratory (with personal information such as your name and signature stored separate from all other information). Electronic files will be password-protected and will be stored locally on a secure server on the Ryerson network.

If you are interested, we would be happy to provide you with the final results of the study when they appear in press via email. After the study has been published or presented, or after a maximum of 7 years, all paper records (form, questionnaires, etc.) we collect from you will be shredded.

Risks and Benefits: The results of this study may help us better understand how people make decisions that involve probability. Your participation will not directly benefit you, but knowledge will be gained that may benefit others. Furthermore, you may find the study fun and educational. However, being involved in a study may be a new experience for you, and you may feel uncomfortable answering some of the questions, some of which you may find disturbing or upsetting. However, any potential discomfort is expected to be temporary and will not be greater than what you might experience in a typical day. Remember that you do not have to answer all questions and that you may discontinue participation at any time with no penalty or loss of benefits to which you are entitled.

Incentives to participate: Participants will receive an incentive of \$15 for participating. You will receive this incentive even if you do not complete the study for any reason. Participants will receive an additional \$5 if 80% or more of the decisions made in the Decision Game computer task are correct.

Voluntary nature of participation: Participation in this study is voluntary. Your choice to participate, or not to participate, will not influence your future relations with Ryerson University. If you decide to participate, you are free to withdraw your consent and to stop your participation at any time, before or during the experiment, without penalty or loss of benefits to which you are

entitled. If closed, no data will be collected. If you do, you will still receive \$15 for your participation. If you would like to participate but wish your data to be excluded from the analyses you can request this at any point during the study without penalty or loss of benefits to which you are entitled. You must notify the researchers by April 1, 2018 if you wish to have your data removed from the study.

Questions about the study: If you have any questions about the research experience please feel free to ask the experimenter at any time (now or during the study). If you have any questions about this research after you have completed the study you may contact Bonnie Armstrong (phone: 416.979.5000 ext. 2193, or bonnie.armstrong@psych.ryerson.ca) or Dr. Julia Spaniol (416.979.5000 ext. 2268, or jspaniol@psych.ryerson.ca).

Questions about your rights as a participant: This research has been reviewed and has received ethical approval by the Research Ethics Committee at Ryerson University. If you have questions about your participation in the research or about your treatment as a research participant, you may contact the Research Ethics Board (Contact: Research Ethics Board, c/o Office of the Vice President, Research and Innovation, Ryerson University, 350 Victoria St., Toronto ON, M5B 2K3, 416.979.5000, ext. 5042, or fax 416.979.5336).

Agreement:

Your signature below indicates that you have read the information in this agreement and have had a chance to ask any questions you have about the study. Your signature also indicates that you agree to be in the study and have been told that you can change your mind and withdraw your consent to participate at any time. You have been given a copy of this agreement. You have been told that by signing this consent agreement you are not giving up any of your legal rights.

Name of Participant (please print)

Signature of Participant

Date

If you would like to be contacted in the future in the event that the experimenter conducts a follow-up study, please sign and date below.

Signature of Participant

Date

☐ Please check this box if you would like to receive a copy of the final report of this research

Signature of Investigator

Date

Appendix II: Debriefing Form for Experiment 1

DEBRIEFING FORM DECISION GAME I

Purpose of Study:

If you were to decide whether to hold an afternoon wedding ceremony outdoors at 6pm on the basis of a) the morning weather forecast and b) actual weather conditions at 2pm, you would need to combine this information in such a way that would inform your decision. People have great difficulty “updating” their knowledge with new information to make a decision¹ (e.g., updating one’s knowledge of the weather forecast announced in the morning, with actual weather conditions in the early afternoon in order to decide whether to have the wedding indoors or outdoors at 6pm). People have a tendency to either place more value on their prior knowledge (e.g., the morning weather forecast) in which a *conservatism* heuristic is used or place more value on new information (e.g., weather conditions at 2pm) in which a *representativeness* heuristic is used². Previous research from our lab examined how younger and older adults combine old information with new information, and found that both age groups were able to combine information more accurately when information was experienced, compared to when information was described to them³. However, the cognitive processes underlying younger and older adults’ performance in information integration tasks are still unknown. In particular, it is unclear whether there are age differences in the use of heuristics (e.g., do younger adults use heuristics more or less than older adults?). In addition, it is unknown whether there are age differences in the type of heuristic used (e.g., do younger adults prefer to use a different heuristic than older adults do?). The goal of the current study is to reveal whether age differences exist when old and new information are to be combined, and decisions are to be made based on judgments of the information.

Main Question:

a) Are there age differences in the use of heuristics? b) Are there age differences in the type of heuristic used?

Study Summary:

You made many decisions in which you had to decide which urn a sample of blue and green balls were drawn from. Some of the decision situations you were faced with a) influenced using a heuristic that led to an incorrect decision, b) influenced using a heuristic that led to a correct decision, and c) did not influence any heuristic. Our main interest is whether you used a heuristic that would lead you to make an incorrect choice, or whether you used a different strategy that would lead you to the correct choice.

Study hypothesis:

We hypothesize that age differences exist regarding the frequency of heuristics used, with older adults using heuristics more than younger adults.

Why is this study important?

This study will tell us how susceptible older and younger adults are to heuristics (short-cuts) that lead to incorrect choices in a probability-based task. It will also shed light on whether older and/or younger adults prefer one type of heuristic to another.

Thank you again for your participation! Please contact bonnie.armstrong@psych.ryerson.ca if you have further questions about this study.

Further readings:

1. Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. *Science*. 1974;185:1124-31.
2. Achtziger A, Alos-Ferrer C, Hugschafer S, Steinhäuser M. The neural basis of belief updating and rational decision making. *Soc Cogn Affect Neurosci*. 2014;9:55-62.
3. Armstrong BA, Spaniol J. Experienced probabilities increase understanding of diagnostic test results in younger and older adults. *Med Decis Mak*. 2017; 1-10.

Appendix III: Consent Form for Experiment 2

CONSENT FORM DECISION GAME II

You are being invited to participate in a research study. Please read this consent form so that you understand what your participation will involve. Before you agree to participate, please ask any questions you may have.

Investigators:

Principal Investigator:	Bonnie Armstrong, Department of Psychology, Ryerson University, bonnie.armstrong@psych.ryerson.ca
Faculty Supervisor and Co-Investigator:	Dr. Julia Spaniol, Department of Psychology, Ryerson University, jspaniol@psych.ryerson.ca

This study is being conducted in partial fulfillment of the doctoral degree requirements for Bonnie Armstrong. The study is funded by a grant from the Natural Sciences and Engineering Research Council.

Purpose of Study: In this study we are investigating how people make decisions based on probability information. The results may be presented at conferences and/or published in an academic journal.

Description of Study: The study takes a total of 3 hours to complete at Ryerson University. All participants will undergo electroencephalographic (EEG) recordings during the study. Prior to beginning the study, the experimenter will explain how EEG is recorded, and the details of the experimental tasks. The session will involve two computer-based tasks, as well as paper-and-pencil questionnaires. Participants will receive an incentive of \$30 for participating, as well as an opportunity to win an additional \$5 that will be based on your performance on the Decision Game computer task described below. In accordance with the Personal Health Information Protection Act (PHIPA), if you choose to participate, your identity will be kept confidential and your responses will be recorded in such a way that you cannot be identified. You will receive a copy of this consent form, which contains the names and contact numbers of the study investigators.

Tasks: The session will involve paper-and-pencil questionnaires and two computer tasks. In the questionnaires, you will be asked questions pertaining to various personality traits; dispositions and preferences about certain statements and in given situations; and overall mood and feelings. You will also complete some simple tests of thinking and vocabulary. The first computer task will require you to press keys on the keyboard in response to stimuli on the computer screen. In the Decision Game computer task you will be presented with two jars that each contain blue and

green balls. The computer will draw a sample of balls from one of the jars. You will be given information regarding the likelihood that each jar was chosen. You will then be shown the sample of balls the computer drew, and you will be asked to guess which jar the balls were drawn from. The goal is to guess correctly as often as possible. You will receive feedback regarding whether you were correct or incorrect subsequent to each decision.

Confidentiality: Each individual's results are confidential. Neither your identity nor any personal information will be available to anyone other than the investigator. Each participant will receive an identification number (ID). Only the ID number will appear on electronic files, data files as well as on hard-copy data forms and questionnaires. Only individuals involved in the research team will have access to a central password-protected electronic file that matches participant IDs with identifying information (name, date of birth, responses to demographic and health-related questions). This information is being kept in the event that a follow-up study is conducted in the future and if you acknowledge and agree to this use of your information, as well as to be contacted in the future. If a follow-up study is conducted in the future and you would like to be contacted and be included in the follow-up study, the investigator will contact you in order to obtain your consent to use this information. No personal information will be disclosed in any resulting publication or presentation. All research records will be stored in a locked file cabinet in a locked laboratory (with personal information such as your name and signature stored separate from all other information). Electronic files will be password-protected and will be stored locally on a secure server on the Ryerson network. You have the right to withhold any personal health information if you wish to. You also have the right to discontinue your participation in this study at any point without any negative consequences.

If you are interested, we would be happy to provide you with the final results of the study when they appear in press via email. After the study has been published or presented, or after a maximum of 7 years, all paper records (form, questionnaires, etc.) we collect from you will be shredded.

Risks and Benefits: The results of this study may help us better understand how people make decisions that involve probability. The anticipated benefits of participating in this study include: a) gaining knowledge of psychological experiments and cognitive functions; and 2) contributing to the research of cognitive psychology. Participation in this study is associated with no or minimal risks. Furthermore, you may find the study fun and educational. However, being involved in a study may be a new experience for you, and you may feel uncomfortable answering some of the questions, some of which you may find disturbing or upsetting. However, any potential discomfort is expected to be temporary and will not be greater than what you might experience in a typical day. Remember that you do not have to answer all questions and that you may discontinue participation at any time with no penalty or loss of benefits to which you are entitled. The experimenter will make every effort to ensure that you feel as comfortable as possible throughout the session. There are no long-term risks associated with the recording of EEG, although you might feel short-term discomfort as a result of wearing the electrode cap for a long period of time. **If you have temporal-mandibular joint (TMJ) disease or any recurrent**

problems with your head or neck, then you should not participate in this study. As a reminder, you can discontinue the study at any time without penalty.

Incentives to participate: Participants will receive an incentive of \$30 for participating. You will receive this incentive even if you do not complete the study for any reason. Participants will receive an additional \$5 if 80% or more of the decisions made in the Decision Game computer task are correct.

Voluntary nature of participation: Participation in this study is voluntary. Your choice to participate, or not to participate, will not influence your future relations with Ryerson University. If you decide to participate, you are free to withdraw your consent and to stop your participation at any time, before or during the experiment, without penalty or loss of benefits to which you are entitled. If you discontinue the study, you will still receive \$30 for your participation. If you would like to participate, but wish your data to be excluded from the analyses you can request this at any point during the study without penalty or loss of benefits to which you are entitled. You must notify the researchers by June 1, 2019 if you wish to have your data removed from the study.

Questions about the study: If you have any questions about the research experience please feel free to ask the experimenter at any time (now or during the study). If you have any questions about this research after you have completed the study you may contact Bonnie Armstrong (phone: 416.979.5000 ext. 2193, or bonnie.armstrong@psych.ryerson.ca).

Questions about your rights as a participant: This research has been reviewed and has received ethical approval by the Research Ethics Committee at Ryerson University. If you have questions about your participation in the research or about your treatment as a research participant, you may contact the Research Ethics Board (Contact: Research Ethics Board, c/o Office of the Vice President, Research and Innovation, Ryerson University, 350 Victoria St., Toronto ON, M5B 2K3, 416.979.5000, ext. 5042, or fax 416.979.5336).

Agreement:

Your signature below indicates that you have read the information in this agreement and have had a chance to ask any questions you have about the study. Your signature also indicates that you agree to be in the study and have been told that you can change your mind and withdraw your consent to participate at any time. You have been given a copy of this agreement. You have been told that by signing this consent agreement you are not giving up any of your legal rights.

Name of Participant (please print)

Signature of Participant

Date

If you would like to be contacted in the future in the event that the experimenter conducts a follow-up study, please sign and date below.

Signature of Participant

Date

☐ Please check this box if you would like to receive a copy of the final report of this research

Signature of Investigator

Date

Appendix IV: Debriefing Form for Experiment 2

DEBRIEFING FORM

DECISION GAME II

Purpose of Study:

If you were to decide whether to hold an afternoon wedding ceremony outdoors at 6pm on the basis of a) the morning weather forecast and b) actual weather conditions at 2pm, you would need to combine this information in such a way that would inform your decision. People have great difficulty “updating” their knowledge with new information to make a decision¹ (e.g., updating one’s knowledge of the weather forecast announced in the morning, with actual weather conditions in the early afternoon in order to decide whether to have the wedding indoors or outdoors at 6pm). People have a tendency to either place more value on their prior knowledge (e.g., the morning weather forecast) in which a *conservatism* heuristic is used or place more value on new information (e.g., weather conditions at 2pm) in which a *representativeness* heuristic is used². Previous research from our lab examined how younger and older adults combine old information with new information, and found that both age groups were able to combine information more accurately when information was experienced, compared to when information was described to them³. However, the cognitive and neural processes underlying younger and older adults’ performance in information integration tasks are still unknown. In particular, it is unclear whether there are age differences in the use of heuristics (e.g., do younger adults use heuristics more or less than older adults?). In addition, it is unknown whether there are age differences in the type of heuristic used (e.g., do younger adults prefer to use a different heuristic than older adults do?). The goal of the current study is to reveal whether age differences exist when old and new information are to be combined, and decisions are to be made based on judgments of the information. In this study, we also measured electrophysiological activity. We measure brain activity in order to examine the biases people have for prior and new information. For example, are there differences in brain activity when people overweight prior information compared to new information? Do younger and older adults show differences in brain activity when overweighting prior and new information?

Main Question:

a) Are there age differences in the use of heuristics? b) Are there age differences in the type of heuristic used? c) Are there age differences in brain activity across different decision situations?

Study Summary:

You made many decisions in which you had to decide which urn a sample of blue and green balls were drawn from. Some of the decision situations you were faced with may have been a) influenced by a heuristic that led to an incorrect decision, b) influenced by a heuristic that led to a correct decision, and c) were not influenced by any heuristic. Our main interest is whether you

used a heuristic that would lead you to make an incorrect choice, or whether you used a different strategy that would lead you to the correct choice.

Study hypothesis:

We hypothesize that age differences exist regarding the frequency of heuristics used, with older adults using heuristics more than younger adults.^{2,4}

Why is this study important?

This study will tell us how susceptible older and younger adults are to heuristics (short-cuts) that lead to incorrect choices in a probability-based task. It will also shed light on whether older and/or younger adults prefer one type of heuristic to another.

Thank you again for your participation! Please contact bonnie.armstrong@psych.ryerson.ca if you have further questions about this study.

Further readings:

1. Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. *Science*. 1974;185:1124-31.
2. Achtziger A, Alos-Ferrer C, Hugelschafer S, Steinhauser M. The neural basis of belief updating and rational decision making. *Soc Cogn Affect Neurosci*. 2014;9:55-62.
3. Armstrong BA, Spaniol J. Experienced probabilities increase understanding of diagnostic test results in younger and older adults. *Med Decis Mak*. 2017; 1-10.
4. Peters E, Finucane ML, MacGregor DG, & Slovic P. The bearable lightness of aging: Judgment and decision processes in older adults. 2000; 144-65.

Appendix V: Consent Form for Experiment 3

CONSENT FORM DECISION GAME III

You are being invited to participate in a research study. Please read this consent form so that you understand what your participation will involve. Before you agree to participate, please ask any questions you may have.

Investigators:

Principal Investigator: Bonnie Armstrong, Department of Psychology, Ryerson University, bonnie.armstrong@psych.ryerson.ca

Faculty Supervisor and Co-Investigator: Dr. Julia Spaniol, Department of Psychology, Ryerson University
jspaniol@psych.ryerson.ca

This study is being conducted in partial fulfillment of the doctoral degree requirements for Bonnie Armstrong. The study is funded by a grant from the Natural Sciences and Engineering Research Council.

Purpose of Study: In this study we are investigating how people make decisions based on probability information. The results may be presented at conferences and/or published in an academic journal.

Description of Study: The study takes a total of 2.5 hours to complete at Ryerson University. Prior to beginning the study, the experimenter will explain the details of the experimental tasks. The session will involve one computer-based task, as well as paper-and-pencil questionnaires. Participants will receive an incentive of \$25 for participating, as well as an opportunity to win an additional \$5 that will be based on your performance on the Decision Game computer task described below. In accordance with the Personal Health Information Protection Act (PHIPA), if you choose to participate, your identity will be kept confidential and your responses will be recorded in such a way that you cannot be identified. You will receive a copy of this consent form, which contains the names and contact numbers of the study investigators.

Tasks: The session will involve paper-and-pencil questionnaires and one computer task. In the questionnaires, you will be asked questions pertaining to various personality traits, dispositions and preferences about certain statements and in given situations, and overall mood and feelings. You will also complete some simple questionnaires of thinking and vocabulary. In the Decision Game computer task you will be presented with 12 jars that each contains blue and green balls. The computer will draw a sample of balls from one of the 12 jars. You will then be shown the sample of balls the computer drew from the chosen jar, and you will be asked to decide which jar

the balls were drawn from. Sometimes some of the balls in the 12 jars will be hidden, and other times they will all be visible. Additionally, sometimes one of the four balls drawn from one of the jars will be hidden, and sometimes all balls that were drawn will be visible. The goal is to decide which jar the sample was drawn from correctly as often as possible. You will receive feedback regarding whether you were correct or incorrect subsequent to each decision.

Confidentiality: Each individual's results are confidential. Neither your identity nor any personal information will be available to anyone other than the investigator. Each participant will receive an identification number (ID). Only the ID number will appear on electronic files, data files as well as on hard-copy data forms and questionnaires. Only individuals involved in the research team will have access to a central password-protected electronic file that matches participant IDs with identifying information (name, date of birth, responses to demographic and health-related questions). This information is being kept in the event that a follow-up study is conducted in the future and if you acknowledge and agree to this use of your information, as well as to be contacted in the future. If a follow-up study is conducted in the future and you would like to be contacted and be included in the follow-up study, the investigator will contact you in order to obtain your consent to use this information. No personal information will be disclosed in any resulting publication or presentation. All research records will be stored in a locked file cabinet in a locked laboratory (with personal information such as your name and signature stored separate from all other information). Electronic files will be password-protected and will be stored locally on a secure server on the Ryerson network. You have the right to withhold any personal health information if you wish to. You also have the right to discontinue your participation in this study at any point without any negative consequences.

If you are interested, we would be happy to provide you with the final results of the study when they appear in press via email. After the study has been published or presented, or after a maximum of 7 years, all paper records (form, questionnaires, etc.) we collect from you will be shredded.

Risks and Benefits: The results of this study may help us better understand how people make decisions that involve probability in situations that vary in the precision of the information. The anticipated benefits of participating in this study include: a) gaining knowledge of psychological experiments and cognitive functions; and 2) contributing to the research of cognitive psychology. Participation in this study is associated with no or minimal risks. Furthermore, you may find the study fun and educational. However, being involved in a study may be a new experience for you, and you may feel uncomfortable answering some of the questions, some of which you may find disturbing or upsetting. However, any potential discomfort is expected to be temporary and will not be greater than what you might experience in a typical day. Remember that you do not have to answer all questions and that you may discontinue participation at any time with no penalty or loss of benefits to which you are entitled. The experimenter will make every effort to ensure that you feel as comfortable as possible throughout the session.

Incentives to participate: Participants will receive an incentive of \$25 for participating. You will receive this incentive even if you do not complete the study for any reason. Participants will receive an additional \$5 if 80% or more of the decisions made in the Decision Game computer task are correct.

Voluntary nature of participation: Participation in this study is voluntary. Your choice to participate, or not to participate, will not influence your future relations with Ryerson University. If you decide to participate, you are free to withdraw your consent and to stop your participation at any time, before or during the experiment, without penalty or loss of benefits to which you are entitled. If you discontinue the study, you will still receive \$25 for your participation. If you would like to participate, but wish your data to be excluded from the analyses you can request this at any point during the study without penalty or loss of benefits to which you are entitled. You must notify the researchers by June 1, 2019 if you wish to have your data removed from the study.

Questions about the study: If you have any questions about the research experience please feel free to ask the experimenter at any time (now or during the study). If you have any questions about this research after you have completed the study you may contact Bonnie Armstrong (phone: 416.979.5000 ext. 2193, or bonnie.armstrong@psych.ryerson.ca).

Questions about your rights as a participant: This research has been reviewed and has received ethical approval by the Research Ethics Committee at Ryerson University. If you have questions about your participation in the research or about your treatment as a research participant, you may contact the Research Ethics Board (Contact: Research Ethics Board, c/o Office of the Vice President, Research and Innovation, Ryerson University, 350 Victoria St., Toronto ON, M5B 2K3, 416.979.5000, ext. 5042, or fax 416.979.5336).

Agreement:

Your signature below indicates that you have read the information in this agreement and have had a chance to ask any questions you have about the study. Your signature also indicates that you agree to be in the study and have been told that you can change your mind and withdraw your consent to participate at any time. You have been given a copy of this agreement. You have been told that by signing this consent agreement you are not giving up any of your legal rights.

Name of Participant (please print)

Signature of Participant

Date

If you would like to be contacted in the future in the event that the experimenter conducts a follow-up study, please sign and date below.

Signature of Participant

Date

☐ Please check this box if you would like to receive a copy of the final report of this research

Signature of Investigator

Date

Appendix VI: Debriefing Form for Experiment 3

DEBRIEFING FORM

DECISION GAME III

Purpose of Study:

If you were to decide whether to hold an afternoon wedding ceremony outdoors at 6pm on the basis of a) the morning weather forecast and b) actual weather conditions at 2pm, you would need to combine this information in such a way that would inform your decision. People have great difficulty “updating” their knowledge with new information to make a decision¹ (e.g., updating one’s knowledge of the weather forecast announced in the morning, with actual weather conditions in the early afternoon in order to decide whether to have the wedding indoors or outdoors at 6pm). People have a tendency to either place more value on their prior knowledge (e.g., the morning weather forecast) in which a *conservatism* heuristic is used or place more value on new information (e.g., weather conditions at 2pm) in which a *representativeness* heuristic is used². Previous research from our lab examined how younger and older adults combine old information with new information, and found that both age groups were able to combine information more accurately when information was experienced, compared to when information was described to them³. However, the cognitive processes underlying younger and older adults’ performance in information integration tasks are still unknown. In particular, it is unclear whether there are age differences in the use of heuristics (e.g., do younger adults use heuristics more or less than older adults?). In addition, it is unknown whether there are age differences in the type of heuristic used (e.g., do younger adults prefer to use a different heuristic than older adults do?), or whether the precision of information affects the type of decisions made. Some research shows that older adults have difficulty representing uncertain information compared to younger adults.⁵ The goal of the current study is to reveal whether age differences exist when old and new information are to be combined, and decisions are to be made based on judgments of uncertain and certain information.

Main Question:

a) Are there age differences in the use of heuristics? b) Can making accurate judgments of probability information be learned over time with feedback? c) Does the precision of information affect the types of decisions made by younger and older adults?

Study Summary:

You made many decisions in which you had to decide which jar a sample of blue and green balls were drawn from. Some of the decision situations you were faced with may have been a) influenced by a heuristic that led to an incorrect decision, b) influenced by a heuristic that led to a correct decision, and c) were not influenced by any heuristic. Our main interest is whether you used a heuristic that would lead you to make an incorrect choice, or whether you used a different strategy that would lead you to the correct choice.

Study hypothesis:

We hypothesize that age differences exist regarding the frequency of heuristics used, with older adults using heuristics more than younger adults.^{2,4}

Why is this study important?

This study will tell us how susceptible older and younger adults are to heuristics (short-cuts in decision making) that lead to incorrect choices in a probability-based task. It will also shed light on whether older and/or younger adults prefer one type of heuristic to another, and whether the precision of information affects the quality of decisions made based on probability information.

Thank you again for your participation! Please contact bonnie.armstrong@psych.ryerson.ca if you have further questions about this study.

Further readings:

1. Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. *Science*. 1974;185:1124-31.
2. Achtziger A, Alos-Ferrer C, Hugelschafer S, Steinhauser M. The neural basis of belief updating and rational decision making. *Soc Cogn Affect Neurosci*. 2014;9:55-62.
3. Armstrong BA, Spaniol J. Experienced probabilities increase understanding of diagnostic test results in younger and older adults. *Med Decis Mak*. 2017; 1-10.
4. Peters E, Finucane ML, MacGregor DG, & Slovic P. The bearable lightness of aging: Judgment and decision processes in older adults. 2000; 144-65.
5. Nassar MR, Bruckner R, Gold JI, Li SC, Heekweren HR, & Eppinger B. Age differences in learning emerge from an insufficient representation of uncertainty in older adults. *Nat Commun*. 2016;7:1-13.

Appendix VII: Experiment 1 Instructions

Practice:

In this experiment you will see two jars that contain a mixture of blue and green balls. You will be presented with the proportion of blue and green balls within each jar. For example, one jar may have 75% blue and 25% green balls, and the other jar may have 50% blue and 50% green balls. In the middle of the computer task there will be a 2-minute break to relax. After this break, the proportion of blue and green balls within each jar will change. For example, one jar may now have 75% green and 25% blue balls, and the other jar may have 50% green and 50% blue balls. You will see this in the practice task as well. In every decision situation you encounter, the computer will choose one of the two jars to draw balls out of. The catch is, you will not be told which jar the computer will draw from! You will only be presented with the likelihood that each jar will be chosen by the computer. For example, there may be a 75% chance that the computer will choose the left jar, and a 25% chance that the computer will choose the right jar to draw balls out of. The likelihood each jar will be chosen by the computer, will change across decision situations. You will then see the sample of balls the computer drew from one of the two jars. The sample of balls drawn by the computer also change across decision situations.

Your task is to choose which jar you believe the sample of balls was drawn from by pressing the L key for the left jar and the R key for the right jar. The sample will stay on the screen until you have made a decision. Once you have made a decision, the sample of balls will be put back in the jar they were drawn from. Along with the 2-minute break in the middle of the computer task, there are also 6 shorter breaks (around 30 seconds each) scattered throughout the task. In the real experiment, if you make 80% or more CORRECT decisions, you will receive an extra \$5. However, you will not be given feedback on the decisions you make. Whether you win the extra \$5 will be revealed at the end of the computer task. It is important to try to make the correct decision in EVERY decision situation in order to maximize your total reward (amount of money won!). We will first go through a practice task. It is important you ask the experimenter anything you do not understand about the task. If you do not have any questions, please press the spacebar to begin.

Experiment:

In this experiment you will see two jars that contain a mixture of blue and green balls. You will be presented with the proportion of blue and green balls within each jar. For example, one jar may have 75% blue balls and 25% green balls, and the other jar may have 50% blue balls and 50% green balls. In the middle of the computer task there will be a 2 minute break to relax. After this break, the proportion of blue and green balls within each jar will change. For example, one jar may now have 75% green balls and 25% blue balls, and the other jar may have 50% green balls and 50% blue balls. You will see this in the practice task too. In every decision situation you encounter, the computer will choose one of the two jars to draw balls out of. The catch is, you will not be told which jar the computer will draw from! You will only be presented with the likelihood that each jar will be chosen by the computer. For example, there may be a 75% chance that the computer will choose the left jar, and a 25% chance that the computer will choose the

right jar to draw balls out of. The likelihood each jar will be chosen by the computer, will change across decision situations. You will then see the sample of balls the computer drew from one of the two jars. The sample of balls drawn by the computer will also change across decision situations.

Your task is to choose which jar you believe the sample of balls was drawn from by pressing the L key for the left jar and the R key for the right jar. The sample will stay on the screen until you have made a decision. Once you have made a decision, the sample of balls will be put back in the jar they were drawn from. Along with the 2 minute break in the middle of the computer task, there are also 6 shorter breaks (around 30 seconds each) scattered throughout the task. In the real experiment, if you make 80% or more CORRECT decisions, you will receive an extra \$5. However, you will not be given feedback on the decisions you make. Whether you win the extra \$5 will be revealed at the end of the computer task. It is important to try to make the correct decision in EVERY decision situation in order to maximize your total reward (amount of money won!). We will first go through a practice task. It is important you ask the experimenter anything you do not understand about the task. If you do not have any questions, please press the spacebar to begin.

30 second break:

Take a brief break (about 30 seconds). Press the spacebar to continue.

2-minute break:

1: Take this time to take a longer break. In 2-minutes this screen will let you know when you may resume. Remember: the proportion of blue and green balls in each bowl is now going to change.

2: Get ready to continue the experiment. Press the spacebar to continue.

Last screen:

Congratulations! The decisions you made were over 80% correct! The decision game is now complete. Thank you for your participation! Please wait for the experimenter for further instruction.

Appendix VIII: Experiment 2 Instructions

Practice:

In this experiment you will see two jars that contain a mixture of blue and green balls. You will be presented with the proportion of blue and green balls within each jar. For example, one jar may have 75% blue and 25% green balls, and the other jar may have 50% blue and 50% green balls. In the middle of the computer task there will be a 2-minute break to relax. After this break, the proportion of blue and green balls within each jar will change. For example, one jar may now have 75% green and 25% blue balls, and the other jar may have 50% green and 50% blue balls. You will see this in the practice task as well. In every decision situation you encounter, the computer will choose one of the two jars to draw balls out of. The catch is, you will not be told which jar the computer will draw from! You will only be presented with the likelihood that each jar will be chosen by the computer. For example, there may be a 75% chance that the computer will choose the left jar, and a 25% chance that the computer will choose the right jar to draw balls out of. The likelihood each jar will be chosen by the computer, will change across decision situations. You will then see the sample of balls the computer drew from one of the two jars. The sample of balls drawn by the computer also change across decision situations.

Your task is to choose which jar you believe the sample of balls was drawn from by pressing the L key for the left jar and the R key for the right jar. The sample will stay on the screen until you have made a decision. Once you have made a decision, the sample of balls will be put back in the jar they were drawn from. In the actual experiment (not the practice), you will be given feedback on the decisions you make (i.e., whether your decision was correct or incorrect). Along with the 2-minute break in the middle of the computer task, there are also 6 shorter breaks (around 30 seconds each) scattered throughout the task. In the real experiment, if you make 80% or more CORRECT decisions, you will receive an extra \$5. Whether you win the extra \$5 will be revealed at the end of the computer task. It is important to try to make the correct decision in EVERY decision situation in order to maximize your total reward (amount of money won!). We will first go through a practice task. It is important you ask the experimenter anything you do not understand about the task. If you do not have any questions, please press the spacebar to begin.

Experiment:

In this experiment you will see two jars that contain a mixture of blue and green balls. You will be presented with the proportion of blue and green balls within each jar. For example, one jar may have 75% blue balls and 25% green balls, and the other jar may have 50% blue balls and 50% green balls. In the middle of the computer task there will be a 2-minute break to relax. After this break, the proportion of blue and green balls within each jar will change. For example, one jar may now have 75% green balls and 25% blue balls, and the other jar may have 50% green balls and 50% blue balls. You will see this in the practice task too. In every decision situation you encounter, the computer will choose one of the two jars to draw balls out of. The catch is, you will not be told which jar the computer will draw from!

You will only be presented with the likelihood that each jar will be chosen by the computer. For example, there may be a 75% chance that the computer will choose the left jar, and a 25%

chance that the computer will choose the right jar to draw balls out of. The likelihood each jar will be chosen by the computer, will change across decision situations. You will then see the sample of balls the computer drew from one of the two jars. The sample of balls drawn by the computer will also change across decision situations.

Your task is to choose which jar you believe the sample of balls was drawn from by pressing the L key for the left jar and the R key for the right jar. The sample will stay on the screen until you have made a decision. Once you have made a decision, the sample of balls will be put back in the jar they were drawn from. You will be given feedback on the decisions you make (i.e., whether your decision was correct or incorrect). Along with the 2-minute break in the middle of the computer task, there are also 6 shorter breaks (around 30 seconds each) scattered throughout the task. As a reminder, you will be rewarded with an extra \$5 if you answer 80% or more of the decisions correctly. It is important to try to make the correct decision in EVERY decision situation in order to maximize your total reward (amount of money won!). Whether you win the extra \$5 or not will be revealed to you at the end of the study. If you have any questions about the task, please ask the experimenter now. Otherwise press the spacebar to begin.

30 second break:

Take a brief break (about 30 seconds). Press the spacebar to continue.

2-minute break:

1: Take this time to take a longer break. In 2-minutes this screen will let you know when you may resume. Remember: the proportion of blue and green balls in each bowl is now going to change.

2: Get ready to continue the experiment. Press the spacebar to continue.

Last screen:

Congratulations! The decisions you made were over 80% correct! The decision game is now complete. Thank you for your participation! Please wait for the experimenter for further instruction.

Appendix IX: Experiment 3 Instructions

Practice:

In this experiment you will see twelve jars that contain a mixture of blue and green balls. You will be presented with the proportion of blue and green balls within each jar. There are two types of jars. One type of jar will have 75% of one colour and 25% of another colour, and the other type of jar will have 50% of one colour and 50% of another colour. Although the proportions of colours in the two jar types will stay the same, the colours of the balls within each jar and the number of jars of each type will change across decision situations. This will be demonstrated in the practice task we do. In every decision situation you encounter, the computer will choose one of the twelve jars to draw a sample of four balls. The catch is you will not be told which jar the computer will draw from! You will only be presented with the twelve jars and the proportion of blue and green balls in each jar. You will then see the sample of balls the computer drew from one of the twelve jars. The sample of balls drawn by the computer will also change across decision situations. Your task is to choose which *type of jar* you believe the sample was drawn from by pressing the L key for the left jar type or the R key for the right jar type. The twelve jars and the sample will stay on the screen until you have made a decision. Once you have made a decision, the sample of balls will be put back in the jar they were drawn from.

In some decision situations you encounter, the contents of some of the twelve jars will be hidden from you. That is, you will not know the proportion of blue and green balls in the jars that are hidden. Importantly, the hidden jars contain proportions of blue and green balls that are identical to the two visible jars, but the number of each type of jar will remain unknown to you.

In other decision situations you encounter, the sample drawn by the computer from one of the twelve jars will contain hidden information. That is, one of the four balls will be hidden, and remain unknown to you. The hidden ball could be either blue or green.

Altogether, there are three types of decision situations:

1. All 12 jars and the sample are visible.
2. Some of the 12 jars are hidden and the sample is visible.
3. All 12 jars are visible and some of the sample is hidden.

There will be two 1-minute breaks separating each type of decision situation, with three optional 30 second breaks within each decision situation block. If you make 80% or more CORRECT decisions, you will receive an extra \$5! Whether you win the extra \$5 will be revealed at the end of the computer task. It is important to try to make the correct decision in EVERY decision situation in order to maximize your total reward. We will first go through a practice task. It is important you ask the experimenter anything you do not understand about the task at this time. If you do not have any questions, please press the spacebar to begin.

Experiment:

In this experiment you will see twelve jars that contain a mixture of blue and green balls. You will be presented with the proportion of blue and green balls within each jar. There are two types

of jars. One type of jar will have 75% of one colour and 25% of another colour, and the other type of jar will have 50% of one colour and 50% of another colour. Although the proportions of colours in the two jar types will stay the same, the colours of the balls within each jar and the number of jars of each type will change across decision situations. This will be demonstrated in the practice task we do. In every decision situation you encounter, the computer will choose one of the twelve jars to draw a sample of four balls. The catch is you will not be told which jar the computer will draw from! You will only be presented with the twelve jars and the proportion of blue and green balls in each jar. You will then see the sample of balls the computer drew from one of the twelve jars. The sample of balls drawn by the computer will also change across decision situations. Your task is to choose which type of jar you believe the sample was drawn from by pressing the L key for the left jar type or the R key for the right jar type. The twelve jars and the sample will stay on the screen until you have made a decision. Once you have made a decision, the sample of balls will be put back in the jar they were drawn from.

In some decision situations you encounter, the contents of some of the twelve jars will be hidden from you. That is, you will not know the proportion of blue and green balls in the jars that are hidden. Importantly, the hidden jars contain proportions of blue and green balls that are identical to the two visible jars, but the number of each type of jar will remain unknown to you.

In other decision situations you encounter, the sample drawn by the computer from one of the twelve jars will contain hidden information. That is, one of the four balls will be hidden, and remain unknown to you. The hidden ball could be either blue or green.

Altogether, there are three types of decision situations:

1. All 12 jars and the sample are visible.
2. Some of the 12 jars are hidden and the sample is visible.
3. All 12 jars are visible and some of the sample is hidden.

There will be two 1-minute breaks separating each type of decision situation, with three optional 30 second breaks within each decision situation block. You will be given feedback on the decisions you make. If you make 80% or more CORRECT decisions, you will receive an extra \$5! Whether you win the extra \$5 will be revealed at the end of the computer task. It is important to try to make the correct decision in EVERY decision situation in order to maximize your total reward. We are about to begin the computer task. It is important you ask the experimenter anything you do not understand about the task at this time. If you do not have any questions, please press the spacebar to begin.

30 second break:

Take a brief break (about 30 seconds). Press the spacebar to continue.

2-minute break:

1: Take this time to take a break. Remember: the type of decision situation will now change. \n\n In 1 minute this screen will let you know when you may resume.

2: Get ready to continue the experiment. Press the spacebar to continue.

Last screen:

Congratulations! The decisions you made were over 80% correct! The decision game is now complete. Thank you for your participation! Please wait for the experimenter for further instruction.

Appendix X: Posterior Probabilities and Corrected Odds of Decision Situations

The table below provides the posterior probabilities for each prior-sample combination for the *left urn*. Posterior probabilities above 0.5 indicate choosing the *left urn* is the correct choice, and probabilities below 0.5 indicate choosing the *left urn* is the incorrect choice.

Posterior probabilities of the left urn contingent on prior probability and sample drawn

Prior	Number of majority colour balls				
	0	1	2	3	4
75%	0.16	0.36	0.63	0.83	0.94
50%	0.06	0.16	0.36	0.63	0.83
25%	0.02	0.06	0.16	0.36	0.63

Following the paradigm by Achtziger et al. (2014), the corrected posterior odds (i.e., the posterior odds – 1) were used to reflect a measure of inverse difficulty for the more likely urn chosen by the computer (see table below).

Corrected odds of the left urn contingent on prior probability and sample drawn

Prior	Number of majority colour balls				
	0	1	2	3	4
75%	4.25	0.78	0.69	4.06	14.19
50%	14.67	4.25	0.78	0.69	4.06
25%	48	14.67	4.25	0.78	0.69

The corrected odds demonstrate how distant the odds are from 1:1. Therefore, the closer the corrected odds value is to 0, the more difficult the decision situation (see bolded corrected odds in table above). The level of difficulty of the six decision situations of interest is also demonstrated in the table of posterior probabilities (see bolded posterior probabilities in table above). Posterior probabilities closer to .50 reflect more difficult decisions, such that the probability of making a correct choice is closer to chance. Therefore, posterior probabilities of .36 and .63 are prior-sample combinations that are of approximately equal difficulty. The cells with bolded corrected odds of approximately equal difficulty represent the six decision situations of interest in the current study.

Appendix XI: Demographic Questionnaire

DEMOGRAPHICS / PHONE SCREENING

Name: _____ Gender: Male Female Age: _____

Phone: _____ Email: _____ Handed: _____

EDUCATION

Currently Student or Non-student _____

Highest level of education obtained _____ Years _____ Area of study _____

Previous degree / diploma NO YES: Area of study _____

LANGUAGE

What is your first language _____ English learned at age _____ Fluent in English YES NO

Language spoken in grade school _____ Other languages _____

What country were you born in _____ Year moved to Canada _____

HEALTH

Vision Do you wear: GLASSES: _____ CONTACTS _____ NONE _____

Operations on your eyes? NO YES: Specify _____

Are you colourblind? NO YES: Specify _____

Hearing Do you have any problems with your hearing? NO YES

If YES, is your hearing corrected with a hearing aid? Specify: _____

Conditions Have you ever had any of the following conditions?

Stroke	NO	YES	Learning disability	NO	YES
Tumor	NO	YES	Psychiatric illness	NO	YES
Neurological diseases	NO	YES	Epilepsy	NO	YES
Head injury	NO	YES	Cancer	NO	YES
Concussion	NO	YES	High blood pressure	NO	YES
Depression	NO	YES	Diabetes	NO	YES
Anxiety	NO	YES	Thyroid disease	NO	YES
Seizure	NO	YES	Serious car accident	NO	YES
Aneurysm	NO	YES	Been unconscious	NO	YES

Have you been taking any medications in the last six months?: _____

Appendix XII: Self-Assessment for Experiments 1 and 2

Self-Assessment

1. Which of the following influenced your decision the most? Circle **one** option below:
 - a) Mostly likelihood information to make decisions
 - b) Mostly sample information to make decisions
 - c) Used both likelihood and sample information to make decisions

2. What strategy did you use to make decisions? Circle **one** option below:
 - a) I used the information indicated in question 1 consistently to make decisions
 - b) I switched the information I used to make a decision (e.g., some decisions I used likelihood information and other decisions I used sample information to make decisions)
 - c) Neither of the above

3. Please explain, step by step, how you made your decisions in the computer task using the space below:

4. In your opinion, how accurate do you think the decisions you made were? Circle **one** option below:
 - a) Very inaccurate
 - b) Somewhat inaccurate
 - c) Neither accurate nor inaccurate
 - d) Somewhat accurate
 - e) Very accurate

5. How difficult did you find the task? Circle **one** option below:
 - a) Very difficult
 - b) Somewhat difficult
 - c) Neither difficult nor easy
 - d) Somewhat easy
 - e) Very easy

6. How confident are you working with numbers? Circle **one** option below:

- a) Not at all confident
- b) Somewhat confident
- c) Confident
- d) Very confident
- e) Extremely confident

7. Please use the space below to provide comments on the task you just completed:

Appendix XIII: Self-Assessment for Experiment 3

Self-Assessment

1. When likelihood and sample information were both unmasked, which of the following influenced your decision the most? Circle **one** option below:
 - a) Mostly likelihood information to make decisions
 - b) Mostly sample information to make decisions
 - c) Used both likelihood and sample information to make decisions
2. What strategy did you use to make decisions when likelihood and sample information were both unmasked? Circle **one** option below:
 - a) I used the information indicated in question 1 consistently to make decisions
 - b) I switched the information I used to make a decision (e.g., some decisions I used likelihood information and other decisions I used sample information to make decisions)
 - c) Neither of the above
3. When likelihood information was unmasked but sample information was masked, which of the following influenced your decision the most? Circle **one** option below:
 - a) Mostly likelihood information to make decisions
 - b) Mostly sample information to make decisions
 - c) Used both likelihood and sample information to make decisions
4. What strategy did you use to make decisions when likelihood information was unmasked but sample information was masked? Circle **one** option below:
 - a) I used the information indicated in question 3 consistently to make decisions
 - b) I switched the information I used to make a decision (e.g., some decisions I used likelihood information and other decisions I used sample information to make decisions)
 - c) Neither of the above
5. When likelihood information was masked but sample information was unmasked, which of the following influenced your decision the most? Circle **one** option below:
 - a) Mostly likelihood information to make decisions
 - b) Mostly sample information to make decisions
 - c) Used both likelihood and sample information to make decisions

6. What strategy did you use to make decisions when likelihood information was masked but sample information was unmasked? Circle **one** option below:
- a) I used the information indicated in question 5 consistently to make decisions
 - b) I switched the information I used to make a decision (e.g., some decisions I used likelihood information and other decisions I used sample information to make decisions)
 - c) Neither of the above
7. Please explain, step by step, how you made your decisions in the computer task using the space below:
-
-
-
-
-
-
-
-
8. In your opinion, how accurate do you think the decisions you made were? Circle **one** option below:
- a) Very inaccurate
 - b) Somewhat inaccurate
 - c) Neither accurate nor inaccurate
 - d) Somewhat accurate
 - e) Very accurate
9. How difficult did you find the task? Circle **one** option below:
- a) Very difficult
 - b) Somewhat difficult
 - c) Neither difficult nor easy
 - d) Somewhat easy
 - e) Very easy

10. How confident are you working with numbers? Circle **one** option below:

- a) Not at all confident
- b) Somewhat confident
- c) Confident
- d) Very confident
- e) Extremely confident

11. Please use the space below to provide comments on the task you just completed:

Appendix XIV: R Script Used for Multi-Level Modeling

```
setwd("C:/Users/dflora/Dropbox/BonnieJuliaMLM")

setwd("C:/Users/dbf/Dropbox/BonnieJuliaMLM")

young <- read.table("young.txt", header = T)
young$age <- 0
young$ager <- 1

old <- read.table("old.txt", header = T)
old$age <- 1
old$ager <- 0

#concatenate young and old data into a single data frame; age = 0 for young, age = 1 for old
all <- rbind.data.frame(young, old)

library(car)
all$cond <- recode(all$Decision.Situation, "c('25_3','75_2')='R'; c('25_4','75_1')='C'; else='F'")
all2 <- all[ which(all$cond != 'F'),]

#concatenate all2 with assessments
all3 = merge(all2, dsst)

#descriptive stats etc.

library(psych)
describe(young$RT)
describe(old$RT)

describe(young$Accuracy)
describe(old$Accuracy)

#describeBy(young$RT, group=young$)
#describeBy(old$RT, group=old$Colour_CB)

describeBy(all2$Accuracy, group=all2$cond)

describeBy(all2$Accuracy, list(all2$cond,all2$age), mat=TRUE,digits=4)
describeBy(all2$RT, list(all2$cond,all2$age), mat=TRUE,digits=4)

#Load 'lme4' package for fitting multilevel models
library(lme4)

#ICC (proportion of variability due to clustering - range from 0-1, higher score more important
to consider clustering).
```

```

#Divide intercept variance and sum of intercept + residual variance
RTICC <- lmer(RT ~ 1 + (1|Subject), data = all3)
summary(RTICC)

#Regress RT on age - "means as outcomes model" because age is predicting each subject's mean
RT across trials
rtAge <- lmer(RT ~ age + (1|Subject), data = all3)

#REML estimation is the default (rather than FIML)
#restricted maximum likelihood (REML) is preferable when level 2 N is small
#here, subjects are the level 2 units
summary(rtAge)
#Under 'Fixed effects', (Intercept) Estimate represents the mean RT for young
#age estimate represents difference between mean RT for young and mean RT for old

#this is a random-intercepts model; assumes that the effect of Block is constant across subjects
#also including age as Level 2 predictor
#only main effects so far; no interactions
rtCon <- lmer(RT ~ age+cond+(1|Subject), data = all3)
summary(rtCon)

#Under 'Fixed effects',
library(car)
Anova(rtCon) #note the capital A

# random slopes model
#only main effects so far; no interactions
rtConrs <- lmer(RT ~ age+cond+(cond|Subject), data = all3)
summary(rtConrs)
#Fixed effect estimates are very similar
Anova(rtConrs)

anova(rtCon, rtConrs) #note the lower-case A
#but random-effects model fits significantly better; also BIC and AIC are better

#add age interaction:
rtConxAge <- lmer(RT ~ age+cond+(age*cond)+(cond|Subject), data = all2)
summary(rtConxAge)
#interaction is significant; don't interpret coefficients of Cond as "main effects"
#to probe interaction, re-code age so that old = 0 and young = 1
#to get "simple slopes" aka "simple main effects" of condition within old
Anova(rtConxAge)

#change reference category, change first letter: all3$cond<-factor(all3$cond,levels=c("R", "A",
"C"))
all3$cond<-factor(all3$cond,levels=c("C", "R"))

```

```

#ICC (proportion of variability due to clustering - range from 0-1, higher score more important
to consider clustering).
#Divide intercept variance and sum of intercept + residual variance
accICC <- lmer(Accuracy ~ 1 + (1|Subject), data = all2)
summary(accICC)

#regress accuracy on age
accAge <- glmer(Accuracy ~ age + (1|Subject), data=all2, family=binomial)
summary(accAge)

exp(fixef(accAge))

#random-intercepts
accCon <- glmer(Accuracy ~ age + cond + (1|Subject), data=all2, family=binomial)
summary(accCon)
exp(fixef(accCon))

#random-slopes
accConrs <- glmer(Accuracy ~ age + cond + (cond|Subject), data=all2, family=binomial)
summary(accConrs)
exp(fixef(accConrs))

#compare random-intercepts to random-slopes model
anova(accCon, accConrs)

#age by condition interaction
accConxAge <- glmer(Accuracy ~ age + cond + (age*cond) + (cond|Subject), data=all2,
family=binomial, control = glmerControl(optimizer = "bobyqa"))

summary(accConxAge)
Anova(accConxAge) #for overall test of interaction
exp(fixef(accConxAge))

#assessments added as interaction terms
#using Lipkus by condition by age interaction as an example
accConxcLipxage <- glmer(Accuracy ~ cLip + cond + age + (cLip*cond*age) + (cond|Subject),
data=all3, family=binomial,
control = glmerControl(optimizer = "bobyqa"))

summary(accConxcLipxage)
Anova(accConxcLipxage)
exp(fixef(accConxcLipxage))

```

Appendix XV: MATLAB Script Used for Signal Processing and ICA

Signal processing for “N2, pP, P3b, FRN”

```
%Batch EEG analyses
clear all;

subject = {'100 resampled pruned with ICA_othercomponents'...
'101 resampled pruned with ICA_othercomponents'...
'102 resampled pruned with ICA_othercomponents'...
'103 resampled pruned with ICA_othercomponents'...
'104 resampled pruned with ICA_othercomponents'...
'105 resampled pruned with ICA_othercomponents'...
'106 resampled pruned with ICA_othercomponents'...
'107 resampled pruned with ICA_othercomponents'...
'108 resampled pruned with ICA_othercomponents'...
'109 resampled pruned with ICA_othercomponents'...
'110 resampled pruned with ICA_othercomponents'...}
Etc.

pathWrite = 'C:\Users\Carson\Documents\MATLAB\DGII\processed_data\';
pathRawData = '/Users/carson/Documents/MATLAB/Bonnie_DG/original_data/';
pathDefs = '/Users/carson/Documents/MATLAB/Bonnie_DG/scripts_lists/';

nsubj = length(subject);

% set flags to 1 if you want the associated operations to be performed
basic = 1;
filter = 0;

%convert BDF to SET
for s = 1:nsubj

    % Basic processing
    if basic == 1
        fprintf('\n')

        fprintf('*****\n')
        fprintf('BASIC PROCESSING FOR subject #g : %s ...\n', s,
subject{s})
        EEG = pop_loadset('filename', [subject{s} '.set'], 'filepath',
[pathWrite 'pruned_ICA/']);
        %fprintf('\n')

        %fprintf('*****\n')
        %fprintf('ARTIFACT REJECTION AND AVERAGING FOR subject #g : %s
...\n', s, subject{s})
        %EEG = pop_eegthresh(EEG,1,[1:68] ,-75,75,-0.199,0.996,2,0); % reject
epochs amplitues exceeding +/- 75Hz

        % Find out number of total events in file:
        numTrials = EEG.trials;
        halfTrial = round(numTrials/2) - 1;
```

```

EEG_first_half = pop_select(EEG, 'trial', [1:halfTrial]);
EEG_second_half = pop_select(EEG, 'trial', [(halfTrial+1):numTrials]);

ERP_first_half = pop_averager( EEG_first_half, 'DSindex', 1,
'Criterion', 'good'); %compute average waveform for each bin
ERP_second_half = pop_averager( EEG_second_half, 'DSindex', 1,
'Criterion', 'good');

% FIRST HALF PROCESSING

%Collapse across left/right targets
%ERP = pop_erpchanoperator( ERP, {'nchan1 = (ch13-ch50) label
C3_C4'}));
ERP_first_half = pop_binoperator( ERP_first_half, {'b34 = (b10+b12)/2
label RepCorrect'}));
ERP_first_half = pop_binoperator( ERP_first_half, {'b35 = (b14+b16)/2
label ConsCorrect'}));
ERP_first_half = pop_binoperator( ERP_first_half, {'b36 = (b18+b20)/2
label AlignCorrect'}));
ERP_first_half = pop_binoperator( ERP_first_half, {'b37 = b35 - b34
label Neutral-Conf_CORRECT_DifferenceWave'}));

ERP_first_half = pop_binoperator( ERP_first_half, {'b38 = (b11+b13)/2
label RepINCorrect'}));
ERP_first_half = pop_binoperator( ERP_first_half, {'b39 = (b15+b17)/2
label ConsINCorrect'}));
ERP_first_half = pop_binoperator( ERP_first_half, {'b40 = (b19+b21)/2
label AlignINCorrect'}));
ERP_first_half = pop_binoperator( ERP_first_half, {'b41 = b39 - b38
label Neutral-Conf_INCORRECT_DifferenceWave'}));

%ERP = pop_binoperator( ERP, {'nb1 = b1 label GainIncentive_Cue',
'nb2 = b2 label GainNeutral_Cue', 'nb3 = b3 label LossIncentive_Cue',...
'nb4 = b4 label LossNeutral_Cue', 'nb5 = b5 label
NeutralNeutral_Cue', 'nb6 = (b6+b7)/2 label GI_Cong_Tar', 'nb7 = (b8+b9)/2
label GI_Incong_Tar',...
'nb8 = (b10+b11)/2 label GN_Cong_Tar', 'nb9 = (b12+b13)/2 label
GN_Incong_Tar', 'nb10 = (b14+b15)/2 label LI_Cong_Tar',...
'nb11 = (b16+b17)/2 label LI_Incong_Tar', 'nb12 = (b18+b19)/2 label
LN_Cong_Tar', 'nb13 = (b20+b21)/2 label LN_Incong_Tar', 'nb14 = (b22+b23)/2
label NN_Cong_Tar',...
'nb15 = (b24+b25)/2 label NN_Incong_Tar', 'nb16 = b26 label
GI_FastCorr_FDBK', 'nb17 = b27 label GI_Slow_FDBK',...
'nb18 = b28 label LI_FastCorr_FDBK', 'nb19 = b29 label
LI_Slow_FDBK', 'nb20 = b30 label GN_Neutral_FDBK', 'nb21 = b31 label
LN_Neutral_FDBK',...
'nb22 = b32 label NN_Neutral_FDBK'}));

%pop_summary_AR_erp_detection(ERP, [pathWrite 'D_Artifact_Summary\'
subject{s} '.txt']); %Summarize rejected trials, save as txt file
saveERP(ERP_first_half, [pathWrite subject{s}
'_FIRST_half_averaged.erp']); % Save ERPs

```

```

%EEG = pop_saveset( EEG, 'filename', [subject{s} '_rej.set'],
'filepath', [pathWrite 'C_Artifact_Rejected']); %Save EEG file with rejected
trials

% SECOND HALF PROCESSING

%Collapse across left/right targets
%ERP = pop_erpchanoperator( ERP, {'nchan1 = (ch13-ch50) label
C3_C4'});
ERP_second_half = pop_binoperator( ERP_second_half, {'b34 =
(b10+b12)/2 label RepCorrect'});
ERP_second_half = pop_binoperator( ERP_second_half, {'b35 =
(b14+b16)/2 label ConsCorrect'});
ERP_second_half = pop_binoperator( ERP_second_half, {'b36 =
(b18+b20)/2 label AlignCorrect'});
ERP_second_half = pop_binoperator( ERP_second_half, {'b37 = b35 - b34
label Neutral-Conf_CORRECT_DifferenceWave'});

ERP_second_half = pop_binoperator( ERP_second_half, {'b38 =
(b11+b13)/2 label RepINCorrect'});
ERP_second_half = pop_binoperator( ERP_second_half, {'b39 =
(b15+b17)/2 label ConsINCorrect'});
ERP_second_half = pop_binoperator( ERP_second_half, {'b40 =
(b19+b21)/2 label AlignINCorrect'});
ERP_second_half = pop_binoperator( ERP_second_half, {'b41 = b39 - b38
label Neutral-Conf_INCORRECT_DifferenceWave'});

%ERP = pop_binoperator( ERP, {'nb1 = b1 label GainIncentive_Cue',
'nb2 = b2 label GainNeutral_Cue', 'nb3 = b3 label LossIncentive_Cue',...
'nb4 = b4 label LossNeutral_Cue', 'nb5 = b5 label
NeutralNeutral_Cue', 'nb6 = (b6+b7)/2 label GI_Cong_Tar', 'nb7 = (b8+b9)/2
label GI_Incong_Tar',...
'nb8 = (b10+b11)/2 label GN_Cong_Tar', 'nb9 = (b12+b13)/2 label
GN_Incong_Tar', 'nb10 = (b14+b15)/2 label LI_Cong_Tar',...
'nb11 = (b16+b17)/2 label LI_Incong_Tar', 'nb12 = (b18+b19)/2 label
LN_Cong_Tar', 'nb13 = (b20+b21)/2 label LN_Incong_Tar', 'nb14 = (b22+b23)/2
label NN_Cong_Tar',...
'nb15 = (b24+b25)/2 label NN_Incong_Tar', 'nb16 = b26 label
GI_FastCorr_FDBK', 'nb17 = b27 label GI_Slow_FDBK',...
'nb18 = b28 label LI_FastCorr_FDBK', 'nb19 = b29 label
LI_Slow_FDBK', 'nb20 = b30 label GN_Neutral_FDBK', 'nb21 = b31 label
LN_Neutral_FDBK',...
'nb22 = b32 label NN_Neutral_FDBK'});

%pop_summary_AR_erp_detection(ERP, [pathWrite 'D_Artifact_Summary\'
subject{s} '.txt']); %Summarize rejected trials, save as txt file
saveERP(ERP_second_half, [pathWrite subject{s}
'_SECOND_half_averaged.erp']); % Save ERPs
%EEG = pop_saveset( EEG, 'filename', [subject{s} '_rej.set'],
'filepath', [pathWrite 'C_Artifact_Rejected']); %Save EEG file with rejected
trials

EEG_first_half = eeg_checkset( EEG_first_half );
EEG_second_half = eeg_checkset( EEG_second_half );

```

```

end

%filter
if filter == 1
    fprintf('\n')

fprintf('*****\n')
    fprintf('FILTERING subject #g : %s ...\n', s, subject{s})
    ERP = pop_loaderp([pathWrite subject{s} '_averaged.erp'], pathWrite);
    ERP = pop_filterp(ERP, 1:70, 0, 30, 2, 'butter', 0);
    saveERP(ERP, [pathWrite subject{s} '_f.erp']);
end
end

```

Signal processing for LRP (for first and second block)

```

%Batch EEG analyses
clear all;

subject = {'100 resampled pruned with ICA_LRP'...
'101 resampled pruned with ICA_LRP'...
'102 resampled pruned with ICA_LRP'...
'103 resampled pruned with ICA_LRP'...
'104 resampled pruned with ICA_LRP'...
'105 resampled pruned with ICA_LRP'...
'106 resampled pruned with ICA_LRP'...
'107 resampled pruned with ICA_LRP'...
'108 resampled pruned with ICA_LRP'...
'109 resampled pruned with ICA_LRP'...
'110 resampled pruned with ICA_LRP'...}
Etc...

pathWrite = '/Users/carson/Documents/MATLAB/Bonnie_DG/processed_data/';
pathRawData = '/Users/carson/Documents/MATLAB/Bonnie_DG/original_data/';
pathDefs = '/Users/carson/Documents/MATLAB/Bonnie_DG/scripts_lists/';

nsubj = length(subject);

% set flags to 1 if you want the associated operations to be performed
basic = 1;
filter = 0;

%convert BDF to SET
for s = 1:nsubj

    % Basic processing
    if basic == 1
        fprintf('\n')

fprintf('*****\n')
        fprintf('BASIC PROCESSING FOR subject #g : %s ...\n', s,
subject{s})
        EEG = pop_loadset('filename', [subject{s} '.set'], 'filepath',
[pathWrite 'pruned_ICA/']);

```

```

        fprintf('\n')

fprintf('*****\n')
    fprintf('ARTIFACT REJECTION AND AVERAGING FOR subject #g : %s
...\n', s, subject{s})
    EEG = pop_eegthresh(EEG,1,[1:68] ,-75,75,-0.199,0.996,2,0); % reject
epochs amplitudes exceeding +/- 75Hz
    ERP = pop_averager( EEG, 'DSindex', 1, 'Criterion', 'good'); %compute
average waveform for each bin

    %Collapse across left/right targets

    ERP = pop_erpchanoperator( ERP, {'nchan1 = (ch13-ch50) label
C3_C4'});

    % Create new bins:
    ERP = pop_binoperator( ERP, {'b7 = (b1+b3)/2 label K25'});
    ERP = pop_binoperator( ERP, {'b8 = (b2+b4)/2 label K75'});
    ERP = pop_binoperator( ERP, {'b9 = (b5+b6)/2 label K50'});
    ERP = pop_binoperator( ERP, {'b10 = (b8-b7) label LRP'});

    %ERP = pop_binoperator( ERP, {'nb1 = b1 label GainIncentive_Cue',
'nb2 = b2 label GainNeutral_Cue', 'nb3 = b3 label LossIncentive_Cue',...
    %'nb4 = b4 label LossNeutral_Cue', 'nb5 = b5 label
NeutralNeutral_Cue', 'nb6 = (b6+b7)/2 label GI_Cong_Tar', 'nb7 = (b8+b9)/2
label GI_Incong_Tar',...
    %'nb8 = (b10+b11)/2 label GN_Cong_Tar', 'nb9 = (b12+b13)/2 label
GN_Incong_Tar', 'nb10 = (b14+b15)/2 label LI_Cong_Tar',...
    %'nb11 = (b16+b17)/2 label LI_Incong_Tar', 'nb12 = (b18+b19)/2 label
LN_Cong_Tar', 'nb13 = (b20+b21)/2 label LN_Incong_Tar', 'nb14 = (b22+b23)/2
label NN_Cong_Tar',...
    %'nb15 = (b24+b25)/2 label NN_Incong_Tar', 'nb16 = b26 label
GI_FastCorr_FDBK', 'nb17 = b27 label GI_Slow_FDBK',...
    %'nb18 = b28 label LI_FastCorr_FDBK', 'nb19 = b29 label
LI_Slow_FDBK', 'nb20 = b30 label GN_Neutral_FDBK', 'nb21 = b31 label
LN_Neutral_FDBK',...
    %'nb22 = b32 label NN_Neutral_FDBK'});

    %pop_summary_AR_erp_detection(ERP, [pathWrite 'D_Artifact_Summary\'
subject{s} '.txt']); %Summarize rejected trials, save as txt file
    saveERP(ERP, [pathWrite subject{s} '_averaged.erp']); % Save ERPs
    EEG = pop_saveset( EEG, 'filename', [subject{s} '_rej.set'],
'filepath', [pathWrite 'C_Artifact_Rejected']); %Save EEG file with rejected
trials
    EEG = eeg_checkset( EEG );
end

%filter
if filter == 1
    fprintf('\n')

fprintf('*****\n')
    fprintf('FILTERING subject #g : %s ...\n', s, subject{s})
    ERP = pop_loaderp([pathWrite subject{s} '_averaged.erp'], pathWrite);
    ERP = pop_filterp( ERP,1:70, 0, 30, 2, 'butter', 0);
    saveERP(ERP, [pathWrite subject{s} '_f.erp']);

```



```

        end
    end
ICA script

% ICA analysis
    if ICA == 1

        fileTags = [fileTags '_ICA'];

        %EEG = pop_runica(EEG,'extended',1,'logfile',[EEG.setname,'-
ICAllog.txt']);
        EEG = pop_runica(EEG,'extended',1);
        EEG = pop_saveset(EEG,'filename', [subject{s} fileTags '.set'],
'filepath', pathWrite);

        %EEG = pop_runica(EEG,'extended',1,'logfile',[EEG.setname,'-
ICAllog.txt']);
        EEGLRP = pop_runica(EEGLRP,'extended',1);
        EEGLRP = pop_saveset(EEGLRP,'filename', [subject{s} fileTags
'LRP.set'], 'filepath', pathWrite);
    end
end

```

Appendix XVI: ERP Latency Results

Latency

N2, pP latency. The 2 (age) x 2 (decision situation) x 2 (block) x 2 (region) mixed ANOVA on N2 and pP peak latency revealed a significant Age by Region interaction, $F(1, 48) = 18.55, p < .001, \eta_p^2 = .28$. To probe this interaction, independent- and paired-samples t tests were conducted to examine potential differences in region within each age group, as well as the potential differences age differences for each region. For younger adults, the pP ($M = 393.65, SD = 7.59$) had a longer latency than the N2 ($M = 253.59, SD = 10.87$), $t(24) = -52.65, p < .001, d = 10.53$. Similarly, for older adults, the pP ($M = 402.63, SD = 8.19$) had a longer latency than the N2 ($M = 245.66, SD = 10.89$), $t(24) = -54.34, p < .001, d = 10.87$. Further, younger adults showed a longer N2 peak latency than older adults, $t(24) = 2.57, p = .01, d = .73$. However, older adults showed a longer pP latency than younger adults, $t(24) = -4.02, p < .001, d = 1.14$.

P3b latency. The 2 (age) x 2 (decision situation) x 2 (block) mixed ANOVA on P3b peak latency revealed no main effect of age, $F(1, 48) = .73, p > .05, \eta_p^2 = .02$, no main effect of block, $F(1, 48) = 1.08, p > .05, \eta_p^2 = .02$, and no main effect of decision situation, $F(1, 48) = .51, p > .05, \eta_p^2 = .01$. No interactions were observed.

FRN latency. The 2 (age) x 2 (decision situation) x 2 (block) x 2 (feedback: correct, incorrect) mixed ANOVA on FRN peak latency revealed a significant main effect of age, $F(1, 48) = 1648.00, p < .001, \eta_p^2 = .97$, such that younger adults ($M = 265.61, SD = 7.71$) showed earlier FRN peak latencies than older adults ($M = 351.25, SD = 7.20$).

Additionally, an Age by Decision Situation by Block three-way interaction was observed, $F(1, 48) = 6.26, p = .02, \eta_p^2 = .12$. To unpack this interaction, stratifying by age, paired-samples t tests were used to examine differences in decision situations within each block, as well as the potential block effects within each decision situation. For younger adults, no differences in the

RconfB ($M = 267.19$, $SD = 13.71$) or CconfB situation ($M = 266.79$, $SD = 11.06$) were observed in the first block, $t(24) = .15$, $p > .05$, or between the RconfB ($M = 264.92$, $SD = 10.59$) and CconfB situation ($M = 263.52$, $SD = 11.97$), in the second block, $t(24) = .53$, $p > .05$. Further, no differences were observed in the RconfB situation in the first, $t(24) = .64$, $p > .05$, or the second block, $t(24) = 1.16$, $p > .05$. For older adults, no differences in the RconfB ($M = 353.83$, $SD = 11.19$) or CconfB situation ($M = 348.05$, $SD = 13.63$), were observed in the first block, $t(24) = 1.51$, $p > .05$, however a significant difference was observed in FRN peak latency between the RconfB ($M = 347.89$, $SD = 12.81$) and CconfB situation ($M = 355.23$, $SD = 10.52$), in the second block, $t(24) = -3.08$, $p = .01$, $d = -.61$. Further, a significant difference in FRN peak latency was observed in the RconfB situation between the first and second blocks, $t(24) = 1.98$, $p = .05$, $d = .40$, and in the CconfB situation between the first and second blocks, $t(24) = -2.42$, $p = .02$, $d = -.48$, for older adults.

References

- Achtziger, A., Alós-Ferrer, C., Hügelschäfer, S., & Steinhauser, M. (2014). The neural basis of belief updating and rational decision making. *Social Cognitive and Affective Neuroscience*, 9, 55–62. doi: 10.1093/scan/nss099.
- Ajzen, I. (1977). Intuitive theories of events and the effects of base-rate information on prediction. *Journal of Personality and Social Psychology*, 35, 303-314. doi: 10.1037/0022-3514.35.5.303.
- Albar, F. M., & Jetter, A. J. (2009). *Heuristics in Decision Making*, in: *Proceedings of PICMET 2009: Technology Management in the Age of Fundamental Change*. Portland, OR.
- Albert, J., Lopez-Martin, S., Hinojosa, J. A., & Carretie, L. (2013). Spatiotemporal characterization of response inhibition. *NeuroImage*, 76, 272–281. doi: 10.1016/j.neuroimage.2013.03.011.
- Alós-Ferrer, C., Garagnani, M., & Hügelschäfer, S. (2016). Cognitive reflection, decision biases, and response times. *Frontiers in Psychology*, 7, 1402. doi: 10.3389/fpsyg.2016.01402.
- Alós-Ferrer, C., & Hügelschäfer, S. (2012). Faith in intuition and behavioral biases. *Journal of Economic Behavior and Organization*, 84, 182-192. doi: 10.1016/j.jebo.2012.08.004.
- Amodio, D. M., Devine, P. G., & Harmon-Jones, E. (2008). Individual differences in the regulation of intergroup bias: The role of conflict monitoring and neural signals for control. *Journal of Personality and Social Psychology*, 94, 60-74. doi: 10.1037/0022-3514.94.1.60.
- Arkes, H., & Ayton, P. (1999). The sunk cost and Concorde effects: Are humans less rational than lower animals? *Psychological Bulletin*, 125, 591-600. doi: 10.1037/0033-2909.125.5.591.
- Armstrong, B. A., & Spaniol, J. (2017). Experienced probabilities increase understanding of

- diagnostic test results in younger and older adults. *Medical Decision Making*, 37, 670-679. doi:10.1177/0272989X17691954.
- Armstrong, B. A., Spaniol, J., & Persaud, N. (2018). Does exposure to simulated patient cases improve accuracy of clinicians' predictive value estimates of diagnostic test results? A within-subjects experiment at St. Michael's Hospital, Toronto, Canada. *BMJ Open*, 8. doi: e019241:10.1136/bmjopen-2017-019241.
- Armstrong, B. A., Sparrow, E. P., & Spaniol, J. (under review). Bayesian inference is affected by absolute frequencies but not temporal learning or acute stress.
- Bäckman, L., Nyberg, L., Lindenberger, U., Li, S. C., & Farde, L. (2006). The correlative triad among aging, dopamine, and cognition. *Neuroscience and Biobehavioral Reviews*, 30, 791–807. doi: 10.1016/j.neubiorev.2006.06.005.
- Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology*, 49, 5-28. doi: 10.1080/713755608.
- Badre, D., Doll, B. B., Long, N. M., & Frank, M. J. (2012). Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration. *Neuron*, 73, 595-607. doi: 10.1016/j.neuron.2011.12.025.
- Band, G. P. H., Ridderinkhof, K. R., & Segalowitz, S. (2002). Explaining neurocognitive aging: Is one factor enough? *Brain and Cognition*, 49, 259-267. doi: 10.1006/brcg.2001.1499.
- Banks, J., & Oldfield, Z. (2007). Understanding pensions: Cognitive function, numerical ability and retirement saving. *Fiscal Studies*, 28, 143-170. doi: 10.1111/j.1475-5890.2007.00052.x.
- Barbey, A. K., & Sloman, S. A. (2007). Base-rate respect: From ecological rationality to dual processes. *Behavioral and Brain Sciences*, 30, 241-254. doi:10.1017/S0140525X07001653.
- Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. *Acta Psychologica*, 44,

211-233. doi: 10.1016/0001-6918(80)90046-3.

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

Baron, J., Scott, S., Fincher, K., & Metz, E. (2015). Why does the Cognitive Reflection Test (sometimes) predict utilitarian moral judgment (and other things)? *Journal of Applied Research in Memory and Cognition*, 4, 265-284. doi: 10.1016/j.jarmac.2014.09.003.

Barrett, H. C., & Kurzban, R. (2006). Modularity in cognition: Framing the debate. *Psychological Review*, 113, 628-647. doi: 10.1037/0033-295X.113.3.628.

Bartholow, B. D., & Dickter, C. L. (2008). A response conflict account of the effects of stereotypes on racial categorization. *Social Cognition*, 26, 314-332. doi: 10.1521/soco.2008.26.3.314.

Bartholow, B. D., Pearson, M. A., Dickter, C. L., Sher, K. J., Fabiani, M., & Gratton, G. (2005). Strategic control and medial frontal negativity: Beyond errors and response conflict. *Psychophysiology*, 42, 33-42. doi: 10.1111/j.1469-8986.2005.00258.x.

Bartholow, B. D., Riordan, M. A., Saults, J. S., & Lust, S. A. (2009). Psychophysiological evidence of response conflict and strategic control of responses in affective priming. *Journal of Experimental Social Psychology*, 45, 655-666. doi: 10.1016/j.jesp.2009.02.015.

Bayes, T., & Price, M (1763). An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, 53, 370-418. doi: 10.1098/rstl.1763.0053.

Bennett, D., Murawski, & Bode, S. (2015). Single-trial event-related potential correlates of belief updating. *eNeuro*, 15, 2: doi: 10.1523/ENEURO.0076-15.2015.

Best, J. R., Miller, P. H., & Jones, L. L. (2009). Executive functions after age 5: Changes and correlates. *Developmental Review*, 29, 180-200. doi: 10.1016/j.dr.2009.05.002.

- Betsch, C., & Iannello, P. (2010). Measuring individual differences in intuitive and deliberate decision making styles: A comparison of different measures. In A. Glöckner, & C. Witteman (Eds.), *Tracing intuition: Recent methods in measuring intuitive and deliberate processes in decision making* (pp. 251–267). London, UK: Psychology Press.
- Betsch, C., & Iannello, P. (in preparation). A unified scale to assess individual differences in intuition and deliberation (USID).
- Besedes, T., Deck, C., Sarangi, S., & Shor, M. (2012). Age effects and heuristics in decision making. *Review of Economics and Statistics*, 96, 638–647. doi: 10.1162/REST.
- Binmore, K. (2009). *Rational decisions*. Princeton, NJ: Princeton University Press.
- Binmore, K., Stewart, L., & Voorhoeve, A. (2012). How much ambiguity aversion?: Finding indifferences between Ellsberg’s risky and ambiguous bets. *Journal of Risk and Uncertainty*, 45, 215–238. doi: 10.1007/s11166-012-9155-3.
- Blanchard-Fields, F., Hertzog, C., Stein, R., & Pak, R. (2001). Beyond a stereotyped view of older adults’ traditional family values. *Psychology and Aging*, 16, 438–496. doi: 10.1037//0882-7974.16.3.483.
- Bokura, H., Yamaguchi, S., Matsubara, M., & Kobayashi, S. (2002). Frontal lobe contribution to response processes – an ERP study and aging effect. *International Congress Series*, 1232, 17–20. doi: 10.1016/S0531-5131(01)00677-X.
- Bonner, C., & Newell, B. R. (2010). In conflict with ourselves? An investigation of heuristic and analytic processes in decision making. *Memory & Cognition*, 38, 186–196. doi:10.3758/MC.38.2.186.
- Bopp, K. L., & Verhaeghen, P. (2005). Aging and verbal memory span: A meta-analysis. *Journal of Gerontology: Psychological Sciences*, 60, 223–233. doi:

10.1093/geronb/60.5.P223.

Bors, D. A., & Forrin, B. (1995). Age, speed of information processing, recall, and fluid intelligence. *Intelligence*, 20, 229-248. doi: 10.1016/0160-2896(95)90009-8.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, 108, 624-652. doi: 10.1037/0033-295X.108.3.624.

Brainerd, C. J. (2007). Kissing cousins but not identical twins: The denominator neglect and base-rate respect models. *Behavioral and Brain Sciences*, 30, 257–258.
doi:10.1017/S0140525X07001689.

Brand, M., Labudda, K., & Markowitsch, H. J. (2006). Neuropsychological correlates of decision-making in ambiguous and risky situations. *Neural Network*, 19, 1266–1276. doi: 10.1016/j.neunet.2006.03.001.

Braver, T. S., & Barch, D. M. (2002). A theory of cognitive control, aging cognition, and neuromodulation. *Neuroscience and Biobehavioral Reviews*, 26, 809-817. doi: 10.1016/S0149-7634(02)00067-2.

Braver, T. S., Barch, D. M., Gray, J. R., Molfese, D. L., & Snyder, A. (2001). Anterior cingulate cortex and response conflict: Effects of frequency, inhibition, and errors. *Cerebral Cortex*, 11, 825-836. doi: 10.1093/cercor/11.9.825.

Brinley, J. E (1965). Cognitive sets, speed and accuracy of performance in the elderly. In A. T. Welford & J. E. Birren (Eds.), *Behavior, aging and the nervous system* (pp. 114-149). Springfield, IL: Charles C Thomas.

Brosnan, M., Hollinworth, M., Antoniadou, K., & Lewton, M. (2014). Is emphasizing intuitive and systemizing deliberative? *Personality and Individual Differences*, 66, 39-43. doi:

10.1016/j.paid.2014.03.006.

Bruin, K. J., & Wijers, A. A. (2002). Inhibition, response mode, and stimulus probability: A comparative event-related potential study. *Clinical Neurophysiology*, 113, 1172-1182. doi: 10.1016/S1388-2457(02)00141-4.

Bruine De Bruin, W., McNair, S. J., Taylor, A. L., Summers, B., & Strough, J. (2015). "Thinking about numbers is not my idea of fun": Need for cognition mediates age differences in numeracy performance. *Medical Decision Making*, 35, 22-26. doi: 10.1177/0272989X14542485.

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

Bruine De Bruin, W., Parker, A. M., & Fischhoff, B. (2009). *Explaining adult age differences in decision-making competence*. Bi-annual meeting of Subjective Probability and Utility in Decision Making (SPUDM), Rovereto, Italy.

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. doi: 10.1037/0022-3514.92.5.938.

Bruine De Bruin, W., Vanderklaauw, W., Downs, J. S., Fischhoff, B., Topa, G., & Armantier, O. (2009). Expectations of inflation: The role of demographic variables, expectation formation, and financial literacy. *Journal of Consumer Affairs*, 44, 381-402. doi: 10.1111/j.1745-6606.2010.01174.x.

Burgess, G. C., Gray, J. R., Conway, A. R., & Braver, T. S. (2011). Neural mechanisms of interference control underlie the relationship between fluid intelligence and working memory

- span. *Journal of Experimental Psychology: General*, 140, 674-692. doi: 10.1037/a0024695.
- Burns, L. R., & D’Zurilla, T. J. (1999). Individual differences in perceived information processing styles in stress and coping situations: Development and validation of the perceived modes of processing inventory. *Cognitive Therapy and Research*, 23, 345–371. doi: 10.1023/A:1018799700207.
- Cabeza, R., Albert, M., Belleville, S., Craik, F. I., Duarte, A., Grady, C. L., ... Rajah, M. N. (2018). Maintenance, reserve, and compensation: The cognitive neuroscience of healthy ageing. *Nature Reviews Neuroscience*, 19, 701-710.
- Cabeza, R., Dennis, N. A. (2012). *Frontal lobes and aging: Deterioration and compensation*. In Stuss, D. T., Knight, R. T. (Eds.). *Principles of Frontal Lobe Function*, 2nd Edition (pp. 628-652). New York: Oxford University Press.
- Cabeza, R., Nyberg, L., & Park, D. (2005). *Cognitive neuroscience of aging: Linking cognitive and cerebral aging*. New York, NY: Oxford University Press.
- Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42, 116-131. doi: 10.1037/0022-3514.42.1.116.
- Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, 119, 197-253. doi: 10.1037/0033-2909.119.2.197.
- Campitelli, G., & Gerrans, P. (2014). Does the cognitive reflection test measure cognitive reflection? A mathematical modeling approach. *Memory and Cognition*, 42, 434-447. doi: 10.3758/s13421-013-0367-9.
- Campitella, G., & Labollita, M. (2010). Correlations of cognitive reflection with judgments and choices. *Judgment and Decision Making*, 5, 182-191.

- Carruthers, P. (2006). *The architecture of the mind*. New York: Oxford University Press.
- Carstensen, L. L., & Hartel, C. R. (2006). *When I'm 64*. National Academy Press.
- Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston, MA: Houghton Mifflin.
- Chen, Y., & Sun, Y. (2003). Age differences in financial decision-making using simple heuristics. *Educational Gerontology*, 29, 627–635. doi: 10.1080/03601270390218152.
- Charles, S. T., Mather, M., & Carstensen, L. L. (2003). Aging and emotional memory: The forgettable nature of negative images for older adults. *Journal of Experimental Psychology: General*, 132, 310–324. doi: 10.1037/0096-3445.132.2.310.
- Chowdhury, R., Guitart-Masip, M., Lambert, C., Dayan, P., Huys, Q., Düzel, E., & Dolan, R. J. (2013). Dopamine restores reward prediction errors in old age. *Nature Neuroscience*, 16, 648-653. doi: 10.1038/nn.3364.
- Chuderski, A. (2014). The relational integration task explains fluid reasoning above and beyond other working memory tasks. *Memory and Cognition*, 42, 448-463. doi: 10.3758/s13421-013-0366-x.
- Clark, A. (2001). *Mindware: An introduction to the philosophy of cognitive science*. New York: Oxford University Press.
- Cokely, E. T., Feltz, A., Ghazal, S., Allan, J. N., Petrova, D., & Garcia-Retamero, R. (in press). Decision making skill: From intelligence to numeracy and expertise. In K. A. Ericsson, R. R. Hoffman, A. Kozbelt, & A. M. Williams (2nd Eds.), *Cambridge Handbook of Expertise and Expert Performance*. New York, NY: Cambridge University Press.
- Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin numeracy test. *Judgment and Decision Making*, 7, 25–47. doi:

10.1037/t45862-000.

- Cokely, E. T., Ghazal, S., & Garcia-Retamero, R. (2014). Measuring numeracy. In B. L. Anderson & J. Schulkin (Eds.), *Numerical reasoning in judgments and decision making about health* (pp. 11-38). Cambridge University Press.
- Cokely, E. T., & Kelley, C. M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgment and Decision Making*, 4, 20-33. doi: 2009-02933-002.
- Cools, E., & Van den Broeck, H. (2007). Development and validation of the Cognitive Style Indicator. *Journal of Psychology: Interdisciplinary and Applied*, 141, 359–387. doi: 10.3200/JRLP.141.4.359-388.
- d’Acremont, M., Schultz, W., & Bossaerts, P. (2013). The human brain encodes event frequencies while forming subjective beliefs. *Journal of Neuroscience*, 33, 10887-10897. doi: 10.1523/JNEUROSCI.5829-12.2013.
- Dagleish, T., Taghavi, R., Neshat-Doost, H., Moradi, A., Canterbury, R., & Yule, W. (2003). Patterns of processing bias for emotional information across clinical disorders: A comparison of attention, memory, and prospective cognition in children and adolescents with depression, generalized anxiety, and posttraumatic stress disorder. *Journal of Clinical Child and Adolescent Psychology*, 32, 10-12. doi: 10.1207/15374420360533022.
- Dashiell, J. F. (1937). Affective value-distances as a determinant of aesthetic judgment-times. *American Journal of Psychology*, 50, 57–67. doi: 10.2307/1416620.
- Dave, C., & Wolfe, K. W. (2003). On confirmation bias and deviations from Bayesian updating. Working Paper, Department of Economics, University of Pittsburgh.
- Davis, S. W., Dennis, N. A., Daselaar, S. M., Fleck, M. S., & Cabeza, R. (2008). Que PASA?

- The posterior-anterior shift in aging. *Cerebral Cortex*, 18, 1201–1209.
doi:10.1093/cercor/bhm155.
- Daw, N. D., O’Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates of exploratory decisions in humans. *Nature*, 441, 876-879. doi: 10.1038/nature04766.
- Dayan, P. (2012). Twenty-five lessons from computational neuromodulation. *Neuron* 76, 240-256. doi: 10.1016/j.neuron.2012.09.027.
- de Boer, L., Axelsson, J., Riklund, K., Nyberg, L., Dayan, P., Backman, L., & Guitart-Masip, M. (2017). Attenuation of dopamine-modulated prefrontal value signals underlies probabilistic reward learning deficits in old age. *eLIFE*, 6, e26424. doi: 10.7554/eLife.26424.001.
- Delaney, R., Strough, J., Parker, A. M., & Bruine de Bruin, W. (2015). Variations in decision making profiles by age and gender: A cluster-analytic approach. *Personality and Individual Differences*, 85, 19-24. doi: 10.1016/j.paid.2015.04.034.
- Delorme, A., & Makeig, S. (2004). EEGLab: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 15, 9-21. doi: 10.1016/j.jneumeth.2003.10.009.
- De Jong, R., Coles, M. G. H., Logan, G. D., & Gratton, G. (1990). In search of the point of no return: The control of response processes. *Journal of Experimental Psychology: Human Perception & Performance*, 16, 164-182. doi: 10.1037/0096-1523.16.1.164.
- De Neys, W. (2012). Bias and conflict: A case for logical intuitions. *Perspectives on Psychological Science*, 7, 28–38. doi:10.1177/1745691611429354.
- De Neys, W. (2007). Nested sets and base-rate neglect: Two types of reasoning? *Behavioral and Brain Sciences*, 30, 260–261. doi:10.1017/S0140525X07001719.
- De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking.

- Cognition*, 106, 1248–1299. doi:10.1016/j.cognition.2007.06.002.
- De Neys, W., Schaeken, W., & d’Ydewalle, G. (2005). Working memory and everyday conditional reasoning: Retrieval and inhibition of stored counterexamples. *Thinking and Reasoning*, 11, 349–381. doi: 10.1080/13546780442000222.
- De Neys, W., & Van Gelder, E. (2008). Logic and belief across the lifespan: The rise and fall of belief inhibition during syllogistic reasoning. *Developmental Science*, 11, 986–993. doi: 10.1111/j.1467-7687.2008.00746.x.
- De Neys, W., & Verschueren, N. (2006). Working memory capacity and a notorious brain teaser. *Experimental Psychology*, 53, 123–131. doi: 10.1027/1618-3169.53.2.123.
- Denburg, N. L., Tranel, D., & Bechara, A. (2005). The ability to decide advantageous declines prematurely in some normal older persons. *Neuropsychologia*, 43, 1099–1106. doi: 10.1016/j.neuropsychologia.2004.09.012.
- Dennis, N. A., & Cabeza, R. (2008). Neuroimaging of healthy cognitive aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 1–54). New York, NY, US: Psychology Press.
- Diamond, A. (2013). Executive functions. *Annual Reviews of Psychology*, 64, 135–168. doi: 10.1146/annurev-psych-113011-143750.
- Di Rosa, E., Mapelli, D., Arcara, G., Amodio, P., Tamburin, S., & Schiff, S. (2017). Aging and risky decision-making: New ERP evidence from the Iowa Gambling Task. *Neuroscience Letters*, 640, 93–98. doi: 10.1016/j.neulet.2017.01.021.
- Donkers, F. C. L., & van Boxtel, G. J. M. (2004). The N2 in go/no-go tasks reflects conflict monitoring not response inhibition. *Brain and Cognition*, 56, 165–176. doi: 10.1016/j.bandc.2004.04.005.

- Duncan, J., Parr, A., Woolgar, A., Thompson, R., Bright, P., Cox, S., ... Nimmo-Smith, I. (2008). Goal neglect and Spearman's g: Competing parts of a complex task. *Journal of Experimental Psychology: General*, 137, 131-148. doi: 10.1037/0096-3445.137.1.131.
- Edwards, W. (1968). Conservatism in human information processing. In B. Kleinmuntz (Ed.), *Formal representation of human judgment* (pp. 17-52). New York: Wiley.
- Eimer, M. (1998). The lateralized readiness potential as an on-line measure of central response activation processes. *Behavior Research Methods, Instruments, and Computers*, 30, 146-156. doi: 10.1016/S0166-4115(97)80027-1.
- Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *The Quarterly Journal of Economics*, 75, 643-669. doi: 10.2307/1884324.
- Enriquez-Geppert, S., Konrad, C., Pantev, C., & Huster, R. J. (2010). Conflict and inhibition differentially affect the N200/P300 complex in a combined go/nogo and stop-signal task. *Neuroimage*, 51, 877-887. doi: 10.1016/j.neuroimage.2010.02.043.
- Eppinger, B., Hämmerer, D., & Li, S. C. (2011). Neuromodulation of reward-based learning and decision making in human aging. *Annals of the New York Academy of Science*, 1235, 1-17. doi: 10.1111/j.1749-6632.2011.06230.x.
- Eppinger, B., Heekeren, H. R., & Li, S. C. (2015). Age-related prefrontal impairments implicate deficient prediction of future reward in older adults. *Neurobiology of Aging*, 36, 2380-2390. doi: 10.1016/j.neurobiolaging.2015.04.010.
- Eppinger, B., & Kray, J. (2011). To choose or to avoid: Age differences in learning from positive and negative feedback. *Journal of Cognitive Neuroscience*, 23, 41-52. doi: 10.1162/jocn.2009.21364.
- Eppinger, B., Kray, J., Mock, B., & Mecklinger, A. (2008). Better or worse than expected?

- Aging, learning, and the ERN. *Neuropsychologia*, 46, 521-539. doi: 10.1016/j.neuropsychologia.2007.09.001.
- Eppinger, B., Schuck, N. W., Nystrom, L. E., & Cohen, J. D. (2013). Reduced striatal responses to reward prediction errors in older compared with younger adults. *Journal of Neuroscience*, 33, 9905-9912. doi: 10.1523/JNEUROSCI.2942-12.2013.
- Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *The American Psychologist*, 49, 709-724. doi: 10.1037/0003-066X.50.9.799.
- Epstein, S., Pacini, R., Denes-Raj, V., & Heier, H. (1996). Individual differences in intuitive-experiential and analytical-rational thinking styles. *Journal of Personality and Social Psychology*, 71, 390-405. doi: 10.1037/0022-3514.71.2.390.
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception and Psychophysics*, 16, 143-149. doi: 10.3758/BF03203267.
- Erixon-Lindroth, N., Farde, L., Wahlin, T. B., Sovago, J., Halldin, C., & Backman, L. (2005). The role of the striatal dopamine transporter in cognitive aging. *Psychiatry Research*, 138, 1-12. doi: 10.1016/j.psychresns.2004.09.005.
- Evans, J. St. B. T. (2010). *Thinking twice: Two minds in one brain*. Oxford: Oxford University Press.
- Evans, J. St. B. T. (2009). How many dual-process theories do we need? One, two, or many? In J. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 33-54). Oxford: Oxford University Press.
- Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59, 255-278. doi:

10.1146/annurev.psych.59.103006.093629.

Evans, J. St. B. T. (2007). *Hypothetical thinking: Dual processes in reasoning and judgment*. New York: Psychology Press.

Evans, J. S. B. T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences*, 7, 454–459. doi: 10.1016/j.tics.2003.08.012.

Evans, J. St. B. T., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, 8, 223–241. doi: 10.1177/1745691612460685.

Fallgatter, A. J., & Strik, W. K. (1999). The NoGo-anteriorization as a neurophysiological standard-index for cognitive response control. *International Journal of Psychophysiology*, 32, 233–238. doi: 10.1016/S0167-8760(99)00018-5.

Falkenstein, M., Hoormann, J., & Hohnsbein, J. (2002). Inhibition-related ERP components: Variation with modality, age, and time-on-task. *Journal of Psychophysiology*, 16, 167–175. doi: 10.1027//0269-8803.16.3.167.

Falkenstein, M., Hoormann, J., & Hohnsbein, J. (1999). ERP components in Go/Nogo tasks and their relation to inhibition. *Acta Psychologica*, 101, 267–291. doi: 10.1016/S0001-6918(99)00008-6.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2013). Statistical power analyses using g*power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149–1160. doi: 10.3758/BRM.41.4.1149.

Feldman Barrett, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin*, 130, 553–573. doi: 10.1037/0033-2909.130.4.553.

- Fernandes, C., Pasion, R., Goncalves, A. R., Ferreira-Santos, F., Barbosa, F., Martins, I. P., & Marques-Teixeira, J. (2018). Age differences in neural correlates of feedback processing after economic decisions under risk. *Neurobiology of Aging*, 65, 51-59. doi: 10.1016/j.neurobiolaging.2018.01.003.
- Ferreira, M. B., Garcia-Marques, L., Sherman, S. J., & Sherman, J. W. (2006). Automatic and controlled components of judgment and decision making. *Journal of Personality and Social Psychology*, 91, 797–813. doi:10.1037/0022-3514.91.5.797.
- Finucane, M. L., Mertz, C. K., Slovic, P., & Schmidt, E. S. (2005). Task complexity and older adults' decision-making competence. *Psychology and Aging*, 20, 71–84. doi: 10.1037/0882-7974.20.1.71\|10.1037/0882-7974.20.1.71.supp.
- Fiedler, K. (2000). Beware of samples! A cognitive-ecological sampling approach to judgment biases. *Psychological Review*, 107, 659–676. doi: 10.1037/0033-295X.107.4.659.
- Fisk, J. E. (2005). Age and probabilistic reasoning: Biases in conjunctive, disjunctive and Bayesian judgements in early and late adulthood. *Journal of Behavioral Decision Making*, 18, 55–82. doi: 10.1002/bdm.488.
- Fjell, A. M., & Walhovd, K. B. (2010). Structural brain changes in aging: Courses, causes and cognitive consequences. *Reviews in the Neurosciences*, 21, 187–221. doi: 10.1515/REVNEURO.2010.21.3.187.
- Fjell, A. M., & Walhovd, K. B. (2001). P300 and neuropsychological tests as measures of aging: Scalp topography and cognitive changes. *Brain Topography*, 14, 25-40. doi: 10.1023/A:1012563605837.
- Flora, D. B. (2018). *Statistical methods for the social and behavioural sciences: A model based approach*. Thousand Oaks, CA: Sage.

- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189-198. doi: 10.1016/0022-3956(75)90026-6.
- Forder, L., & Dyson, B. J. (2016). Behavioural and neural modulation of win-stay but not lose-shift strategies as a function of outcome value in rock, paper, scissors. *Scientific Reports*, 6, 33809. doi: 10.1038/srep33809.
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, 19, 25–42. doi:10.1257/089533005775196732.
- Friedman, D. (2003). Cognition and aging: A highly selective overview of event-related potential (ERP) data. *Journal of Clinical and Experimental Neuropsychology*, 25, 702-720. doi:10.1076/jcen.25.5.702.14578.
- Friedman, D. (2012). The components of aging. In S. Kappenman & S. J. Luck (Eds.), *The Oxford handbook of event-related potential components* (pp. 1-28). New York, NY:Oxford University Press.
- Furl, N., & Averbeck, B. B. (2011). Parietal cortex and insula relate to evidence seeking relevant to reward-related decisions. *Journal of Neuroscience*, 48, 17572-17582. doi: 10.1523/JNEUROSCI.4236-11.2011.
- Galesic, M., & Garcia-Retamero, R. (2011). Graph literacy: A cross-cultural comparison. *Medical Decision Making*, 31, 444–457. doi: 10.1177/0272989X10373805.
- Galesic, M., Gigerenzer, G., & Straubinger, N. (2009). Natural frequencies help older adults and people with low numeracy to evaluate medical screening tests. *Medical Decision Making*, 29, 368–371. doi: 10.1177/0272989X08329463.
- Garcia-Retamero, R., & Cokely, E. T. (2013). Communicating health risks with visual aids.

- Current Directions in Psychological Science*, 22, 392–399. doi: 10.1177/0963721413491570.
- Garcia-Retamero, R., & Galesic, M. (2010). Who profits from visual aids: Overcoming challenges in people's understanding of risks. *Social Science and Medicine*, 70, 1019–1025. doi: 10.1016/j.socscimed.2009.11.031.
- Garcia-Retamero, R., Galesic, M., & Gigerenzer, G. (2010). Do icon arrays help reduce denominator neglect? *Medical Decision Making*, 30, 672–684. doi: 10.1177/0272989X10369000.
- Garcia-Retamero, R., Wicki, B., Cokely, E. T., & Hanson, B. (2014). Factors predicting surgeons' preferred and actual roles in interactions with their patients. *Health Psychology*, 33, 920–928. doi: 10.1037/hea0000061.
- Gazzaley, A., 2011. Influence of early attentional modulation on working memory. *Neuropsychologia* 49, 1410–1424. doi: 10.1016/j.neuropsychologia.2010.12.022.
- Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: Bridging selective attention and working memory. *Trends in Cognitive Science*, 16, 129–135. doi: 10.1016/j.tics.2011.11.014.
- Ghazal, S., Cokely, E. T., & Garcia-Retamero, R. (2014). Predicting biases in very highly educated samples: Numeracy and metacognitions. *Judgment and Decision Making*, 9, 15–34. doi: 2014-03718-002.
- Gigerenzer, G. (2016). Archival insights into the evolution of economics. In R. Frantz & L. Marsh (Eds.), *Minds, Models, and Milieux*. New York, NY: Palgrave MacMillian.
- Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, 62, 451–482. doi: 10.1146/annurev-psych-120709-145346.
- Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L., M., & Woloshin, S. (2007). Helping doctors and patients make sense of health statistics. *Psychological Science in the*

- Public Interest*, 8, 53-96. doi: 10.1111/j.1539-6053.2008.00033.x.
- Gigerenzer G., Hertwig R. & Pachur T. (2011). *Heuristics: The foundations of adaptive behavior*. New York: Oxford University Press.
- Gigerenzer, G., Todd, P. M., & the ABC Research Group. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Gilinsky, A. S., & Judd, B. B. (1994). Working memory and bias in reasoning across the life span. *Psychology and Aging*, 9, 356-371. doi: 10.1037/0882-7974.9.3.356.
- Glöckner, A., & Witteman, C. (2010). Beyond dual-process models: A categorisation of processes underlying intuitive judgement and decision making. *Thinking & Reasoning*, 16, 1–25. doi: 10.1080/13546780903395748.
- Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, 109, 75–90. doi: 10.1037//0033-295X.109.1.75.
- Goodie, A. S., & Fantino, E. (1999). What does and does not alleviate base-rate neglect under direct experience. *Journal of Behavioral Decision Making*, 12, 307-335. doi: 10.1002/(SICI)1099-0771(199912)12:4<307::AID-BDM324>3.0.CO;2-H.
- Grady, C. L. (2012). The cognitive neuroscience of ageing. *Nature Neuroscience Reviews*, 13, 491-505. doi: 10.1038/nrn3256.
- Grady, C. L. (2008). Cognitive neuroscience of aging. *Annual of the New York Academy of Sciences*, 1124, 127-144. doi: 10.1196/annals.1440.009.
- Grady, C. L., Bernstein, L. J., Beig, S., & Siegenthaler, A. L. (2002). The effects of encoding task on age-related differences in the functional neuroanatomy of face memory. *Psychology and Aging*, 17, 7–23. doi: 10.1037/0882-7974.17.1.7.
- Grady, C. L., Maisog, J. M., Horwitz, B., Ungerleider, L. G., Mentis, M. J., Salerno, J. A., ...

- Haxby, J. V. (1994). Age-related changes in cortical blood flow activation during visual processing of faces and location. *Journal of Neuroscience*, *14*, 1460-1462. doi: 10.1523/JNEUROSCI.14-03-01450.1994.
- Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience*, *6*, 316-322. doi: 10.1038/nn1014.
- Grether, D. M. (1992). Testing Bayes rule and the representativeness heuristic: Some experimental evidence. *Journal of Economic Behavior and Organization*, *17*, 31 –57. doi: 10.1016/0167-2681(92)90078-P.
- Grether, D. M. (1980). Bayes rule as a descriptive model: The representativeness heuristic. *Quarterly Journal of Economics*, *95*, 537–557. doi: 10.2307/1885092.
- Groom, M. J., & Cragg, L. (2015). Differential modulation of the N2 and P3 event-related potentials by response conflict and inhibition. *Brain and Cognition*, *97*, 1-9. doi: 10.1016/j.bandc.2015.04.004.
- Grützmann, R., Riesel, A., Klawohn, J., Kathmann, N., & Endrass, T. (2014). Complementary modulation of N2 and CRN by conflict frequency. *Psychophysiology*, *51*, 761-772. doi: 10.1111/psyp.12222.
- Hamm, R. M. (1993). Explanations for common responses to the Blue/Green cab probabilistic inference word problem. *Psychological Reports*, *27*, 219-242. doi: 10.2466/pr0.1993.72.1.219.
- Hämmerer, D., Li, S. C., Müller, V., & Lindenberger, U. (2011). Life span differences electrophysiological correlates of monitoring gains and losses during probabilistic reinforcement learning. *Journal of Cognitive Neuroscience*, *23*, 579-592. doi: 10.1162/jocn.2010.21475.

- Hämmerer, D., Li, S. C., Müller, V., & Lindenberger, U. (2010). An electrophysiological study of response conflict processing across the lifespan: Assessing the roles of conflict monitoring, cue utilization, response anticipation, and response suppression. *Neuropsychologia*, 48, 3305-3316. doi: 10.1016/j.neuropsychologia.2010.07.014.
- Hämmerer, D., Schwartenbeck, P., Gallagher, M., FitzGerald, T. H. B., Düzel, E., & Dolan, R. J. (2019). Older adults fail to form stable task representations during model-based reversal inference. *Neurobiology of Aging*, 74, 90-100. doi: 10.1016/j.neurobiolaging.2018.10.009.
- Handley, S. J., Capon, A., Beveridge, M., Dennis, A., & Evans, J. St. B. T. (2004). Working memory, inhibitory control and the development of children's reasoning. *Thinking and Reasoning*, 10, 175-195. doi: 10.1080/13546780442000051.
- Hanoch, Y., Wood, S., & Rice, T. (2007). Bounded rationality, emotions and older adult decision making: Not so fast and yet so frugal. *Human Development*, 50, 333–358. doi: 10.1159/000109835.
- Harada, C. N., Natelson Love, M. C., & Triebel, K. L. (2013). Normal cognitive aging. *Clinics in Geriatric Medicine*, 29, 737-752. doi: 10.1016/j.cger.2013.07.002.
- Hartshorne, J. K., & Germine, L. T. (2015). When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life span. *Psychological Science*, 26, 433–443. doi:10.1177/0956797614567339
- Hasher, L., Lustig, C., & Zacks, R. (2007). Inhibitory mechanisms and the control of attention. In A. Conway, C. Jarrold, M. Kane, A. Miyake, & J. Towse. (Eds.), *Variation in working memory* (pp. 227-249). New York: Oxford University Press.
- Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. *Psychology of Learning and Motivation*, 22, 193-225. doi: 10.1016/S0079-

7421(08)60041-9.

- Hauser, T. U., Iannaccone, R., Stämpfli, P., Drechsler, R., Brandeis, D., Walitza, S., & Brem, S. (2014). The feedback-related negativity (FRN) revisited: New insights into the localization, meaning and network organization. *NeuroImage*, 84, 159-168. doi: 10.1016/j.neuroimage.2013.08.028.
- Heath, C., & Tversky, A. (1991). Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty*, 4, 5-28. doi: 10.1007/BF00057884.
- Hedden, T., & Gabrieli, J. D. E. (2004). Insights into the ageing mind: A view from cognitive neuroscience. *Nature Reviews Neuroscience*, 5, 87-97. doi: 10.1038/nrn1323.
- Heldmann, M., Markgraf, U., Rodríguez-Fornells, A., & Münte, T. F. (2008). Brain potentials reveal the role of conflict in human errorful and errorless learning. *Neuroscience Letters*, 444, 64–68. doi: 10.1016/j.neulet.2008.07.042.
- Herbert, M., Eppinger, B., & Kray, J. (2011). Younger but not older adults benefit from salient feedback during learning. *Frontiers in Psychology*, 2, 171. doi: 10.3389/fpsyg.2011.00171.
- Hertzog, C. R., Smith, M., & Ariel, R. (2017). Does the cognitive reflection test actually capture heuristics versus analytic reasoning styles in older adults? *Experimental Aging Research*, 44, 18-34. doi: 10.1080/0361073X.2017.1398508.
- Hess, T. M. (2001). Ageing-related influences on personal need for structure. *International Journal of Behavioral Development*, 25, 482-490. doi: 10.1080/01650250042000429.
- Hess, T. M. (2014). Selective engagement of cognitive resources: Motivational influences on older adults' cognitive functioning. *Perspective on Psychological Science*, 9, 388-407. doi: 10.1177/1745691614527465.
- Hess, T. M., & Queen, T. L. (2014). Aging influences on judgments and decision processes:

- Interactions between ability and experience. In P. Verhaeghen & C. Hertzog (Eds.), *The Oxford handbook of emotion, social cognition, and problem solving in adulthood* (pp. 238–255). New York, NY: Oxford University Press.
- Hess, T. M., Waters, S. J., & Bolstad, C. A. (2000). Motivational and cognitive influences on affective priming in adulthood. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 55, 193-204. doi: 10.1093/geronb/55.4.P193.
- Hicks, K. L., Harrison, T. L., & Engle, R. W. (2015). Wonderlic, working memory capacity, and fluid intelligence. *Intelligence*, 50, 186-195. doi: 10.1016/j.intell.2015.03.005.
- Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, 109, 679-709. doi: 10.1037//0033-295X.109.4.679.
- Holroyd, C. B., & Krigolson, O. E. (2007). Reward prediction error signals associated with a modified time estimation task. *Psychophysiology*, 44, 913-917. doi: 10.1111/j.1469-8986.2007.00561.x.
- Hoppe, E. I., & Kusterer, D. J. (2011). Behavioral biases and cognitive reflection. *Economics Letters*, 110, 97-100. doi: 10.1016/j.econlet.2010.11.015.
- Houdé, O. (1997). The problem of deductive competence and the inhibitory control of cognition. *Current Psychology of Cognition*, 16, 108-113.
- Horn, J. L. (1970). Organization of data on life-span development of human abilities. In L. R. Goulet & P. B. Baltes (Eds.), *Life-span developmental psychology: Research and theory*. New York: Academic Press.
- Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica*, 26, 107-129. doi: 10.1016/0001-6918(67)90011-X.

- Horn, J. L., & Donaldson, G. (1976). On the myth of intellectual decline in adulthood. *American Psychologist*, 31, 701–719. doi:10.1037/0003-066x.31.10.701
- Huettel, S. A., Stowe, C. J., Gordon, E. M., Warner, B. T., & Platt, M. L. (2006). Neural signatures of economic preferences for risk and ambiguity. *Neuron*, 49, 765-775. doi: 10.1016/j.neuron.2006.01.024.
- Hügelschäfer, S. (2011). Bayesian updating in the EEG: Differentiation between automatic and controlled processes of human economic decision making. (Unpublished doctoral dissertation), University of Konstanz, Konstanz, Germany.
- Huster, R.J., Enriquez-Geppert, S., Lavallee, C. F., Falkenstein, M., & Herrmann, C. S. (2013). Electroencephalography of response inhibition tasks: Functional networks and cognitive contributions. *International Journal of Psychophysiology*, 87, 217-233. doi: 10.1016/j.ijpsycho.2012.08.001
- Hsieh, S., Wu, M., & Tang, C. H. (2016). Adaptive strategies for the elderly in inhibiting irrelevant and conflict no-go trials while performing the Go/No-Go task. *Frontiers in Aging Neuroscience*, 7, 243. doi: 10.3389/fnagi.2015.00243.
- Ito, S., Stuphorn, V., Brown, J. W., & Schall, J. D. (2003). Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science*, 302, 120-122. doi: 10.1126/science.1087847.
- Jarvik, M. E. (1951). Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *Journal of Experimental Psychology*, 41, 291–297. doi: 10.1037/h0056878.
- Jiang, J., Beck, J., Heller, K., & Egner, T. (2015). An insula-frontostriatal network mediates flexible cognitive control by adaptively predicting changing control demands. *Nature*

Communications, 6, 8165. doi: 10.1038/ncomms9165.

Johnson, M. M. S. (1990). Age differences in decision making: A process methodology for examining strategic information processing. *Journal of Gerontology: Psychological Sciences*, 45, P75-P78. doi: 10.1093/geronj/45.2.P75.

Jones, M., & Sugden, R. (2001). Positive confirmation bias in the acquisition of information. *Theory and Decision*, 50, 59–99. doi: 10.1023/A:1005296023424.

Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. *Neuron* 89, 221–234. doi: 10.1016/j.neuron.2015.11.028.

Juslin, P. (2015). Controlled information integration and Bayesian inference. *Frontiers in Psychology*, 6, 2009–2012. doi: 10.3389/fpsyg.2015.00070.

Juslin, P., Nilsson, H., & Winman, A. (2009). Probability theory, not the very guide of life. *Psychological Review*, 116, 856–874. doi: 10.1037/a0016979.

Juslin, P., Nilsson, H., Winman, A., & Lindskog, M. (2011). Reducing cognitive biases in probabilistic reasoning by the use of logarithm formats. *Cognition*, 120, 248–267. doi: 10.1016/j.cognition.2011.05.004.

Juslin, P., & Persson, M. (2002). PROBABILITIES from EXemplars (PROBEX): A “lazy” algorithm for probabilistic inference from generic knowledge. *Cognitive Science*, 26, 563–607. doi: 10.1016/S0364-0213(02)00083-6.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin and Review*, 9, 637–671. doi: 10.3758/BF03196323.

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

- Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *The American Economic Review*, 93, 1449–1475. doi: 10.1257/000282803322655392.
- Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases* (pp. 49–81). New York, NY: Cambridge University Press.
doi:10.1017/CBO9780511808098.004.
- Kahneman, D., & Tversky, A. (1973). Psychology of prediction. *Psychological Review*, 80, 237–251. doi:10.1037/h0034747.
- Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3, 430–454. doi: 10.1016/0010-0285(72)90016-3.
- Keller, J., Bohner, G., & Erb, H. P. (2000). Intuitive and heuristic judgment – different processes? Presentation of a German version of the rational-experiential inventory and of new self-report scales of heuristic use. *Zeitschrift für Sozialpsychologie*, 31, 87–101. doi: 10.1024//0044-3514.31.2.87.
- Kelley, C. M., & McLaughlin, A. C. (2012). Individual differences in the benefits of feedback for learning. *Human Factors*, 54, 26-35. doi: 10.1177/0018720811423919.
- Kim, S., & Hasher, L. (2005). The attraction effect in decision making: Superior performance in older adults. *Quarterly Journal of Experimental Psychology*, 58A, 120-133. doi: 10.1080/02724980443000160.
- Kirmizi-Alsan, E., Bayraktaroglu, Z., Gurvit, H., Keskin, Y. H., Emre, M., & Demiralp, T. (2006). Comparative analysis of event-related potentials during Go/NoGo and CPT: Decomposition of electrophysiological markers of response inhibition and sustained attention. *Brain Research*, 1104, 114-128. doi: 10.1016/j.brainres.2006.03.010.

- Klaczynski, P. A. (2014). Heuristics and biases: Interactions among numeracy, ability, and reflectiveness predict normative responding. *Frontiers in Psychology*, 5, 1-13. doi: 10.3389/fpsyg.2014.00665.
- Klayman, J. (1995). Varieties of confirmation bias. *Psychology of Learning and Motivation: Advances in Research and Theory*, 32, 385–418. doi: 10.1016/S0079-7421(08)60315-1.
- Knight, F. (1921). *Risk, Uncertainty and Profit. Reprints of Economic Classics*. New York, NY: Sentry Press.
- Koehler, J. J. (1996). The base rate fallacy reconsidered: Descriptive, normative, and methodological challenges. *Behavioral and Brain Sciences*, 19, 1-53. doi: 10.1017/S0140525X00041157.
- Kolossa, A., Kopp, B., & Fingscheidt. (2015). A computational analysis of the neural bases of Bayesian inference. *Neuroimage*, 106, 222-237. Doi: 10.1016/j.neuroimage.2014.11.007.
- Kopp, B., Seer, C., Lange, F., Kluytmans, A., Kolossa, A., Fingscheidt, T., & Hoijtink, H. (2016). P300 amplitude variations, prior probabilities, and likelihoods: A Bayesian ERP study. *Cognitive, Affective, & Behavioral Neuroscience*, 16, 911-928. doi: 10.3758/s13415-016-0442-3.
- Korsch, M., Frühholz, S., & Herrmann, M. (2016). Conflict-specific aging effects mainly manifest in early information processing stages – An ERP study with different conflict types. *Frontiers in Aging Neuroscience*, 8, 53. doi: 10.3389/fnagi.2016.00053.
- Kropotov, J., Ponomarev, V., Tereshchenko, E. P., Müller, A., & Jäncke, L. (2016). Effect of aging on ERP components of cognitive control. *Frontiers in Aging Neuroscience*, 8, 1-15. doi: 10.3389/fnagi.2016.00069.

- Kutas, M., & Donchin, E. (1980). Preparation to respond as manifested by movement-related brain potentials. *Brain Research*, 202, 95-115. doi: 10.1016/S0006-8993(80)80037-0.
- Larsen, M. J., Clayson, P. E., Primosch, M., Leyton, M. & Steffensen, S. C. (2015). The effects of acute dopamine precursor depletion on the cognitive control functions of performance monitoring and conflict processing: An event-related potential (ERP) study. *PloS One*, 22, e0140770. doi: 10.1371/journal.pone.0140770.
- Lenartowicz, A., Kalar, D. J., Congdon, E., & Poldrack, R. A. (2010). Towards an ontology of cognitive control. *Topics in Cognitive Science*, 2, 678-692. doi: 10.1111/j.1756-8765.2010.01100.x.
- Leuthold, H., Sommer, W., & Ulrich, R. (1996). Partial advance information and response preparation: Inferences from the lateralized readiness potential. *Journal of Experimental Psychology: General*, 125, 307-323. doi: 10.1037/0096-3445.125.3.307.
- Liberali, J.M., Reyna, V. F., Furlan, S., Stein, L. M., & Pardo, S. T. (2012). Individual differences in numeracy and cognitive reflection, with implications for biases and fallacies in probability judgment. *Journal of Behavioral Decision Making*, 25, 361-381. doi: 10.1002/bdm.752.
- Lindenberger, U., Mayr, U., & Kliegl, R. (1993). Speed and intelligence in old age. *Psychology and Aging*, 8, 207–220. doi: 10.1037/0882-7974.8.2.207.
- Lipkus, I., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making*, 21, 37–44. doi:10.1177/0272989X0102100105.
- Liu, P. J., Wood, S., & Hanoch, Y. (2015). Choice and aging: Less is more. In T. M. Hess, J. N. Strough, & C. E. Lockenhoff (Eds.), *Aging and Decision Making: Empirical and Applied*

- Perspectives* (pp. 17–40). Academic Press. doi: 10.1016/B978-0-12-417148-0.00002-9.
- Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the Depression Anxiety Stress Scales*. Sydney: Psychology Foundation.
- Lu, Y. (2015). Is experiential-intuitive cognitive style more inclined to err on conjunction fallacy than analytical-rational cognitive style? *Frontiers in Psychology*, 6, 85. doi: 10.3389/fpsyg.2015.00085.
- Lucci, G., Berchicci, M., Spinelli, D., Taddei, F., & Di Russo, F. (2013). The effects of aging on conflict detection. *PLoS One*, 8, e56566. doi: 10.1371/journal.pone.0056566.
- Luck, S. J. (2014). *An introduction to the event-related potential technique* (2nd ed.). Cambridge, MA: MIT Press.
- Lusardi, A., & Mitchell, O. S. (2014). The economic importance of financial literacy: Theory and evidence. *Journal of Economic Literature*, 52, 5-44. doi: 10.1257/jel.52.1.5.
- Lustig, C., Hasher, L., & Zacks, R. (2007). Inhibitory deficit theory: Recent developments in a “New View”. In D. Gorfein & C. M. MacLeod (Eds.), *Inhibition in cognition* (pp. 145-162). Washington, DC: American Psychological Association.
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109, 163-203. doi: 10.1037/0033-2909.109.2.163.
- MacLeod, C. M., & McLaughlin, K. (1995). Implicit and explicit memory bias in anxiety: A conceptual replication. *Behaviour Research and Therapy*, 33, 1-14. doi: 10.1016/00057967(94)E0004-3.
- Madden, D. J., Bennett, I. J., Burzynska, A., Potter, G. G., Chen, N., & Song, A. W. (2012). Diffusion tensor imaging of cerebral white matter integrity in cognitive aging. *Biochimica et Biophysica Acta*, 1822, 386-400. doi: 10.1016/j.bbadis.2011.08.003.

- Manard, M., Carabin, D., Jaspar, M., & Collette, F. (2014). Age-related decline in cognitive control: The role of fluid intelligence and processing speed. *BMC Neuroscience*, 15, 7. doi: 10.1186/1471-2202-15-7.
- Markovits, H., & Doyon, C. (2004). Information processing and reasoning with premises that are empirically false: Interference, working memory, and processing speed. *Memory and Cognition*, 32, 592-601. doi: 10.3758/BF03195850.
- Mars, R. B., Debener, S., Gladwin, T. E., Harrison, L. M., Haggard, P., Rothwell, J. C., & Bestmann, S. (2008). Trial-by-trial fluctuations in the event-related electroencephalogram reflect dynamic changes in the degree of surprise. *Journal of Neuroscience*, 28, 12539-12545. doi: 10.1523/JNEUROSCI.2925-08.2008.
- Marschner, A., Mell, T., Wartenburger, I., Villringer, A., Reischies, F. M., & Heekeren, H. R. (2005). Reward-based decision-making and aging. *Brain Research Bulletin*, 67, 382-390. doi: 10.1016/j.brainresbull.2005.06.010.
- Mata, R., Josef, A. K., Samanez-Larkin, G. R., & Hertwig, R. (2011). Age differences in risky choice: A meta-analysis. *Annals of the New York Academy of Sciences*, 1235, 18-29. doi: 10.1111/j.1749-6632.2011.06200.x.
- Mata, R., & Nunes, L. (2010). When less is enough: Cognitive aging, information search, and decision quality in consumer choice. *Psychology and Aging*, 25, 289–298. doi: 10.1037/a0017927.
- Mata, R., Schooler, L. J., & Rieskamp, J. (2007). The aging decision maker: Cognitive aging and the adaptive selection of decision strategies. *Psychology and Aging*, 22, 796–810. doi: 10.1093/acprof:oso/9780199744282.003.0022.
- Mather, M. (2006). *A review of decision-making processes: Weighing the risks and benefits of*

- aging. (L. L. Carstensen & C. R. Hartel, Eds.). Washington, DC: The National Academies Press. doi: 10.1017/CBO9781107415324.004.
- Mather, M., & Carstensen, L. L. (2005). Aging and motivated cognition: The positivity effect in attention and memory. *Trends in Cognitive Sciences*, 9, 496-502. doi: 10.1016/j.tics.2005.08.005.
- Mather, M., & Johnson, M. K. (2000). Choice-supportive source monitoring: Do our decisions seem better to us as we age? *Psychology and Aging*, 15, 596–606. doi: 10.1037/0882-7974.15.4.596.
- Mather, M., & Knight, M. (2005). Goal-directed memory: The role of cognitive control in older adults' emotional memory. *Psychology and Aging*, 20, 554–570. doi: 10.1037/0882-7974.20.4.554.
- Mathewson, K. J., Dywan, J., & Segalowitz, S. J. (2005). Brain bases of error-related ERPs as influenced by age and task. *Biological Psychology*, 70, 88-104. doi: 10.1016/j.biopsycho.2004.12.005.
- Mathewson, K. J., Dywan, J., Snyder, P. J., Tays, W. J., & Segalowitz, S. J. (2008). Aging and electrocortical response to error feedback during a spatial learning task. *Psychophysiology*, 45, 936-948. doi: 10.1111/j.1469-8986.2008.00699.x.
- Matsumoto, K., Suzuki, W., & Tanaka, K. (2003). Neuronal correlates of goal-based motor selection in the prefrontal cortex. *Science*, 301, 229-232. doi: 10.1126/science.1084204.
- McDowd, J. M., & Craik, F. I. (1988). Effects of aging and task difficulty on divided attention performance. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 267–280. doi: 10.1038/ng749.
- McGuire, J. T., Nassar, M. R., Gold, J. I., & Kable, J. W. (2014). Functionally dissociable

- influences on learning rate in a dynamic environment. *Neuron*, 84, 870-881. doi: 10.1016/j.neuron.2014.10.013.
- McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, 37, 1-10. doi: 10.1016/j.intell.2008.08.004.
- McVay, J. C., & Kane, M. J. (2012). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of Experimental Psychology: General*, 141, 302-320. doi: 10.1037/a0025250.
- Miltner, W. H., Braun, C. H., & Coles, M. G. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a “generic” neural system for error detection. *Journal of Cognitive Neuroscience*, 9, 788-798. doi: 10.1162/jocn.1997.9.6.788.
- Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, 21, 8-14. doi: 10.1177/0963721411429458.
- Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin*, 132, 297-326. doi: 10.1037/0033-2909.132.2.297.
- Morcom, A. M., & Johnson, W. (2015). Neural reorganization and compensation in aging. *Journal of Cognitive Neuroscience*, 27, 1275-1285. doi: 10.1162/jocn_a_00783.
- Morse, C. K. (1993). Does variability increase with age? An archival study of cognitive measures. *Psychology and Aging*, 8, 156-164. doi: 10.1037/0882-7974.8.2.156.
- Moutier, S., Plagne-Cayeux, S., Melot, A. M., & Houdé, O. (2006). Syllogistic reasoning and belief-bias inhibition in school children: Evidence from a negative priming paradigm. *Developmental Science*, 9, 166-172. doi: 10.1111/j.1467-7687.2006.00476.x.

- Mutter, S. A., & Pliske, R. M. (1994). Aging and illusory correlation in judgments of co-occurrence. *Psychology and Aging, 9*, 53-63. doi: 10.1037/0882-7974.9.1.53.
- Mussweiler, T., & Strack, F. (1999). Hypothesis-consistent testing and semantic priming in the anchoring paradigm: A selective accessibility model. *Journal of Experimental Social Psychology, 35*, 136–164. doi: 10.1006/jesp.1998.1364.
- Nassar, M. R., Bruckner, R., Gold, J. I., Li, S., Heekeren, H. R., & Eppinger, B. (2016). Age differences in learning emerge from an insufficient representation of uncertainty in older adults. *Nature Communications, 7*, 11609. doi: 10.1038/ncomms11609.
- Nelson, W., Reyna, V. F., Fagerlin, A., Lipkus, I., and Peters, E. (2008). Clinical implications of numeracy: Theory and practice. *Annals of Behavioral Medicine, 35*, 261–274. doi:10.1007/s12160-008-9037-8.
- Nieuwenhuis, S., Ridderinkhof, K. R., Talsma, D., Coles, M. G. H., Holroyd, C. B., Kok, A., & van der Molen, M. W. (2002). A computational account of altered error processing in older age: Dopamine and the error-related negativity. *Cognitive, Affective and Behavioral Neuroscience, 2*, 19-36. doi: 10.3758/CABN.2.1.19.
- Norris, J. E., McGeown, W. J., Guerrini, C., & Castronovo, J. (2015). Aging and the number sense: Preserved basic non-symbolic numerical processing and enhanced basic symbolic processing. *Frontiers in Psychology, 6*, 1-13. doi: 10.3389/fpsyg.2015.00999.
- Nyberg, L., Lovden, M., Riklund, K., Lindenberger, U., & Backman, L. (2012). Memory aging and brain maintenance. *Trends in Cognitive Sciences, 16*, 292–305. doi: 10.1016/j.tics.2012.04.005.
- Nyberg, L., Salami, A., Andersson, M., Eriksson, J., Kalpouzos, G., Kauppi, K.,... Nilsson, L. G. (2010). Longitudinal evidence for diminished frontal cortex function in aging. *Proceedings of*

- the National Academy of Sciences of the United States of America*, 107, 22682-22686. doi: 10.1073/pnas.1012651108.
- O'Connell, R. G., Balsters, J. H., Kilcullen, S. M., Campbell, W., Bokde, A. W., Lai, R., ... Robertson, I. H. (2012). A simultaneous ERP/fMRI investigation of the P300 aging effect. *Neurobiology of Aging*, 33, 2448-2461. doi: 10.1016/j.neurobiolaging.2011.12.021
- Oechssler, J., Roider, A., & Schmitz, P. W. (2009). Cognitive abilities and behavioral biases. *Journal of Economic Behavior and Organization*, 72, 147-152. doi: 10.1016/j.jebo.2009.04.018.
- Oldrati, V., Patricelli, J., Colombo, B., & Antonietti, A. (2016). The role of dorsolateral prefrontal cortex in inhibition mechanism: A study on cognitive reflection test and similar tasks through neuromodulation. *Neuropsychologia*, 91, 499-508. doi: 10.1016/j.neuropsychologia.2016.09.010.
- O'Reilly, J. X., Schüffelgen, U., Cuell, S. F., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2013). Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proceedings of the National Academy of Sciences*, 110, E3660-E3669. doi: 10.1073/pnas.1305373110.
- Osman, M. (2004). An evaluation of dual-process theories of reasoning. *Psychonomic Bulletin & Review*, 11, 988–1010. doi: 10.3758/BF03196730.
- Ostwald, D., Spitzer, B., Guggenmos, M., Schmidt, T. T., Kiebel, S. J., & Blankenburg, F. (2012). Evidence for neural encoding of Bayesian surprise in human somatosensation. *Neuroimage*, 62, 177-188. doi: 10.1016/j.neuroimage.2012.04.050.

- Pacini, R., & Epstein, S. (1999). The relation of rational and experiential information processing styles to personality, basic beliefs and the ratio-bias phenomenon. *Journal of Personality and Social Psychology*, 76, 972–987. doi: 10.1037/0022-3514.76.6.972.
- Park, D. C. (2000). The basic mechanisms accounting for age-related decline in cognitive function. In D. C. Park & N. Schwarz (Eds.), *Cognitive aging: A primer* (pp. 3–21). Philadelphia: Psychology Press.
- Park, D. C., & Reuter-Lorenz, P. (2009). The adaptive brain: Aging and neurocognitive scaffolding. *Annual Review of Psychology*, 60, 173-196. doi: 10.1146/annurev.psych.59.103006.093656.
- Paus, T. (2001). Primate anterior cingulate cortex: Where motor control, drive and cognition interface. *Nature Neuroscience Reviews*, 2, 417-424. doi: 10.1038/35077500.
- Payzan-LeNestour, E., & Bossaerts, P. (2011). Risk, unexpected uncertainty, and estimation uncertainty: Bayesian learning in unstable settings. *PLoS Computational Biology*, 7, e1001048. doi: 10.1371/journal.pcbi.1001048.
- Pennycook, G. (2017). A perspective on the theoretical foundation of dual-process models. In W. De Neys (Ed.), *Dual Process Theory 2.0* (pp. 5-39). London, England: Routledge. doi:10.4324/9781315204550
- Pennycook, G., Cheyne, J. A., Koehler, D. J., & Fugelsang, J. A. (2016). Is the cognitive reflection test a measure of both reflection and intuition? *Behavior Research Methods*, 48, 341-348. doi: 10.3758/s13428-015-0576-1.
- Pennycook, G., & Thompson, V. A. (2012). Reasoning with base-rates is routine, relatively effortless, and context dependent. *Psychonomic Bulletin & Review*, 19, 528–534. doi:10.3758/s13423-012-0249-3.

- Pennycook, G., Trippas, D., Handley, S. J., & Thompson, V. A. (2013). Base rates: Both neglected and intuitive. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40, 544-554. doi: 10.1037/a0034887.
- Perkins, D. N. (1995). *Outsmarting IQ: The emerging science of learnable intelligence*. New York: Free Press.
- Peters, E. (2012). Beyond comprehension: The role of numeracy in judgments and decisions. *Current Directions in Psychological Science*, 21, 31-35. doi: 10.1177/0963721411429960
- Peters, E., & Bjälkebring, P. (2015). Multiple numeric competencies: When a number is not just a number. *Journal of Personality and Social Psychology*, 108, 802-822. doi: 10.1037/pspp0000019.
- Peters, E., Diefenbach, M. A., Hess, T. M., & Västfjäll, D. (2008). Age differences in dual information-processing modes: Implications for cancer decision making. *Cancer*, 113, 3556-3567. doi: 10.1002/cncr.23944.
- Peters, E., Finucane, M., MacGregor, D., & Slovic, P. (2000). The bearable lightness of aging: Judgment and decision processes in older adults. In P. Stern & L. L. Carstensen (Eds.), *The aging mind: Opportunities in cognitive research* (pp. 144 –165). Washington, DC: National Academies Press.
- Peters, E., Hess, T. M., Västfjäll, D., & Auman, C. (2007). Adult age differences in dual information processes: Implications for the role of affective and deliberative processes in older adults' decision making. *Perspectives on Psychological Science*, 2, 1-23. doi: doi.org/10.1111/j.1745-6916.2007.00025.x.

- Peters, E., Hess, T. M., Auman, C., & Västfjäll, D. (2007). Adult age differences in dual information processes and their influence on judgments and decisions: A review. *Perspectives on Psychological Science*, 2, 1-23. doi: 10.1111/j.1745-6916.2007.00025.x.
- Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological Science*, 17, 407-413. doi: 10.1111/j.1467-9280.2006.01720.x.
- Piazza, M., & Izard, V. (2009). How humans count: Numerosity and the parietal cortex. *The Neuroscientist*, 15, 261-273. doi: 10.1177/1073858409333073.
- Piazza, M., & Sousa, P. (2014). Religiosity, political orientation, and consequentialist moral thinking. *Social Psychological and Personality Science*, 5, 334-342. doi: 10.1177/1948550613492826.
- Pietschmann, M., Endrass, T., Czerwon, B., & Kathmann, N. (2011). Aging, probabilistic learning and performance monitoring. *Biological Psychology*, 86, 74-82. doi: 10.1016/j.biopsycho.2010.10.009.
- Pietschmann, M., Simon, K., Endrass, T., & Kathmann, N. (2008). Changes of performance monitoring with learning in older and younger adults. *Psychophysiology*, 45, 67-75. doi: 10.1111/j.1469-8986.2008.00651.x.
- Pires, L., Leitão, J., Guerrini, C., Simões, M. R. (2014). Event-related brain potentials in the study of inhibition: Cognitive control, source localization and age-related modulations. *Neuropsychology*, 24, 461-490. doi: 10.1007%2Fs11065-014-9275-4.
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118, 2128-2148. doi: 10.1016/j.clinph.2007.04.019.

- Poudel, G. R., Bhattarai, A., Dickinson, D. L., and Drummond, S. P. A. (2017). Neural correlates of decision-making during a Bayesian choice task. *NeuroReport*, 28, 193-199. doi: 10.1097/WNR.0000000000000730.
- Qiu, C., & Fratiglioni, L. (2018). Aging without dementia is achievable: Current evidence from epidemiological research. *Journal of Alzheimer's Disease*, 62, 933-942. doi: 10.3233/JAD-171037.
- Rajah, M. M., Bastianetto, S., Bromley-Brits, K., Cools, R., D'Esposito, M., Grady, C. L., ... Pruessner, J. (2009). Biological changes associated with healthy versus pathological aging: A symposium review. *Ageing Research Reviews*, 8, 140–146. doi: 10.1124/dmd.107.016501.CYP3A4-Mediated.
- Raven, J. C. (1982). *Revised Manual for Raven's Progressive Matrices and Vocabulary Scale*. NFER Nelson, Windsor, England.
- Raz, N. (2005). The aging brain observed in vivo: Differential changes and their modifiers. In R. Cabeza, L. Nyberg, & D. L. Park (Eds.), *Cognitive neuroscience of aging* (pp. 19–57). New York: Oxford University Press.
- Reed, A. E., & Carstensen, L. L. (2012). The theory behind the age-related positivity effect. *Frontiers in Psychology*, 2, 1-9. doi: 10.3389/fpsyg.2012.00339.
- Reuter-Lorenz, P. A. (2002). New visions of the aging mind and brain. *Trends in Cognitive Science*, 6, 394-400. doi: 10.1016/S1364-6613(02)01957-5.
- Reuter-Lorenz, P. A., Jonides, J., Smith, E. E., Hartley, A., Miller, A., Marshuetz, C., ... Koeppe, R. A. (2000). Age differences in the frontal lateralization of verbal and spatial working memory revealed by PET. *Journal of Cognitive Neuroscience*, 12, 174–187. doi: 10.1162/089892900561814.

- Reyna, V. F. (2004). How people make decisions that involve risk: A dual-processes approach. *Current Directions in Psychological Science*, 13, 60–66. doi:10.1111/j.0963-7214.2004.00275.x.
- Reyna, V. F., & Brainerd, C. J. (2008). Numeracy, ratio bias, and denominator neglect in judgments of risk and probability. *Learning and Individual Differences*, 18, 89–107. doi:10.1016/j.lindif.2007.03.011.
- Reyna, V. F., & Farley, F. (2006). Risk and rationality in adolescent decision making: Implications for theory, practice, and public policy. *Psychological Science in the Public Interest*, 7, 1–44. doi: 10.1111/j.1529-1006.2006.00026.x.
- Reyna, V. F., & Mills, B. (2007). Converging evidence supports fuzzy trace theory’s nested sets hypothesis, but not the frequency hypothesis. *Behavioral and Brain Sciences*, 30, 278–280. doi:10.1017/ S0140525X07001872.
- 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, 943-973. doi: 10.1037/a0017327.
- Rizio, A. A., & Dennis, N. A. (2014). The cognitive control of memory: Age differences in the neural correlates of successful remembering and intentional forgetting. *PloS One*, 9, e87010. doi: 10.1371/journal.pone.0087010.
- Roggeveen, A. B., Prime, D. J., & Ward, L. M. (2007). Lateralized readiness potentials reveal motor slowing in the aging brain. *Journal of Gerontology, Series B, Psychological Sciences and Social Sciences*, 62, P78-P84. doi: 10.1093/geronb/62.2.p78.
- Salthouse, T. A. (2019). Trajectories of cognitive aging. *Psychology and Aging*, 34, 17-24. doi: 10.1037/pag0000288.

- Salthouse, T. A. (2012). Consequences of age-related cognitive declines. *Annual Review of Psychology*, 63, 201-226. doi: 10.1146/annurev-psych-120710-100328.
- Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, 103, 403-428. doi: 10.1037/0033-295X.103.3.403.
- Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experimental Psychology: General*, 132, 566-594. doi: 10.1037/0096-3445.132.4.566.
- Samanez-Larkin, G. R., Worthy, D. A., Mata, R., McClure, S. M., & Knutson, B. (2014). Adult age differences in frontostriatal representation of prediction error but not reward outcome. *Cognitive, Affective & Behavioral Neuroscience*, 14, 672-682. doi: 10.3758/s13415-014-0297-4.
- Savage, L. J. (1954). *The foundations of statistics*. New York: Dover 1972 edition of the original 1954 John Wiley and Sons.
- Scheffers, M. K., & Coles, M. G. H. (2000). Performance monitoring in a confusing world: Error-related brain activity, judgments of response accuracy, and types of errors. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 141-151. doi: 10.1037/0096-1523.26.1.141.
- Schmiedt-Fehr, C., and Basar-Eroglu, C. (2011). Event-related delta and theta brain oscillations reflect age-related changes in both a general and a specific neuronal inhibitory mechanism. *Clinical Neurophysiology*, 122, 1156-1167. doi: 10.1016/j.clinph.2010.10.045.
- Schwartenbeck, P., FitzGerald, T. H. B., & Dolan, R. (2016). Neural signals encoding shifts in beliefs. *Neuroimage*, 125, 578-586. doi: 10.1016/j.neuroimage.2015.10.067.
- Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H. G. (1997). The role of numeracy in

- understanding the benefit of screening mammography. *Annals of Internal Medicine*, 127, 966–972. doi:10.7326/0003-4819-127-11-199712010-00003.
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275, 1593-1599. doi: 10.1126/science.275.5306.1593.
- Schultz, W., Preuschoff, K., Camerer, C., Hsu, M., Fiorillo, C. D., Tobler, P. N., & Bossaerts, P. (2008). Explicit neural signals reflecting reward uncertainty. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 363, 3801–3811.
- Scott, S. G., & Bruce, R. A. (1995). Decision-making style: The development and assessment of a new measure. *Educational and Psychological Measurement*, 55, 818–831. doi: 10.1177/0013164495055005017.
- Seer, C., Lange, F., Boos, M., Dengler, R., & Kopp, B. (2016). Prior probabilities modulate cortical surprise responses: A study of event-related potentials. *Brain and Cognition*, 106, 78-89. doi: 10.1016/j.bandc.2016.04.011.
- Shenhav, A., Rand, D. G., & Greene, J. D. (2012). Divine intuition: Cognitive style influences belief in God. *Journal of Experimental Psychology: General*, 141, 423-428. doi: 10.1037/a0025391.
- Shiloh, S., Salton, E., & Sharabi, D. (2002). Individual differences in rational and intuitive thinking styles as predictors of heuristic responses and framing effects. *Personality and Individual Differences*, 32, 415-429. doi: 10.1016/S0191-8869(01)00034-4.
- Simon, H. (1972). *Theories of Bounded Rationality*. (M. C & R. R, Eds.), *Decision and Organization*. North-Holland, Amsterdam.
- Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, 119, 3-22. doi: 10.1037/0033-2909.119.1.3.

- Smith, J. L., Johnstone, S. J., & Barry, R. J. (2008). Movement-related potentials in the Go/NoGo task: The P3 reflects both cognitive and motor inhibition. *Clinical Neurophysiology*, *119*, 704-714. doi: 10.1016/j.clinph.2007.11.042.
- Smith, J. L., Johnstone, S. J., & Barry, R. J. (2007). Response priming in the Go/NoGo task: The N2 reflects neither inhibition nor conflict. *Clinical Neurophysiology*, *118*, 343-355. doi: 10.1016/j.clinph.2006.09.027.
- Spaniol, J., & Bayen, U. J. (2005). Aging and conditional probability judgments: A global matching approach. *Psychology and Aging*, *20*, 165–181. doi: 10.1037/0882-7974.20.1.165.
- Spreng, R. N., Wojtowicz, M., & Grady, C. L. (2010). Reliable differences in brain activity between young and old adults: A quantitative meta-analysis across multiple cognitive domains. *Neuroscience and Biobehavioral Reviews*, *34*, 1178-1194. doi: 10.1016/j.neubiorev.2010.01.009.
- Stanovich, K. E. (2018). Miserliness in human cognition: The interaction of detection, override and mindware. *Thinking and Reasoning*, *24*, 423-444. doi: 10.1080/13546783.2018.1459314.
- Stanovich, K. E. (2011). *Rationality and the reflective mind*. New York: Oxford University Press.
- Stanovich, K. E. (2009). Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory? In J. S. B. T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 55-88). New York, NY, US: Oxford University Press.
- Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah, NJ: Erlbaum.
- Stanovich, K. E., & Evans, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, *8*, 223-241. doi:

10.1177/1745691612460685.

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

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, 23, 645-726. doi: 10.1017/CBO9780511808098.026.

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

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

Staudinger, U. M., & Glück, J. Psychological wisdom research: Commonalities and differences in a growing field. *Annual Review of Psychology*, 62, 215-241. doi: 10.1146/annurev.psych.121208.131659.

Steinhauser, M., Hübner, R., & Druey, M. (2009). Adaptive control of response preparedness in task switching. *Neuropsychologia*, 47, 1826-1835. doi:10.1016/j.neuropsychologia.2009.02.022.

Stern, E. R., Gonzalez, R., Welsh, R. C., & Taylor, S. F. (2010). Updating beliefs for a decision: Neural correlates of uncertainty and underconfidence. *Journal of Neuroscience*, 30, 8032-8041. doi: 10.1523/JNEUROSCI.4729-09.2010.

Sternberg, R. J. (1985). Implicit theories of intelligence, creativity, and wisdom. *Journal of Personality and Social Psychology*, 49, 607-627. doi: 10.1037/0022-3514.49.3.607.

- Sterr, A., & Dean, P. (2008). Neural correlates of movement preparation in healthy ageing. *European Journal of Neuroscience*, 27, 254-260. doi: 10.1111/j.1460-9568.2007.05975.x.
- Stigler, S. M. (1986). *The history of statistics: The measurement of uncertainty before 1900*. Cambridge, MA: Harvard University Press.
- Stigler, S. M. (1983). Who discovered Bayes's theorem? *American Statistician*, 37, 290-296. doi: 10.1080/00031305.1983.10483122.
- St-Laurent, M., Abdi, H., Burianova, H., & Grady, C. L. (2011). Influence of aging on the neural correlates of autobiographical, episodic and semantic memory retrieval. *Journal of Cognitive Neuroscience*, 23, 4150–4163. doi: 10.1162/jocn.
- Störmer, V. S., Winther, G. N., Li, S. C., Andersen, S. K. (2013). Sustained multifocal attentional enhancement of stimulus processing in early visual areas predicts tracking performance. *Journal of Neuroscience*, 33, 5346-5351. doi: 10.1523/JNEUROSCI.4015-12.2013.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643-662. doi: 10.1037/h0054651.
- Strough, J., Karns, T. E., & Schlosnagle, L. (2011). Decision-making heuristics and biases across the life span. *Annals of the New York Academy of Sciences*, 1235, 57-74. doi: 10.1111/j.1749-6632.2011.06208.x.
- Summerfield, C., & Koechlin, E. (2008). A neural representation of prior information during perceptual inference. *Neuron*, 59, 336-346. doi: 10.1016/j.neuron.2008.05.021.
- Sutton, R. and Barto, A. (1998). *Reinforcement learning: An introduction*. Cambridge, MA, MIT Press.

- Uddin, L. Q., Nomi, J. S., Hebert-Seropian, B., Ghaziri, J., & Boucher, O. (2017). Structure and function of the human insula. *Journal of Clinical Neurophysiology*, 34, 300-306. doi: 10.1097/WNP.0000000000000377.
- Taleb, N. N. (2010). *The black swan: The impact of the highly improbable*. New York, NY: Random House.
- Tentori, K., Osherson, D., Hasher, L., & May, C. (2001). Wisdom and aging: Irrational preferences in college students but not in older adults. *Cognition*, 81, 87-96. doi: 10.1016/S0010-0277(01)00137-8.
- Thompson, V. A. (2009). Dual-process theories: A metacognitive perspective. In J. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 171-195). Oxford: Oxford University Press.
- Todd, P. M., Gigerenzer, G., & ABC Research Group. (2011). *Ecological rationality: Intelligence in the world*. New York: Oxford University Press.
- Toplak, M. E., 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. doi: 10.1002/bdm.544.
- Toplak, M. E., West, R. F., & Stanovich, K. E. (2014a). Assessing miserly processing: An expansion of the Cognitive Reflection Test. *Thinking and Reasoning*, 20, 147-168. doi: 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. doi: 10.1037/a0034910.

- 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 and Cognition*, 39, 1275-1289. doi: 10.3758/s13421-011-0104-1.
- Traczyk, J., & Fulawka, K. (2016). Numeracy moderates the influence of task-irrelevant affect on probability weighting. *Cognition*, 151, 37-41. doi: 10.1016/j.cognition.2016.03.002.
- Trapp, S., Lepsien, J., Kotz, S. A., & Bar, M. (2016). Prior probability modulates anticipatory activity in category-specific areas. *Cognitive, Affective, & Behavioral Neuroscience*, 16, 135-144. doi: 10.3758/s13415-015-0373-4.
- Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90, 293-315. doi: 10.1037/0033-295X.90.4.293.
- Tversky, A., & Kahneman, D. (1982). Evidential impact of base rates. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 153-160). Cambridge: Cambridge University Press.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1131. doi:10.1126/science.185.4157.1124.
- Tymula, A., Rosenberg Belmaker, L. A., Ruderman, L., Glimcher, P. W., & Levy, I. (2013). Like cognitive function, decision making across the life span shows profound age-related changes. *Proceedings of the National Academy of Sciences*, 110, 17143-17148. doi: 10.1073/pnas.1309909110.
- Vallesi, A. (2011). Targets and non-targets in the aging brain: A go/nogo event-related potential study. *Neuroscience Letters*, 487, 313-317. doi: 10.1016/j.neulet.2010.10.046.

- Vallesi, A., McIntosh, A. R., & Stuss, D. T. (2011). Overrecruitment in the aging brain as a function of task demands: Evidence for a compensatory view. *Journal of Cognitive Neuroscience*, 23, 801-815. doi: 10.1162/jocn.2010.21490.
- Vallesi, A., & Stuss, D. T. (2010). Excessive sub-threshold motor preparation for non-target stimuli in normal aging. *NeuroImage*, 50, 1251-1257. doi: 10.1016/j.neuroimage.2010.01.022.
- Vallesi, A., Stuss, D. T., McIntosh, A. R., & Picton, T. W. (2009). Age-related differences in processing irrelevant information: Evidence from event-related potentials. *Neuropsychologia*, 47, 577-586. doi: 10.1016/j.neuropsychologia.2008.10.018.
- Vilares, I., Howard, J. D., Fernandes, H. L., Gottfried, J. A., & Kording, K. P. (2012). Differential representations of prior and likelihood uncertainty in the human brain. *Current Biology*, 22, 1641-1648. doi: 10.1016/j.cub.2012.07.010.
- Volz, K. G., & Gigerenzer, G. (2012). Cognitive processes in decisions under risk are not the same as in decisions under uncertainty. *Frontiers in Neuroscience*, 6, 105. doi: 10.3389/fnins.2012.00105.
- Vossel, S., Mathys, C., Stephan, K. E., & Friston, K. J. (2015). Cortical coupling reflects Bayesian belief updating in the deployment of spatial attention. *Journal of Neuroscience*, 35, 11532-11542. doi: 10.1523/JNEUROSCI.1382-15.2015.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54, 1063-1070. doi: 10.1037/0022-3514.54.6.1063.
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale*. 3rd ed. San Antonio, Texas: Psychological Corp.

- Wegier, P., Armstrong, B. A., & Shaffer, V. A. (2019). Aiding risk information learning through simulated experience (ARISE): A comparison of the communication of diagnostic screening test information in explicit and simulated experience formats. *Medical Decision Making*, 39, 196-207. doi:10.1177/0272989X19832882.
- Weiler, J. A., Bellebaum, C., & Daum, I. (2008). Aging effects acquisition and reversal of reward-based associative learning. *Learning and Memory*, 15, 190-197. doi: 10.1101/lm.890408.
- Weller, J. A., Dieckmann, N. F., Tusler, M., Mertz, C. K., Burns, W. J., & Peters, E. (2013). Development and testing of an abbreviated numeracy scale: A Rasch analysis approach. *Journal of Behavioral Decision Making*, 26, 198-212. doi: 10.1002/bdm.1751.
- West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin*, 120, 272-292. doi: 10.1037/0033-2909.120.2.272.
- West, R. L. (2004). The effects of aging on controlled attention and conflict processing in the Stroop task. *Journal of Cognitive Neuroscience*, 16, 103-113. doi: 10.1162/089892904322755593.
- West, R., & Alain, C. (2000). Age-related decline in inhibitory control contributes to the increased Stroop effect observed in older adults. *Psychophysiology*, 37, 179-189. doi: 10.1111/1469-8986.3720179.
- Westbrook, A., Kester, D., & Braver, T. S. (2013). What Is the subjective cost of cognitive effort? Load, trait, and aging effects revealed by economic preference. *PLoS ONE*, 8, 1–8. doi: 10.1371/journal.pone.0068210.
- Wolfe, C. R., & Fisher, C. R. (2013). Individual differences in base rate neglect: A fuzzy processing preference index. *Learning and Individual Differences*, 25, 1–11.

doi:10.1016/j.lindif.2013.03.003.

- Worthy, D. A., Gorlick, M. A., Pacheco, J. L., Schnyer, D. M., & Maddox, T. W. (2011). With age comes wisdom: Decision-making in younger and older adults. *Psychological Science*, 22, 1375-1380. doi: 10.1177/0956797611420301.
- Yates, J. F., & Patalano, A. L. (1999). Decision making and aging. In D. C. Park, R. W. Morrell, & K. Shifren (Eds.), *Processing of medical information in aging patients: Cognitive and human factors perspectives* (pp. 31-54). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Yoon, C., Cole, C. A., & Lee, M. P. (2009). Consumer decision making and aging: Current knowledge and future directions. *Journal of Consumer Psychology*, 19, 2–16. doi: 10.1016/j.jcps.2008.12.002.
- Zamarian, L., Sinz, H., Bonatti, E., Gamboz, N., & Delazer, M. (2008). Normal aging affects decisions under ambiguity, but not decisions under risk. *Neuropsychology*, 22, 645–657. doi: 10.1037/0894-4105.22.5.645.