THE ROLE OF MEMORY AND IDEOLOGICAL BIASES IN THE CORRECTION OF MISINFORMATION

Briony Swire-Thompson
Bachelor of Arts (Psychology), Honours

This thesis is presented for the degree of
Doctor of Philosophy of the University of Western Australia
School of Psychological Science
2017
Thesis Declaration

I, Briony Swire-Thompson, certify that:

This thesis has been substantially accomplished during enrolment in the degree.

This thesis does not contain material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution.

No part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of The University of Western Australia and where applicable, any partner institution responsible for the joint-award of this degree.

This thesis does not contain any material previously published or written by another person, except where due reference has been made in the text.

The work(s) are not in any way a violation or infringement of any copyright, trademark, patent, or other rights whatsoever of any person.

The research involving human data reported in this thesis was assessed and approved by The University of Western Australia Human Research Ethics Committee. Approval #: [RA/4/1/6289] and Massachusetts Institute of Technology Institutional Review Board [COUHES Protocol 1603512596].

Written patient consent has been received and archived for the research involving patient data reported in this thesis.

This thesis contains published work and/or work prepared for publication, some of which has been co-authored.

This thesis was funded by a University of Western Australia Postgraduate Award, a University of Western Australia Safety-Net Top-Up scholarship and a Postgraduate America-Australia Fulbright Scholarship.

This research was supported by an Australian Government Research Training Program (RTP) Scholarship.

Signature:

Date: 25/06/2017
“[Memory] selects, eliminates, alters, exaggerates, minimizes, glorifies, and vilifies also; but in the end it creates its own reality, its heterogeneous but usually coherent version of events; and no sane human being ever trusts someone else's version more than his own.”

Salman Rushdie, *Midnight's Children*
Abstract

Once an individual concludes that information is true, it is remarkably difficult to change their mind, even if valid counter-evidence is presented. This is known as the continued influence effect of misinformation. Persisting belief in inaccurate information can have serious ramifications, as our convictions regarding what is fact or fiction greatly shape how we see the world, and ultimately impact our decisions. These convictions thus have the potential to influence how we dedicate our time, what we purchase, and who we vote for. If we are to provide sound, evidence-based recommendations for how to ameliorate the continued influence effect, we need a better understanding of the cognitive mechanisms underlying it. To this end, the empirical chapters of this thesis address three questions: (1) Does the familiarity of the misinformation play a role in the continued influence effect? (2) Are some correction formats better at promoting belief change than others? (3) Does credibility of the information’s source influence how the misinformation or the correction is processed?

The thesis first delves into memory processes and investigates whether the misinformation’s familiarity is a contributing factor to the continued influence effect. As familiarity-based acceptance of misinformation is most likely to occur in the absence of strategic memory processes, the experiments described in Chapter 3 examined factors known to affect whether strategic memory processes can be utilised. These factors were age, the amount of detail provided with the correction, and the amount of time between encoding and retrieval attempt. We found that older adults are worse at sustaining belief change after learning misinformation is false, and that a greater level of explanatory detail promoted more sustained belief change.
Moreover, we found that corrections of misinformation over the long term were significantly less effective than affirmations of facts, supporting the notion that information’s familiarity does contribute to misinformation effects.

Second, this thesis explores whether corrections can be designed to optimise their corrective impact, and whether repeating the misinformation within the retraction hinders its corrective efficacy. This is a question of both practical and theoretical significance. In the experiment described in Chapter 4, we investigated the efficacy of four explanation formats. Results indicated that the type of explanation does not dramatically predict whether belief change is sustained. It was found that repeating the original misinformation within the retraction was not detrimental to belief updating, contrary to some claims found in the literature. It can be concluded that the frame—i.e., how the retraction is formatted—is unlikely to have a strong impact on the retraction’s efficacy.

Even if an ideal corrective format were to exist, people do not process information in an ideological vacuum—deeply held beliefs regarding what should be true about the world inevitably come into play. In the final empirical chapter, this thesis uses political stimuli to examine biases that contribute to the continued influence effect, namely motivated reasoning and differences in perceived source credibility. Specifically, the experiments reported in Chapter 5 investigated the assessment of veracity when the information stems from a polarising source. We found that people are more likely to believe information if it comes from a trusted political candidate. Moreover, even if belief in misinformation was reduced post-correction, voting preferences for the political candidate spreading the misinformation remained stable. This suggests that people do not insist on veracity as a prerequisite for supporting political candidates.
This thesis thus aims to demonstrate why people continue to believe in inaccurate information, how corrections can be designed to maximise impact, and whether certain individuals are predisposed to refrain from belief change. The results of this thesis contribute to the broader theoretical conceptualisations regarding memory and belief, as well as practical applications as to how to best correct misinformation.
“That's what the world is, after all: an endless battle of contrasting memories.”

Haruki Murakami, *IQ84*
Table of Contents

Thesis Declaration........................................................................................................i
Abstract.......................................................................................................................... iii
Table of Contents ........................................................................................................... vii
List of Figures ................................................................................................................ xi
List of Tables .................................................................................................................... xiii
Acknowledgements ....................................................................................................... xv
Publications Included in the Thesis................................................................................ xvii
Statement of Candidate Contribution.............................................................................. xix
Preamble ........................................................................................................................... xxi
CHAPTER 1: OVERVIEW........................................................................................................ 1
  1.1 Introduction.................................................................................................................. 2
  1.2 Thesis Outline ............................................................................................................. 6
  1.3 References................................................................................................................... 8

CHAPTER 2: CORRECTING MISINFORMATION—A CHALLENGE FOR EDUCATION
AND COGNITIVE SCIENCES.............................................................................................. 13
  2.1 Introduction.................................................................................................................. 14
  2.2 Misinformation Effects and Individual-Level Cognition............................................ 16
      2.2.1 Retractions and the Provision of Alternatives—The Continued Influence Effect ..... 17
      2.2.2 A Dual-Process Account of Misinformation Effects ........................................... 20
      2.2.3 Why Retractions Fail: The Effects of Repetition and Familiarity ....................... 24
  2.3 Misinformation Effects and Social Cognition............................................................ 29
      2.3.1 The Role of Attitudes for Misinformation Effects.............................................. 30
      2.3.2 Motivated Reasoning and Misinformation Effects............................................ 31
  2.4 Correcting Misinformation ......................................................................................... 36
  2.5 References................................................................................................................... 42

CHAPTER 3: THE ROLE OF FAMILIARITY IN CORRECTING INACCURATE
INFORMATION.................................................................................................................... 55
  3.1 Foreword..................................................................................................................... 56
  3.2 Abstract...................................................................................................................... 59
  3.3 Introduction.................................................................................................................. 60
      3.3.1 The Illusory Truth Effect..................................................................................... 61
      3.3.2 Strategic and Automatic Memory Processes...................................................... 61
      3.3.3 The Familiarity Backfire Effect......................................................................... 63
      3.3.4 Factors Likely to Influence the Correction of Information: Detail, Time, and Age .. 65
  3.4 Experiment 1 ................................................................................................................ 67
      3.4.1 Method................................................................................................................. 68
      3.4.2 Results ............................................................................................................... 73
  3.5 Experiment 2 ................................................................................................................ 78
      3.5.1 Method................................................................................................................. 78
      3.5.2 Results ............................................................................................................... 79
  3.6 General Discussion ..................................................................................................... 85
      3.6.1 Differential Forgetting of Myths and Facts Over Time .................................. 86
List of Figures

Figure 1. Post-manipulation belief ratings over time in Experiment 1. Dotted lines indicate the pre-manipulation belief ratings’ means............................................................74

Figure 2. Post-manipulation inference scores in Experiment 1 over time........................77

Figure 3. Post-manipulation belief ratings in older adults in Experiment 2. Dotted lines indicate the pre-manipulation belief ratings’ mean..................................................80

Figure 4. Post-manipulation myth belief ratings presented in an age-based median split in the older adult sample.....................................................................................................83

Figure 5. Post-manipulation inference scores in an older adult population in Experiment 2.....84

Figure 6. Belief scores in Experiment 3 over time by explanation formats. Dotted lines illustrate control condition.....................................................................................................116

Figure 7. Inference scores in Experiment 3 over time by explanation formats. Dotted lines illustrate control condition......................................................................................118

Figure 8. Design schematic of Experiment 4 .........................................................................142

Figure 9. Pre-explanation Democratic and Republican belief in statements associated with Trump or presented unattributed. Error bars denote 95% confidence intervals ...........147

Figure 10. Belief in Trump and unattributed misinformation and facts over time, across Trump support groups and source conditions. Rep = Republican, Misinfo = Misinformation. Dotted lines show misinformation items. Error bars denote 95% confidence intervals ......................................................................................148

Figure 11. Accuracy scores—misinformation scores subtracted from fact scores— across Trump support and source. Error bars denote 95% confidence intervals ..................149

Figure 12. Influence of Trump—Trump-attributed scores subtracted from unattributed scores— across Trump support. Misinfo = Misinformation. Rep = Republican. Dotted lines show misinformation items .....................................................................................151

Figure 13. Likelihood-to-vote ratings over time between Trump support and source. Error bars denote 95% confidence intervals. Rep = Republican, Misinfo = Misinformation.153

Figure 14. Design schematic of Experiment 5 .........................................................................157

Figure 15. Pre-explanation Democratic and Republican belief in statements associated with Trump. Error bars denote 95% confidence intervals ..............................................159
Figure 16. Belief in Trump misinformation and facts after partisan explanations. Rep = Republican, Misinfo = Misinformation. Dotted lines show misinformation items. Error bars denote 95% confidence intervals ................................................................. 160

Figure 17. Likelihood of voting for Trump across Trump support groups. Error bars denote 95% confidence interval ................................................................. 162

Figure 18. Feelings-thermometer scores over time across partisanship and sources. Error bars denote 95% confidence intervals ................................................................. 253
List of Tables

Table 1. Example of a myth and fact, corresponding explanations, and inference questions..........................................................................................................................70

Table 2. Number of myth and fact items presented in each explanation condition..............112

Table 3. Example of standard format myth retraction, reverse order retraction and facts-only frame .........................................................................................................................114

Table 4. F and p values of myth condition belief ratings planned comparisons ..................117

Table 5. F and p values of fact condition belief ratings planned comparisons ......................118

Table 6. F and p values of myth condition inference score planned comparisons ..................119

Table 7. F and p values of fact condition inference score planned comparisons ..................120

Table 8. Examples of Trump and unattributed misinformation with corresponding corrections ........................................................................................................................144

Table 9. Items and their associated brief and detailed explanations ......................................214

Table 10. Items and their associated inference questions ........................................................228

Table 11. Claims and their associated brief and detailed explanation ....................................237

Table 12. Trump and unattributed items and their corresponding explanations ....................248
Acknowledgements

I would first like to thank my supervisors, Ullrich Ecker and Stephan Lewandowsky. I’m extremely grateful for all the support during the complex PhD process, and for consistently providing the fastest feedback ever known in academic history! Ullrich has unceasingly provided insightful recommendations not only for how to improve my work, but how to navigate the world of academia in general. Stephan, thank you for making your first-year psychology lecture series in 2003 so thrilling. Even though I was one of 600 students in the lecture hall, I genuinely attribute these to be the reason I chose cognitive psychology as my field of research.

A significant thanks to my host supervisor, Adam Berinsky, for welcoming me so wholeheartedly into the Massachusetts Institute of Technology community. Our meetings almost exclusively consisted of walks around the Charles River brainstorming experiments and discussing exciting psychological and political theory. This experience would not have been possible without the Australian-American Fulbright Commission. I am immensely grateful to the commission for sending me on the Fulbright adventure, and for enabling the three-continent collaboration between Ullrich, Stephan, Adam, and I.

I would like to acknowledge the CogSciLab members, in particular our laboratory manager, Charles Hanich, for his friendship and for being my voice when I had tonsillitis. I’d additionally like to thank Carmen Lawrence and Mark Hurlstone for their advice during the candidature review and panel meetings. Furthermore, the psychological office staff have been tremendously helpful, namely Maria del Pilar Puerta Francos, for assistance when navigating this PhD from abroad.
These years of study wouldn’t have been half as exhilarating without the Sanders community. This friendship group has had many names over the years, the “friendship bubble”, the “warm hug”, and “super-best-friends”, all of which are incredibly fitting. A special thanks to Jessica Moncrieff-Boyd, Isabel Rossen, Susie Wang, Alice Mason, Zenobia Talati, Louise Delane, Stephanie Wade, Stephanie Whitworth, Gareth Lee, Sarah George, and Shraddah Kashyap. The challenging components of this doctorate were always more than compensated by the fun of writing retreats, psychology run club outings, cake-day initiatives, social experiments, and copious amounts of tea. I’m so grateful to be part of such a thought-provoking, inspiring, hilarious group.

I would like to express my immense thanks to my parents. Both Sue and Jon Swire-Thompson have shared my excitement from the first day of enrolment through to completion. A massive thank you also goes to my brother, Rob Swire, for years of sympathetic conversations regarding how to be productive in a framework that necessitates self-motivation—it turns out musicians have incredibly similar tribulations as scientists. Finally, I’m extremely grateful for the unwavering support of my partner, Joe DeGutis. Thank you for cheering me across the finish line, and for showing me that research becomes even more fun on the other side of a PhD!
Publications Included in the Thesis

Publication 1: Published Book Chapter (Chapter 2).


Publication 2: Published paper (Chapter 3)


Publication 3: Unsubmitted paper (Chapter 4)


Publication 4: Published paper (Chapter 5)

Statement of Candidate Contribution

Publication 1: 40% (Chapter 2).

The candidate reviewed literature and contributed to the preparation of the manuscript. Associate Professor Ullrich Ecker and Professor Stephan Lewandowsy provided feedback and also contributed to literature review and manuscript preparation.

Publication 2: 80% (Chapter 3)

The candidate completed the literature review, designed the study, programmed the survey, analysed the data, completed the manuscript preparation and completed the manuscript revisions. Associate Professor Ullrich Ecker and Professor Stephan Lewandowsy provided guidance and feedback on the design, analysis, and manuscript preparation.

Publication 3: 90% (Chapter 4)

The candidate completed the literature review, designed the study, programmed the survey, analysed the data, completed the manuscript preparation and completed the manuscript revisions. Associate Professor Ullrich Ecker and Professor Stephan Lewandowsy provided guidance and feedback on the design, analysis, and manuscript preparation.
Publication 4: 80% (Chapter 5)

The candidate completed the literature review, designed the study, programmed the survey, analysed the data, completed the manuscript preparation and completed the manuscript revisions. Associate Professor Ullrich Ecker, Professor Adam Berinsky, and Professor Stephan Lewandowsy provided guidance and feedback on the design, analysis, and manuscript preparation.

Publication 5: 95% (Chapter 6)

The candidate completed the literature review and manuscript preparation, and Associate Professor Ullrich Ecker provided guidance and feedback.

Student signature:  
Date: 21/06/2017

On behalf of all co-authors, I, Associate Professor Ullrich Ecker certify that the student statements regarding their contributing to each of the works listed above are correct.

Coordinating Supervisor signature:  
Date: 25/06/2017
Preamble

This thesis is presented as a series of empirical papers and theoretical chapters that have been published or have been written in manuscript format. Two empirical papers have been published in peer-reviewed journals (Chapter 3 and Chapter 5). One paper has not been submitted (Chapter 4), but has been written in the format of a manuscript. The introductory theoretical review has been published as a book chapter (Chapter 2), and the summary of major findings and recommendations is currently in press as a book chapter (Chapter 6). To facilitate review, literature references for in-text citations can be found directly after each chapter. As the manuscripts were written as independent works, some definitions and themes are presented multiple times throughout the thesis.
Chapter 1: Overview
1.1 Introduction

With “post-truth” being declared the word of the year in 2016 (Fox, 2016), and concepts such as “alternative facts” becoming front page news (Blake, 2017), misinformation has seemingly never been more topical. In the present context, we define misinformation as information that is initially assumed to be valid but subsequently turns out to be false (Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012). Due to technological advances such as the internet and social media, we now live in an age overflowing with information. This increased accessibility to information has undoubtedly had positive impacts upon society (Kaplan & Haenlein, 2010), yet an unfortunate by-product of this information accessibility has been that there is also a greater quantity of inaccurate information in the public sphere. While misinformation is by no means a new phenomenon, historical barriers to information access such as literacy, language, or availability of texts are fast diminishing (Roser & Prtiz-Ospina, 2016), and one can now easily find support for almost any preconceived belief on the internet with a few clicks of the computer mouse.

An added complication is that even when people are presented with evidence-based corrections, inaccurate beliefs are extraordinarily difficult to correct. This is known as the continued influence effect of misinformation (Johnson & Seifert, 1994; Seifert, 2002), and it can have serious ramifications. For example, despite an abundance of evidence-based research countering the baseless claim that the measles mumps rubella (MMR) vaccine causes autism, vaccination rates have been lower than desired due to this misconception (Madsen et al., 2002; Poland & Spier, 2010). In May 2017, this led to more measles cases recorded in Minnesota alone than in the entire USA the previous year (Howard, 2017). The present thesis explores why attempts to correct inaccurate information often fail. We examine two elements that
are theoretical underpinnings of the continued influenced effect, viz. (1) processes that dictate how people remember information and corrections, and (2) how people’s group identities and beliefs influence information processing.

On the one hand, we can look to the memory processes governing information retrieval to understand the continued influence effect. For example, a common (though not universal) assumption among memory researchers is that there are two separate types of memory retrieval, strategic and automatic (Yonelinas, 2002). Based on such a dual-process model, one potential explanation of the continued influence effect is that it is likely to occur when misinformation has been automatically activated but strategic memory processes have failed (Ecker, Lewandowsky, & Tang, 2010). To illustrate, a person may encounter a myth that may automatically trigger a feeling of familiarity, and familiar information in general is more likely to be accepted as valid (e.g., Begg, Anas, & Farinacci, 1992). The person may thus accept the myth as true, unless they are able to use strategic processes to also recollect previously encountering a correction of the myth.

It has even been proposed that the act of boosting the misinformation’s familiarity by repeating the misconception within a correction could lead to a backfire effect—this is where a correction ironically leads a person to believe even more in the misinformation that the correction is aiming to rectify (Cook & Lewandowsky, 2011; Lewandowsky et al., 2012). However, this well-cited notion has not actually been thoroughly tested, and the original claims were made in an unpublished manuscript (Skurnik, Yoon, & Schwarz, 2007; as cited in Schwarz, Sanna, Skurnik, & Yoon, 2007).

1 Henceforth, the term “myth” is used to refer to pieces of misinformation encountered in the real world, such as the false claim that the MMR vaccine cause autism.
Within the dual-process framework, the role of familiarity for the continued influence effect can be investigated in two ways: by manipulating the familiarity of the misinformation, or by manipulating strategic memory processes. In previous work using the former approach, we have manipulated the number of times a misconception is repeated within a correction text, finding tentative evidence that a larger number of corrections is beneficial even if each correction repeats the misinformation (Ecker, Lewandowsky, Swire, & Chang, 2011). In addition, it seems that misinformation that is explicitly stated is easier to correct than if the misinformation is only implied (Ecker, Hogan, & Lewandowsky, 2017; Rich & Zaragoza, 2016). The experiments presented in Chapter 3 of the present thesis used the latter approach of experimentally testing the relevance of myth familiarity for the continued influence effect by manipulating factors known to affect strategic memory processes. In previous work, tentative evidence was found that bolstering strategic memory processes can reduce the continued influence effect (Ecker et al., 2010). Vice versa, if familiarity is indeed a contributing factor, then reducing strategic memory processes should lead to greater continued influence effects and perhaps even a familiarity backfire effect.

Apart from the basic retrieval mechanisms underlying the continued influence effect, research has only just begun to investigate the socio-cognitive factors that may contribute to the effect. There is a tapestry of cognitive biases that can bring about motivated reasoning and distorted perceptions of source credibility. Motivated reasoning is when people evaluate information more critically when the information is counter to their beliefs (Kunda, 1990; Kraft, Lodge, & Taber, 2015). As such, motivated reasoning is associated with the content of the information, while biases regarding source credibility are more concerned with where the information comes
from. It is not fully understood how these factors impact on the processing of misinformation and corrections, and how this affects beliefs and behaviours. This thesis specifically examines the impact of source credibility on the assessment of veracity when information comes from a polarising source, as well as the effectiveness of explanations when they come from one’s own political party or an opposition party. By keeping the content of the initial information and explanations stable across conditions, we were able to provide a novel and accurate measure of a source’s impact upon misinformation processing.

The body of work contained in this thesis extends the current misinformation literature in two important ways. Firstly, the measurement of misinformation effects has largely focused upon fictitious scenarios that retract one key piece of misinformation (Ecker et al., 2011, Johnson & Seifert, 1994; Wilkes and Leatherbarrow, 1988), or it has focused upon one topic of real-world misinformation (cf. Schwarz et al, 2007; Nyhan & Reifler, 2010). As the central concern of this thesis is both with the underlying mechanisms and the implications of the continued influence effect, all experiments utilised multiple real-world stimuli from various topics in an attempt to measure constructs in a manner that would allow for greater generalisation and application.

Secondly, experiments investigating misinformation effects are traditionally fairly short, often providing the misinformation, retracting the misinformation, and measuring misinformation belief within the one experimental session. The interval between the encoding of misinformation and the measurement of participants’ post-correction beliefs or inferential reasoning typically ranges from 5 minutes to 40 minutes (Ecker et al., 2010; Ecker et al., 2011, Ecker, Lewandowsky, Cheung, & Maybery, 2015; Johnson & Seifert, 1994), with some studies neglecting to mention
the interval at all (e.g., van Oostendorp & Bonebakker, 1999; Nyhan & Reifler, 2015). While educational literature has measured post-correction belief after a longer period of time, this has largely been in primary school children (Guzzetti, 1993). In the adult misinformation literature, only a very recent study by Rich, Van Loon, Dunlosky, and Zaragoza (2017) has measured post-correction belief one week after a correction. The current studies aimed to systematically track belief at multiple post-correction time points (i.e. immediately, 30-minutes, one week, and three weeks post-correction) to gain a better understanding of how belief changes are sustained over time.

The results of this thesis thus advance the broader theoretical conceptualisation of continued-influence misinformation effects. In particular, it examines the contribution of various memory and reasoning processes to these effects over time. It also provides a novel investigation into the assessment of information veracity given the credibility of the source, as well as fact-checking effects on existing attitudes and voter behaviour. This research is essential so that we can give practical advice regarding how best to remediate misinformation’s persistence.

1.2 Thesis Outline

This thesis consists of seven chapters. Following this introductory chapter, Chapter 2 provides theoretical context and rationale for the empirical chapters that follow. It discusses misinformation effects at both the individual-level and the level of social cognition. Dual-processing accounts are introduced, which is relevant background for Chapter 3 and 4, and motivated reasoning is discussed, which is applicable to Chapter 5.
The three empirical chapters—Chapters 3, 4, and 5—are presented in journal-article manuscript formats. Chapter 3 consists of two pilot studies and two large-scale experiments. It investigates the parameters of familiarity’s hindrance on corrective information in an undergraduate and older adult population, and whether affirmations of facts are more effective than retractions of myths. Chapter 4 investigates the efficacy of different correction formats over a one and three-week period. Chapter 5 shifts to the political realm in order to investigate the impact of pre-existing attitudes on the assessment of information veracity. To this effect, two experiments examine the impact of misinformation corrections on people’s beliefs and voting intentions. This is the final empirical chapter.

Chapter 6 subsequently pairs our current knowledge of cognitive mechanisms with practical recommendations. To conclude, Chapter 7 deliberates thesis inconsistencies, and the wider implications of this thesis are collectively considered.
1.3 References


Chapter 2: Correcting Misinformation—A Challenge for Education and Cognitive Sciences
2.1 Introduction

Jenny McCarthy is a popular and successful U.S. TV show host, actress, and author. She is also the mother of a child diagnosed with autism, which has inspired her to become an activist and serve on the board-of-directors of Generation Rescue, an organisation dedicated to informing the public about the recovery of children with autism spectrum disorders, and the presumed risks of vaccinations.

In March 2012, she wrote: “MMR [i.e., the common measles-mumps-rubella vaccine], by far, has been the vaccine most commonly cited by parents as a trigger for a regression into autism” (McCarthy, 2012). This argument concerning the cause of autism originates from a 1998 publication in the medical journal The Lancet, which suggested a relationship between the MMR vaccine and the onset of autism (Wakefield et al., 1998).

Celebrities commonly use their popularity to rally for worthy causes. Angelina Jolie raises awareness of humanitarian crises and refugee issues, Oprah Winfrey promotes education for disadvantaged girls in Africa and the U.S., Salma Hayek condemns violence against women and promotes UNICEF vaccination programs. So why did we choose McCarthy’s example to open a chapter on misinformation? What’s wrong with McCarthy’s activism?

As it turns out, quite a few things. First and foremost, there is a strong consensus in the medical science community that there is no causal link between the MMR vaccine (or any other vaccine) and autism. This consensus is based on exhaustive and widely published global research efforts to replicate the Wakefield et
al. (1998) findings. For example, a retrospective cohort study by Madsen et al. (2002) reviewed all children born in Denmark between January 1991 and December 1998, and concluded that there was no increase in the incidence of autism for vaccinated as opposed to unvaccinated children. In a review paper, DeStefano and Thompson (2004) concluded that “evidence now is convincing that the measles-mumps-rubella vaccine does not cause autism or any particular subtypes of autistic spectrum disorder.” (DeStefano & Thompson, 2004, p. 19). In fact, it has become clear that the Wakefield et al. (1998) study was a fraud. The Lancet officially retracted the article in 2010 (Editors of the Lancet, 2010), and the UK’s General Medical Council found the lead author guilty of professional misconduct and removed him from the medical register. Hence, McCarthy’s public claims represent misinformation.

The second issue regards the use of one’s celebrity status to comment on scientific questions without being an expert on those issues. This is relevant because the public is demonstrably receptive to non-expert opinions. For example, roughly a quarter of survey respondents in the U.S. placed ‘some’ (24 %) or even ‘a lot’ (2 %) of trust in the opinions of celebrities on vaccine safety (Freed, Clark, Butchart, Singer, & Davis, 2011). In the present context, this affirms the likelihood that McCarthy’s claims will affect people despite the plethora of evidence that shows those claims to be wrong.

The fact that people receive much of their information from potentially unreliable sources such as celebrities, popular “infotainment” TV shows and non-expert websites poses a problem in particular because it is known that misinformation continues to exert an influence on people’s opinions even after it has been retracted. To illustrate, in 2002, when it had become clear and widely
publicised that the Wakefield et al. (1998) findings were invalid, 24% of mothers in the UK still erroneously considered the vaccine a greater risk than the disease it was preventing (Smith & Yarwood, 2007). Even after a decade of rectification—efforts by doctors, scientists, health agencies, and some media outlets to correct the misconceptions—as well as events such as the vaccine-preventable mumps epidemic in the UK in 2005, rates of immunisation in the UK still had not returned to the level of 1996-1997 (Health Protection Agency, 2011). Clearly, erroneous information can have adverse consequences for public health and society at large, despite subsequent corrections and widespread efforts to disseminate the correct information.

In this chapter, we outline how retracted misinformation still influences people’s reasoning, and why misinformation can persist in society, being surprisingly difficult to eradicate. We discuss how attempts to correct misinformation can under some circumstances even worsen the problem, but we will also explore ways in which misinformation can be useful for educational purposes. We use the term ‘misinformation’ for any information that is (or might reasonably be) believed to be valid when first acquired, but which is subsequently shown to be incorrect or superseded by updated information.

2.2 Misinformation Effects and Individual-Level Cognition

In contemporary information-driven societies we are confronted with a myriad of information sources, and it can be difficult to distinguish fact from fiction. Considering how much information we process and how quickly the world can change, people actually do a remarkable job at updating their knowledge and mental models of the world—hardly anyone would think George W. Bush was still U.S. President, and we can generally keep track of the things we have done versus those that are still on our to-do list. However, human memory is faced with the conundrum
of maintaining stable memory representations (which is the whole point of having a memory in the first place) while also allowing for flexible modulation of memory representations to keep up-to-date with reality. Memory has evolved to achieve both of these aims, and hence it does not work like a blackboard: Out-dated things are rarely actually wiped out and over-written; instead, they tend to linger in the background, and access to them is only gradually lost (De Beni & Palladino, 2004; Oberauer & Vockenberg, 2009).

2.2.1 Retractions and the Provision of Alternatives—The Continued Influence Effect

The degree to which out-dated information lingers depends on the likelihood that it will become relevant again in the future. Hence, when people are provided with a plausible alternative to replace the original misinformation, memory updating is usually successful and people no longer rely on the initial, later retracted, information (H. Johnson & Seifert, 1994; Ecker, Lewandowsky, & Tang, 2010). Once you understand that your parents sneakily exchanged the milk teeth under your pillow for dollar coins, you no longer have a reason to believe in the tooth fairy.

However, if the valid alternative is unknown, or difficult to understand, misinformation will prevail. In the case of autism, the retraction of Wakefield et al.’s (1998) claims might have been more successful had the actual causes of autism been better understood and offered to the general public as an alternative account. A retraction creates a gap in a person’s mental model of the world, and without a plausible alternative account, people may continue to make use of retracted misinformation, preferring an out-dated, possibly incorrect model of the world to an incomplete model of the world. Hence, plain retractions—simply stating that something is incorrect—are usually rather ineffective in correcting misinformation (Ecker, Lewandowsky, Swire, & Chang, 2011a; Ecker et al., 2010; H. Johnson &
People’s ongoing reliance on corrected misinformation is known in the literature as the ‘continued influence effect’ (H. Johnson & Seifert, 1994; see Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012, for a review).

Classic laboratory studies by Wilkes and Leatherbarrow (1988) and H. Johnson and Seifert (1994) have demonstrated that people rely on information after it is no longer relevant (see also Ecker, Lewandowsky, & Apai, 2011b; Ecker et al., 2010). In these studies, participants were given a transcript of a fictitious event, such as a warehouse fire. As the story is being told, negligent storage of volatile materials (oil paint and pressurised gas cylinders) is originally imputed as a likely cause of the fire. Similar to a news story changing as new information comes to light, some participants were then told that actually no such materials had been found. When participants were asked about the presence of corrections in the report, they acknowledged the absence of oil paint and gas cylinders. However, when the same participants were making inferences to more indirect questions such as “What was the cause of the explosions?” people seemed to fall back on the only explanation they knew, namely the pressurised gas cylinders. Explicit knowledge of a correction is therefore no barrier to people using the out-dated information during their reasoning. In these studies, reliance on out-dated misinformation was eliminated only if people were given an alternative account of the fire (i.e., evidence for arson, such as petrol-soaked rags, were found at the scene). Access to this alternative explanation enabled people to (largely) discount the volatile materials as the cause of the fire.

However, providing an alternative account is not always a solution to the problem. Specifically, providing an alternative may be helpful in simple scenarios with a straightforward cause-and-effect chain (e.g., arson, not negligence, caused the
fire), but it may be potentially less useful in scenarios with many alternatives and complex interactions. In particular, if the alternative account is more complicated than the original misinformation-based account, the alternative may not be fully embraced. People generally prefer simple explanations to those which require more cognitive effort to understand and remember (Lombrozo, 2007). For example, Heit and Rotello (2012) reported that valid explanations are considered stronger when they are short, whereas, ironically, illogical explanations are considered more convincing when they are long. Those results mesh well with Chater’s (1999) proposal that finding the simplest explanatory pattern from complex data is a fundamental principle of cognition that guides perception, learning, memory, and reasoning. Chater argued that simplicity was not only a crucial factor in describing people’s behaviour across a range of cognitive domains, but that the application of a simplicity principle was also normatively justified (meaning essentially that all other variables being equal, simple explanations are more likely to be correct and therefore helpful in making predictions and decisions). On that basis, a simple mental model of an event is not only easy to understand, efficiently stored and remembered, but it is also likely to be adequate. Hence, plausible and simple explanations will be more readily accepted, remembered, and utilised than a correct but over-complicated account.

To summarise, misinformation effects are usually not an issue when there is a simple and plausible alternative to fill the gap created by a retraction in a person’s mental model of an event or links between concepts. When there is no simple alternative available, however, people often continue to rely on retracted misinformation. To better understand this reliance on out-dated or incorrect information, it is useful to consider the relevant memory processes in more detail.
We frame our discussion within a view that proposes two distinct sets of memory processes: One based on familiarity and one based on a more targeted, strategic memory search to recollect contextual details.

### 2.2.2 A Dual-Process Account of Misinformation Effects

Familiarity is considered a quick and automatic appraisal process. When people encounter a statement, they judge its familiarity automatically and independently of the recollection of its source and its validity assessment (Begg, Anas, & Farinacci, 1992; Hintzman, 2010; Jacoby, 1991; Zimmer & Ecker, 2010). For example, unfounded statements such as ‘the Great Wall of China can be seen from the moon,’ ‘sugar causes hyperactivity in children,’ or ‘humans use only 10% of their brain’ will be immediately familiar to many people even if they can also recollect having learned that these statements are, in fact, myths. Yet, familiarity can lead to errors of judgment because it is often difficult to determine why a piece of information is familiar. Hence, the simple fact that an item is familiar can lead to its acceptance as true (Schwarz, Sanna, Skurnik, & Yoon, 2007) or the impression that the item had been encountered just recently. An example of the latter can be seen in the ‘mirror effect’ in episodic recognition studies, where unstudied but highly familiar items (e.g., high-frequency words) typically attract more false alarms than less familiar items; cf. Reder, Angstadt, Cary, Erickson, & Ayers, 2002; Reder, Nhouyvanisvong, Schunn, Ayers, Angstadt, & Hiraki, 2000). In this case, people misattribute a sense of familiarity that arises from an item’s high natural-language frequency (i.e., its general long-term familiarity) to its presumed occurrence in an experiment’s study phase (i.e., a specific short-term familiarity or recency).

Likewise, a question about an event or the relationship among concepts will automatically activate related memory representations, some of which will be
relevant and valid, but some of which may supply irrelevant, incorrect, or subsequently invalidated information (Ayers & Reder, 1998). For example, in the well-studied post-event misinformation literature (Loftus & Hoffman, 1989; Loftus & Palmer, 1974), people may retrieve information merely suggested to them after witnessing an event because that suggestive misinformation is more recent and hence potentially more familiar at retrieval (Ayers & Reder, 1998; Lindsay & Johnson, 1989). It might be suggested to a crime witness, for instance, that the victim’s arm was hurt during a theft incidence, when in actual fact the witness saw the victim’s neck being hurt. Because of the suggestion, the witness may later remember seeing the victim’s arm being hurt rather than their neck (Okado & Stark, 2003).

The reliance on familiarity alone is therefore insufficient for accurate information processing. This implies the need for strategic memory processes: there needs to be a filter or monitoring process that can scrutinise the retrieved memory output, and that can direct a potentially more targeted memory search process with the aim of recollecting specific contextual details (M. Johnson, Hashtroudi, & Lindsay, 1993; Schacter, Norman, & Koutstaal, 1998; Zimmer & Ecker, 2010). This filter and directed search is the domain of strategic memory processes. For example, recollecting how your PSYC101 lecturer showed you the evidence that people (mostly) use more than 10% of their brain will prevent you from accepting the myth as true even though it is familiar. Likewise, a specific search for information regarding an item’s source (e.g., whether it was actually encountered within the context of an experiment’s study list, or whether it was actually encountered during a witnessed event) will—if successful—preclude its false acceptance as valid. In other words, if strategic memory processes are utilised and functional, misinformation effects based on familiarity will not occur. Alas, strategic memory processes often
fail; they require more executive control and mental effort than automatic memory processes (Herron & Rugg, 2003; Jacoby, 1991), and hence they are more readily compromised by factors such as divided attention (Jennings & Jacoby, 1993; Troyer & Craik, 2000), old age (Jennings & Jacoby, 1993; Spencer & Raz, 1995), alcohol intoxication (Bisby, Leitz, Morgan, & Curran, 2010), or simply a lack of motivation.

It follows that people will utilise misinformation when (a) an invalid memory representation is activated by a cue, and (b) strategic memory processes fail (M. Johnson et al., 1993; Okado & Stark, 2003; Roediger & McDermott, 1995; Schacter et al., 1998).

This basic idea of a failed strategic memory process contributing to misinformation effects has been applied in various forms. Sometimes, one correctly remembers an aspect of an event but misremembers important details, such as the source of the information. Incorrect recollection of the source may contribute to misinformation effects, for example, when one erroneously remembers receiving a piece of information from a reliable source (e.g., reading it in a scientific review paper) when in fact one received it from an unreliable source such as a blog, or, as discussed above, when one remembers witnessing something (e.g., an injury of a crime-victim’s arm) that was in fact merely suggested (M. Johnson et al., 1993; Ost, Granhag, Udell, & af Hjelmsäter, 2008; Ruffman, Rustin, Garnham, & Parkin, 2001; Zaragoza & Lane, 1994).

Sometimes misinformation effects may arise when people remember the misinformation but do not recollect the retraction. Considering the earlier emphasis on people being able to remember the retraction but not using it during inferencing, this may seem like a contradiction. However, because recollection is a controlled process, there is a difference between being able to recollect something (in principle)
and actually recollecting it (cf. Ecker, Zimmer, & Groh-Bordin, 2007). This is why people may use misinformation in their reasoning in response to indirect questions about the cause of an event, while still recollecting the correction when questioned about the event’s cause, or asked even more directly whether a presented report contained any corrections (as discussed above). The assumption that misinformation effects may occur when people remember the misinformation but fail to recollect the retraction is supported by research on negations. These studies assume that representations of information that is retracted or negated remain intact but the information’s impact is usually offset by concurrent retrieval of a retraction/negation “tag” affixed to the misinformation representation in memory (Ecker et al., 2011a; Gilbert, Krull, & Malone, 1990; Wilson & Park, 2008)—when the tag is lost or is not retrieved alongside ‘its’ misinformation, misinformation effects occur.

Some evidence for the importance of strategic monitoring processes for the continued-influence effects of misinformation comes from a study by Ecker et al. (2010). In that study, people read news reports containing corrections. Participants relied less on retracted misinformation in their reasoning when they were informed about the basic notion of misinformation effects, and were warned about the possible presence of misinformation, before reading. The authors suggested that warnings put participants into a more strategic processing mode, allowing for both more efficient “tagging” of misinformation at encoding and better strategic monitoring at retrieval (cf. also Echterhoff, Hirst, & Hussy, 2005).

As mentioned above, the efficiency of strategic memory and monitoring processes is known to depend on age, with both children and the elderly having less efficient strategic processes. It is therefore unsurprising that misinformation effects tend to be particularly pronounced in children (Ruffman et al., 2001) and in the
elderly (Ansburg & Heiss, 2012; Dehon, 2006; Skurnik, Yoon, Park, & Schwarz, 2005; cf. also Spencer & Raz, 1995).

By the same token, because strategic monitoring is an effortful process, misinformation effects are generally more likely to occur when people operate under high cognitive load or with few cognitive resources available (Ansburg & Heiss, 2012; Dehon, 2006; Gilbert et al., 1990; Wilson & Wolf, 2009). For example, when participants are sleep-deprived at retrieval, they are more likely to produce false memories in the Deese-Roediger-McDermott (DRM) paradigm (Diekelmann, Landolt, Lahl, Born, & Wagner, 2008; note: in the DRM paradigm, participants study the associates of a focal word [e.g., cloud, water, etc.] but not the focal word itself [e.g., rain]; yet participants reliably misremember studying the focal word). In Diekelmann et al.’s study, sleep-deprivation (i.e., fatigue, not wakefulness after study) led to higher levels of false recognition of the non-studied focal word.

In summary, people will tend to rely on automatically retrieved, familiar information when they do not have the cognitive resources to invest in monitoring their memory and reasoning. Resources might be compromised by many factors including age and the demands of the situation. Whenever people rely on automatic memory processes alone, the potential for inadvertent reliance on information that is known to be false is greatest.

2.2.3 Why Retractions Fail: The Effects of Repetition and Familiarity

Because of the impact of automatic memory processes, making (mis)information more familiar is one of the most powerful ways to make people believe and use it. It is for this reason that advertisements are put on high rotation, and this also explains why politicians repeat the same talking points ad infinitum (Singh, Kristensen, & Villaseñor, 2009). Moons, Mackie, and Garcia-Marques
(2009) reported that repeating an argument led to stronger agreement with it; if the argument was a weak one, however, repetition led to stronger agreement only in participants with low induced motivation for analytic processing—a finding that supports the link between low cognitive effort / strategic memory processing and familiarity-based misinformation effects.

It may seem rational to more strongly believe in statements one has heard repeatedly from various sources—in fact, a statement encountered in different contexts will be particularly memorable (e.g., Verkoeijen, Rikers, & Schmidt, 2004), and true information is more likely to be encountered in a large number of contexts than a fabrication. However, even repetition from a single source will make a statement more believable (Weaver, Garcia, Schwarz, & Miller, 2007), and internally repeating a statement by retrieving it from memory may be even more powerful in generating perceived validity than actual (external) repetition (Ozubko & Fugelsang, 2011). It is particularly remarkable that even when people accurately recognise that a statement is familiar simply because it was just encountered in a lab experiment, or when the source is known to be unreliable, the statement’s familiarity will still influence people’s belief in it (Begg et al., 1992; Henkel & Mattson, 2011).

Given how generally effective repetition seems to be as a memorial agent, it is tempting to assume that the same benefits of repetition and familiarity might also arise for retractions: Perhaps three corrections are better than one? Contrary to that expectation, the repetition of retractions appears to be less powerful than the repetition of initial information. In a study by Ecker et al. (2011a), repeating a piece of misinformation strongly increased people’s reliance on it, but repeating the retraction only had a limited effect. In their study, three retractions were more effective than a single retraction after misinformation was encoded three times, but
three retractions were as (in-)effective as one retraction when a piece of misinformation was encoded only once. The authors explained this in the framework of “retraction-tags” that was introduced earlier. Ecker et al. argued that each repetition of misinformation lays down a separate memory representation, and repetition thus leads to higher overall availability of misinformation in memory (because there are multiple traces). However, each representation can only be offset by one retraction, which “tags” the representation as being incorrect. Repeated retractions therefore have only limited efficacy because once a representation has been tagged, additional repetitions of the retraction have no effect on that particular representation.

Another explanation for why retractions, and repeated retractions in particular, are so ineffective can be derived from the dual-process perspective. We have already discussed how retractions are often ineffective: It turns out that, paradoxically, retractions may even increase an individual’s acceptance of incorrect information, a class of phenomena known as “backfire” or “boomerang” effects. One example of this is the “familiarity backfire effect.” This counterintuitive phenomenon occurs when a retraction repeats the misinformation, thereby inadvertently increasing its familiarity. For example, the statement “the belief that the MMR vaccine causes autism is false” repeats the association between MMR and autism and may make the statement “MMR causes autism” seem more familiar. When recollection of specific contextual information fails—memory for details such as the source of the information or its “tagging” as false can be expected to deteriorate over time—only the enhanced familiarity of the statement may remain (Skurnik et al., 2005; also see Schwarz et al., 2007). The marketing industry is well aware of this effect: Even negative advertisements and negative publicity can have a positive long-term effect...
on sales when it serves to make people more familiar with a product or brand (Berger, Sorensen, & Rasmussen, 2010).

The familiarity backfire effect has been demonstrated most clearly by Skurnik et al. (2005). These authors presented participants with health-related claims such as “Aspirin destroys tooth enamel.” Claims were either labeled as valid or invalid (because of ethical concerns with presenting misleading health claims, all claims were actually valid but were not obviously true or false to the non-expert participants), and were presented once or three times. Thus, an invalid item presented three times would have been identified as invalid three times, as opposed to just once for the once-presented counterpart. The results were intriguing. On the one hand, Skurnik et al. found that repeating claims that were labeled invalid led older adults to misremember them as valid less often after a short delay of 30 minutes (presented three times: 17 % vs. once: 28%), thus attesting to the success of repeated retractions (young adults had much lower rates, misremembering 7 % of invalid claims presented three times and 10 % of invalid claims presented once). On the other hand, after a delay of three days, the effect was reversed, with older people misremembering the thrice-repeated invalid items as valid more often (40 %) than their once-presented counterparts (28 %). (Young adults did not show this effect, with rates for invalid claims presented thrice vs. once at 14 and 24 %, respectively.) Skurnik et al. argued that although repetition served to highlight that a claim was invalid, thereby facilitating its immediate rejection, the repetition also increased the claim’s familiarity and thus led to higher levels of misinformation acceptance—at least in older participants—after a delay. After a delay, strategic recollection processes were particularly important, lest they be overwhelmed by relatively intact familiarity-based memory. Because strategic processes are generally weaker in the
elderly, they were insufficient to counteract the (misleading) familiarity signal arising from the repetition of the invalid information (cf. Prull, Dawes, Martin, Rosenberg, & Light, 2006; Spencer & Raz, 1995).

Wilson and Park (2008) subsequently adapted the statements used by Skurnik et al. (2005). However, rather than labeling statements as valid or invalid, Wilson and Park simply negated some of the statements (e.g., “Aspirin does not destroy tooth enamel”). Again, elderly people were found to be more likely to misremember negated statements as valid, this time after a mere 45 minutes.

Turning to the other end of the adult lifespan, research on misinformation in younger adults after shorter delays has hitherto yielded inconclusive evidence. Recall that Skurnik et al. (2005) found no evidence for a familiarity backfire effect in young adults after 30 minutes. Likewise, Ecker, Lewandowsky, and Hogan (2017) presented participants with fictional news reports similar to the “warehouse fire” scenario, in which an initially-presented cause of an event subsequently turned out to be false, and was replaced by a valid cause. Ecker et al. found that repeating the initial, outdated cause during the retraction made the retraction more, not less, effective (cf. also Johar & Roggeveen, 2007).

In contrast, Skurnik, Yoon, and Schwarz (2007; as discussed in Schwarz et al., 2007) reported evidence for a familiarity backfire effect even in young adults after only 30 minutes of retention. They presented their participants with a flyer regarding the common flu vaccine, which employed the popular “myths vs. facts” format. For example, it stated that “The side effects [of the vaccine] are worse than the flu.—FALSE. The worst side effect you’re likely to get […] is a sore arm […].” Skurnik et al. found that retracting (and thereby repeating) common myths led to a marked increase in accepting those myths as true. Interestingly, attitudes towards the
vaccine had also changed—participants who had read the flyer rated the vaccine less favourably than those who had not. Berinsky (2012) reported that corrections of the rumor that U.S. health care changes would promote euthanasia tended to be more effective when the correction did not repeat the rumor.

In summary, with young adults the balance of evidence does not warrant undue concerns about the familiarity backfire effect. By contrast, in old adults, the presence of familiarity backfire effects seems to be well established and must give rise to concern.

We have identified the reliance on automatic memory processes, in the absence of effortful strategic processing, as a key contributing factor to the persistence of misinformation. That is, in the absence of a particular motivation to believe one version of an event, or one set of concept-relations, over another, and when there is no reason to doubt the validity of the retraction, people will rely on misinformation to the extent that strategic memory processes fail to retrieve the retraction. This could be due to a genuine failure, but could also happen when the retraction-information is available in principle, but is not retrieved because the retrieval cue is too indirect or people simply do not invest the effort required. The latter variable is particularly relevant because in many circumstances, people will be motivated to believe some pieces of information but not others. We therefore next consider the heuristics involved in people’s judgment of the believability of incoming information, and the role of people’s existing beliefs, attitudes, and worldviews in shaping their beliefs and memory.

2.3 Misinformation Effects and Social Cognition

It is impossible to critically assess the veracity of each and every piece of information one encounters, which makes the use of heuristics that can yield quick
judgments of credibility particularly attractive. Thus people will often assess the validity of a claim by assessing how it fits in with what they (or their peers) already know or believe. This can be a very efficient strategy and is usually adaptive (e.g., anyone would be sceptical about claims of a newly discovered Pacific island on which there is no gravity). In support, Richter, Schroeder, and Wöhrmann (2009) demonstrated how relevant pre-existing knowledge is used by people to reject false assertions quickly and automatically, provided that the pre-existing knowledge is accessible and people are certain that the knowledge is accurate.

2.3.1 The Role of Attitudes for Misinformation Effects

Potential problems arise from use of this heuristic when the existing knowledge is false, beliefs are biased, or people hold very strong attitudes that run counter to the evidence. Ecker, Lewandowsky, Fenton, and Martin (2014) showed that pre-existing attitudes determine how often a piece of attitude-relevant (mis-)information is used in reasoning, but they also demonstrated that pre-existing attitudes do not necessarily compromise the effectiveness of retractions. In their study, people scoring high or low on a racial-prejudice scale read a news report involving either an Aboriginal robber or an Aboriginal hero; the involvement of the Aboriginal was retracted in some conditions. Ecker et al. found that references to retracted racial misinformation co-varied with people’s racial prejudice (e.g., people with high racial prejudice referred to the Aboriginal robber more often, but to the Aboriginal hero less often, than people lower in prejudice). Ecker et al. also found that people in both high- and low-prejudice groups reduced their reliance on the racial information after a retraction to equivalent degrees, meaning that there was no evidence that people ignored the retraction when it ran counter to their attitudes.
However, this outcome stands in striking contrast to other research which has found that people holding a strong belief may reject any information that runs counter to that belief. Thus, attempts to retract attitude-congruent misinformation can be futile with strongly partisan individuals. Indeed, it can even backfire and strengthen the misconception (“the worldview backfire effect;” see Lewandowsky et al., 2012, for a review). For example, Nyhan and Reifler (2010) reported that attempts to refute non-fictitious misconceptions (e.g., that weapons of mass destruction had been found in Iraq during the 2003 invasion) were successful when the retractions were in line with people’s worldview, but that worldview-incongruent retractions were ineffective or ironically strengthened the misconceptions. It is presently unclear under what circumstances such worldview backfire effects can be observed, but Ecker et al. (2012b) argued that it will depend on whether acknowledging the retraction will require a change of attitudes: In the case of an arbitrary incident such as a robbery, one can uphold one’s belief (e.g., that most Aboriginals are criminal) even when accepting that the particular incident (e.g., that a particular crime was not committed by an Aboriginal) did not support the belief. In contrast, a refutation will likely not be accepted when it necessitates a change in belief, and potentially entails a loss of face. For example, for a person who supported the 2003 invasion of Iraq, acknowledging that no weapons of mass destruction were found in Iraq may require a re-assessment of political decisions made at the time, and re-assessment of one’s own support.

2.3.2 Motivated Reasoning and Misinformation Effects

The rejection of attitude-incongruent information can be seen as a case of motivated reasoning—biased information processing that seeks to confirm existing beliefs rather than update one’s knowledge based on an unbiased assessment of the
evidence (for reviews of motivated cognition, see Kunda, 1990; Nickerson, 1998; Redlawsk, 2002). Motivated reasoning behaviour is particularly prevalent when the belief or attitude in question is held with confidence and conviction because it is central to a person’s value system and defines their identity (Brannon, Tagler, & Eagly, 2007; Hart, Albarracín, Eagly, Brechan, Lindberg, & Merrill, 2009). This means that efforts to correct common misperceptions may have to focus on people with moderate—rather than extremely strong—convictions. Even in people with moderate attitudes, however, motivated reasoning seems sufficiently common (cf. Prasad et al., 2009) to warrant a more detailed examination.

The confirmation bias behind motivated reasoning can lead people to selectively expose themselves only to (“cherry-picked”) evidence that supports their prior beliefs. An example of such behaviour is the tendency of consumers to choose newspapers that are known to be biased in their reporting, or internet users to frequent websites that they know to support their worldviews (T. Johnson, Bichard, & Zhang, 2009). Other strategies include denying the existence of contrasting evidence or—when presented with such evidence as in the case of misinformation retractions discussed here—deflecting or discrediting the evidence or its source. For example, when a psychologist corrects the common myth that there is a debate amongst climate scientists as to whether climate change is anthropogenic, explaining how there is in fact a strong consensus on this (cf. Anderegg, Prall, Harold, & Schneider, 2010; Doran & Zimmerman, 2009), a frequent strategy of people who oppose the findings from climate science is to deflect and discredit the source by questioning the psychologist’s expertise to speak about climatology. Supporting this anecdotal evidence, derogating the source of the evidence was one of the strategies identified by Jacks and Cameron (2003) in their study on motivated reasoning. The
authors presented counter-attitudinal messages (arguing against the death penalty) to participants in favor of the death penalty, and measured people’s responses. The strategy of derogating the source was used by their participants in particular when the source of the counter-attitudinal message was perceived as lacking expertise, and when participants’ pro death penalty attitudes were particularly strong.

The two most frequently used strategies in the study by Jacks and Cameron (2003), however, were counter-arguing and attitude bolstering. Counter-arguing refers to the covert or overt generation of arguments to counter the corrective information being presented, a common and principally rational strategy as long as the counterarguments maintain ties to empirical reality. Attitude bolstering means to completely ignore the refutation and focus on supporting evidence.

In a study by Prasad and colleagues (2009), participants were presented with convincing evidence that was incongruent with both a specific misconception they held and their general political worldview (specifically, Republican voters who believed in a link between Saddam Hussein and the terror attacks of 9/11 were presented with contrary evidence including a statement from President G. W. Bush himself). Faced with the evidence, only 2% of participants explicitly acknowledged the inadequacy of their misconception and updated their beliefs. A higher proportion (14%) denied holding the misconception in the first place, a surprising behaviour that may have served to avoid “losing face.” Most participants, however, displayed some form of motivated reasoning. The utilised strategies included the above-mentioned counter-arguing of facts and attitude bolstering. Prasad et al. also identified two other strategies. One was ‘inferred justification,’ which describes a form of flawed backward reasoning where people recursively infer the reasons that justify a present situation, ignoring evidence to the contrary. For example, some
people argued that there must have been a link between Hussein and 9/11 because there must have been a reason for the U.S. invasion of Iraq. The other strategy identified by Prasad et al. was ‘disputing rationality;’ this is when people insist on their right to an opinion without factual reasoning, often combined with assertions of confidence that nothing could change their mind.

Munro (2010) additionally described a situation in which people discounted evidence by denying the in-principle amenability of a topic to scientific investigation (‘scientific impotence discounting’). In that study, participants who believed that homosexuality was (or was not) a mental illness rejected evidence that homosexuals are not (or are) over-represented in psychological treatment facilities, and were more likely to agree with the suggestion that the topic could not be studied scientifically. In fact, the rejection of science even generalised to other topics, implying that when given evidence inconsistent with their beliefs, people became more sceptical of the scientific method in general.

Not surprisingly, motivated reasoning tendencies are less prevalent in open-minded people (Hart et al., 2009). Levitan and Visser (2008) asked their participants to rate the heterogeneity of attitudes in their social network by having them list five members of their social network, and rate how much they agreed with each of those people’s worldviews. The study demonstrated that people exposed to diverse attitudes through their social network are more open to change their attitudes in response to counter-attitudinal persuasive arguments, provided arguments are sufficiently strong (i.e., attitude diversity did not lead to attitude change in response to weak arguments).

Some researchers have argued that personality traits such as “social vigilantism”—the tendency to impose one’s own beliefs on others—predict greater
expressions of belief superiority and increased counter-arguing tendencies (Saucier & Webster, 2010). Other researchers have expressed the somewhat contrasting view that motivated reasoning tendencies are linked to low self-esteem (Wiersema, van Harreveld, & van der Pligt, 2012), negative mood (Jonas, Graupmann, & Frey, 2006), and threats of social exclusion (Greitemeyer, Fischer, & Kastenmüller, 2012).

The idea that people with low self-esteem (or people in a situation of reduced emotional or social stability) have a stronger need to defend their identity-defining attitudes is in line with self-affirmation theory. Self-affirmation theory claims that people use a variety of psychological adaptations, strategies, and distortions to protect the self from the threats of perceived failure or deficiencies (cf. Sherman & Cohen 2006). In the present case, this implies that people with low self-esteem and strong attitudes will find counter-attitudinal evidence particularly threatening to their identity and self-worth, and are hence more likely to engage in motivated reasoning. Motivated reasoning tendencies can therefore be softened by affirming people’s self-worth, for example by focusing them on personal achievements before confronting them with attitude-incongruent corrective information (Cohen, Bastardi, Sherman, Hsu, McGoey, & Ross, 2007).

The motivated-reasoning research implies that pre-existing attitudes can create a major obstacle for attempts to correct misinformation, in particular when corrections are perceived to require attitudinal change. The literature just reviewed suggests that counter-attitudinal rejections are more likely to be accepted when
people are reminded of attitudinal diversity in their social network, when they are in
a good mood and self-affirmed, and not threatened by social exclusion.²

2.4 Correcting Misinformation

We have already discussed that mere retractions (even when repeated) are
fairly ineffective in reducing reliance on invalid or outdated information. Further, the
above discussion of the familiarity backfire effect suggests that at least in older
people, repeating misinformation during attempts to correct it carries the inherent
risk of ironically strengthening the misconception (Skurnik et al., 2005, 2007). This
finding presents a pragmatic challenge, because often a correction will require a
myth to be repeated; for example, the statement ‘the Great Wall of China is not
visible from the moon’ may inadvertently strengthen the link between ‘the Great wall
of China’ and ‘visible from the moon’, yet it is unclear how the debunking could be
performed in this instance without mentioning the misinformation. Finally, we have
discussed how motivated reasoning can make it difficult to correct misconceptions
when the correction runs counter to strongly held beliefs. In some instances, pre-
existing beliefs can be so powerful that the presentation of corrective information can
ironically fortify the misconception (viz. the ‘worldview backfire effect’; Nyhan &
Reifler, 2010).

In opposition to those concerns involving backfire effects, there is a growing
literature on the deliberate use of misinformation, and its refutation, as an
educational tool. Students acquiring basic knowledge in any subject will typically

² While it goes beyond the scope of this chapter, there are also strategies that aim more at changing
behavioral tendencies rather than the attitudes and beliefs per se. These include the design of choice
architectures (e.g., opt-in vs. opt-out approaches to superannuation or organ donation schemes) and
the framing of information with respect to attitude-congruence (e.g., a person concerned about climate
change might be persuaded to install solar panels on their house by arguments referring to their carbon
footprint, but persons who are not concerned about climate change may find economical arguments
more convincing). See Lewandowsky et al. (2012) for a brief review.
harbor a variety of misconceptions, and techniques that can effectively reduce these are of interest to all teachers and science communicators. Directly addressing and refuting such misconceptions to introduce the valid information has been shown to be more effective than presenting the same valid information in a ‘standard’ teaching format (Kowalski & Taylor, 2009). Unlike plain retractions which simply state that some piece of information is not true, a refutational text is more detailed; although it explicitly presents the misinformation, it also provides a comprehensive explanation of why it is incorrect. For the ‘Great Wall of China’ example, a refutation might include supporting evidence such as the narrow width of the wall relative to the distance between the moon and the earth, and how that translates into a visual angle too small for the human eye to resolve. A refutational text might also try to explain why the misinformation was presented in the first place. For example, to understand why Andrew Wakefield suggested that autism was linked to the MMR vaccine, it helps to know that he received around half a million pounds in undisclosed payments from a lawyer preparing class action against the producers of the compound vaccine, and that there were plans to start a company to sell diagnostic tests (Deer, 2011). Arguably, such explanations will facilitate belief updating because they foster people’s scepticism regarding the initial misinformation and its source (cf. Lewandowsky, Stritzke, Oberauer, & Morales, 2005, 2009).

A meta-analysis by Guzzetti, Snyder, Glass, and Gamas (1993) explored the efficacy of different interventions intended to reduce misconceptions in the classroom. Participants ranged from primary school students to university postgraduates, and a wide range of science misconceptions were incorporated. The authors found seven types of intervention being utilised within these studies, with refutational text and non-refutation texts (such as simple narratives, or expository
texts where new information is presented with no reference to the misconception), being the most common. Refutation texts were found to be the most effective strategy of all interventions, and significantly more effective in producing conceptual change than non-refutational text. Several studies included a post-test one or two months later, and refutational text was the only intervention to foster long-term conceptual change.

Efficacy of a refutation is dependent on a number of factors. Guzzetti, Williams, Skeels, and Wu (1997) observed three high school physics classes for an 8-month period, accumulating qualitative data through interviews and questionnaires. They found that students ignored the refutation text and persisted with misinformation if (1) the refutation was poorly constructed and lacked clarity or (2) the students self-reported that they were only skimming the text. When students did not engage with the text, they mistakenly found evidence for their prior misconception within the refutation. It is also important that the misconception is based upon a lack of knowledge (which may be corrected by a valid explanation), rather than an individual’s attitudes or belief system, as previously discussed.

The effectiveness of refutation texts is often explained within the ‘conceptual change model’ proposed by Posner, Stike, Hewson, and Gertzog (1982). This model suggests that there are four stages necessary for conceptual change to take place. The first step is dissatisfaction with one’s own current understanding, which instigates cognitive conflict. The proposed replacement construct then needs to be intelligible (i.e., easily understood), plausible, and potentially fruitful (i.e., the student should be able to see the relevance of a correct understanding of matters and the potential applicability of the corrected knowledge to future problems).
Yet, considering the first step in Posner et al.’s (1982) model, to become dissatisfied with one’s insufficient understanding, one would first need to notice an incongruity between one’s current conception and the evidence. In fact, before one can notice such a discrepancy, one’s misconception, the presented evidence (and the associated correction) would have to be co-activated and aligned in working memory (cf. McCrudden & Kendeou, 2014; van den Broek & Kendeou, 2008).

McCrudden (2012) looked at students’ understanding of the biological term ‘fitness,’ which refers specifically to the number of an animal’s offspring that survive into adulthood, a meaning quite different from the common interpretation in terms of physical fitness. Based on an analysis of think-aloud protocols and interviews, McCrudden argued that when reading a refutational text, co-activation of the misconception and the corrective evidence may occur automatically, but that the detection of a discrepancy between the two may require strategic monitoring processes (cf. Ecker et al., 2010). Using think-aloud protocols, McCrudden further demonstrated how merely noticing the discrepancy between misconception and evidence is not sufficient for a change in belief—the reader needs to strategically engage in resolving the discrepancy. Kendeou and van den Broek (2007) found that students with prior misconceptions engaged in such ‘conceptual change strategies’ only when reading refutation texts. These conceptual change or resolution strategies were reflected through participants vocalising the juxtaposition between their misconception and the correct information, making text-based inferences, and paraphrasing (McCrudden, 2012; Kendeou & van den Broek, 2007).

The above discussion implies that one of the main reasons why refutational texts may be particularly effective in reducing misconceptions is that they set the stage for strategic ‘conceptual change’ processing. That is, refutational texts allow
people to co-activate, align, and integrate their misconceptions with corrective evidence, which then facilitates the updating of their beliefs (cf. Ecker et al., 2012a; Johar & Roggeveen, 2007).

This view of corrections via refutation is consistent (although it does not require) the notion of memory reconsolidation, which has gained recent popularity (see Hardt, Einarsson, & Nader, 2010, for a review, but also Ecker & Lewandowsky, 2012, for a critique). Reconsolidation theory proposes that when an item in memory is retrieved, it re-enters a transient, labile state, which may allow it to be modified, updated, or weakened, before is then re-stabilised (i.e., re-consolidated). When a refutation is read, the original misconception is necessarily recalled, and according to reconsolidation, this would render the misconception’s representation labile and amenable to be updated and to accommodate the correct information.

Apart from the ‘aligning’ of misconception and correction, the effectiveness of using refutational materials in a classroom setting may have another ‘operative ingredient:’ It fosters critical thinking. Refutational texts encourage students to critically appraise what they know, to sceptically assess empirical evidence, and to endeavour to draw valid conclusions from the evidence (cf. Baker & Piburn, 1991; Berland & Reiser, 2008; Kuhn & Crowell, 2011; Manson, Gava, & Boldrin, 2008; ten Dam & Volman, 2004). Bedford (2010) suggested that the in-depth study of misinforming materials itself could improve students’ understanding of scientific concepts, by motivating students to acquire the necessary knowledge and critical thinking skills to convincingly argue against misrepresentations. The explicit study of misinforming materials and the analysis of controversial issues can thus ameliorate students’ critical thinking and reasoning skills. That is, reading refutational texts and discussing them in-class, building arguments based on
evidence, will help students improve both their understanding of the subject matter and their argumentative reasoning skills. Alas, knowledge is currently often imparted as a set of unequivocal facts, and there is a lack of argument and debate in science classrooms (Osborne, 2010).

To conclude, our analysis of misinformation correction strategies suggests that familiarity backfire effects may only be a concern when there is little explanation regarding the motivation behind the initial misinformation and the evidence (i.e., in simple “this is the myth, this is the fact” situations). In situations in which more extensive explanations are possible, taking up and addressing the misinformation seems beneficial. It allows misconceptions and corrective evidence to be aligned, and seems to facilitate belief updating, in particular when a ‘factual wrapper’ can be established to encapsulate misinformation. This can be achieved by providing clear-cut evidence, and explaining the motivation behind the initial spreading of misinformation (i.e., a “this is the myth, this is why the myth was spread, this is the truth, this is the evidence” approach). Fostering critical thinking, scepticism, epistemological knowledge and an understanding of science from a young age thus seem crucial educational aims for contemporary information societies.
2.5 References


(retracted paper)


Chapter 3: The Role of Familiarity in Correcting Inaccurate Information
3.1 Foreword

One particularly counterintuitive phenomenon discussed in Chapter 2 is the backfire effect, where an individual’s belief *increases* after a correction has been elicited. One example described was the *familiarity backfire effect*. This is presumed to occur when the misinformation is repeated within the retraction, therefore increasing the misinformation’s familiarity and inadvertently the acceptance of the misconception being true (Cook & Lewandowsky, 2011; Lewandowsky, Ecker, Seifert, Schwarz, & J. Cook, 2012; Skurnik, Yoon, Park, & Schwarz, 2005). The familiarity backfire effect warrants further exploration as (1) reports of when the effect occurs have been inconsistent (which may relate to inconsistent definitions of the phenomenon) and (2) the existing literature does not conclusively demonstrate that a backfire effect caused solely by familiarity really exists. Chapter 3 thus seeks to consolidate theoretical inconsistencies specifying what a familiarity backfire effect should entail, improve upon previous experimental designs, and empirically examine the existence of the familiar backfire effect.

Relating to point (1), the term “familiarity backfire effect” has been used inconsistently within the literature, often taking on different meanings. The term is sometimes used to refer to occasions where myths are simply misremembered as facts, without a control condition or baseline comparison (cf. Peter & Koch, 2016). By contrast, we argue it should only pertain to cases where a correction inadvertently *increases* myth belief relative to a no-correction or pre-correction baseline. However, irrespective of the specific definition applied, reports of when and under what circumstances the familiarity backfire effect occurs have been inconsistent. Some studies found that the effect occurs after a retention interval of just half an hour (Skurnik, Yoon, and Schwarz 2007, as cited in Schwarz, Sanna, Skurnik, & Yoon,
2007), while others reported that it only occurs after a delay of three days and only in older participants (Skurnik et al., 2005). Other studies yielded results that did not support the notion of a familiarity backfire effect (Cameron et al., 2013; Ecker et al., 2010), although it must be noted that these studies were not primarily designed to assess the existence of the effect.

Regarding point (2), the two frequently cited papers providing the strongest evidence for a familiarity backfire effect do not conclusively demonstrate that the observed backfire effects were due to familiarity alone. One paper, Skurnik et al. (2007), is an unpublished manuscript. This is problematic as the exact details of the experiment are not available, such as whether certain conditions were compared in statistical analyses. Moreover, this study investigated processing of misinformation regarding vaccines. The topic of vaccines is a contentious issue, and as it is known that attempts to correct misinformation about contentious topics can backfire because people’s worldview is being challenged (Nyhan & Reifler, 2010; Nyhan & Reifler, 2015). It is thus unclear whether the backfire effect described by Skurnik et al. arose solely due to familiarity. In order to better measure the effects of familiarity, the experiments in Chapter 3 do not exclusively focus upon politicised information, but provides 20 myths and 20 facts covering many different topics.

Similarly, it is also difficult to draw firm conclusions from the Skurnik et al. (2005) study for several reasons. For instance, this study did not have a control condition where pre-correction belief ratings were measured. Therefore, we do not know whether or not participants’ beliefs increased after the correction was presented, as our conception of the familiarity backfire notion implies. Additionally, the generalisability of the findings is limited, as the study used older participants but did not use a cognitive screening task to exclude participants with cognitive impairments. Finally, all the claims presented to participants were arbitrarily labelled
as ‘myths’ and ‘facts’, but were in fact all true. As an individual’s belief in the correction impacts its efficacy (Rich, Van Loon, Dunlosky, Zaragoza, 2017), some of the effect may have been driven by participants not trusting the experimenters to provide valid information. In contrast, the experiments described in Chapter 3 measured participants’ belief both prior to and after corrections were presented; both young and older adults were tested, with older adults being screened for cognitive decline using the Montreal Cognitive Assessment; and materials were myths and facts that were indeed false and true, respectively.

Moreover, Chapter 3 takes a novel, theory-driven approach to measuring familiarity’s role in the processing of misinformation. We use dual-processing accounts of memory (Yonelinas, 2002), to help us understand misinformation effects, theorising that familiarity-based acceptance of myths should be most likely to occur in the absence of strategic memory processes. As we know under what circumstances strategic memory processes are more likely to fail—i.e., in an older adult population, when there is not sufficient recollectable detail in an explanation, and after a long delay between encoding of a correction and its retrieval—we were able to manipulate the extent to which people rely upon familiarity. This enabled us to determine if and under what circumstances corrections backfire due to familiarity.
3.2 Abstract

People frequently continue to use inaccurate information in their reasoning even after a credible retraction has been presented. This phenomenon is often referred to as the continued influence effect of misinformation. The repetition of the original misconception within a retraction could contribute to this phenomenon, as it could inadvertently make the “myth” more familiar—and familiar information is more likely to be accepted as true. From a dual-process perspective, familiarity-based acceptance of myths is most likely to occur in the absence of strategic memory processes. We thus examined factors known to affect whether strategic memory processes can be utilised; age, detail and time. Participants rated their belief in various statements of unclear veracity, and facts were subsequently affirmed and myths were retracted. Participants then re-rated their belief either immediately or after a delay. We compared groups of young and older participants, and we manipulated the amount of detail presented in the affirmative/corrective explanations, as well as the retention interval between encoding and a retrieval attempt. We found that (a) older adults over the age of 65 were worse at sustaining their postcorrection belief that myths were inaccurate, (b) a greater level of explanatory detail promoted more sustained belief change, and (c) fact affirmations promoted more sustained belief change in comparison to myth retractions over the course of one week (but not over three weeks). This supports the notion that familiarity is indeed a driver of continued influence effects.
3.3 Introduction

Every day we process an extraordinary amount of information, and it is often up to the individual to discern fact from fiction. A proportion of this information is inevitably inaccurate and deserves to be corrected after initial encoding. In order to maintain an accurate and up-to-date representation of the world, ideally we would disregard invalidated information. However, we are far from perfect at performing this task, as corrected misinformation often continues to influence memory and reasoning—this persistence is termed the continued influence effect of misinformation (cf. Ecker, Lewandowsky, Swire, & Chang, 2011; Johnson & Seifert, 1994). With “fake news” fast becoming a global issue, and with the increased spread of misinformation over social media, the ability to effectively correct misinformation has never been more important (Connolly et al., 2016; Lavoipierre, 2017). From relatively benign misconceptions such as “ostriches hide their heads in the sand” to more malignant misinformation such as “the MMR vaccine causes autism” (Poland & Spier, 2010), studies have consistently observed a continued influence effect. In other words, simply stating that information is incorrect is often ineffective, with simple retractions typically only halving the number of references to a critical piece of misinformation relative to a no-retraction control condition (see Lewandowsky et al., 2012, for a review). Part of the reason why corrections are often ineffective may arise because corrections typically repeat the misinformation, thereby making it more familiar. The present paper investigates whether the continued influence effect is at least partially familiarity-driven, and how beliefs change over time after a credible correction has been presented.
3.3.1 The Illusory Truth Effect

The illusory truth effect occurs when increased familiarity gives rise to the illusion that information is valid and inadvertently increases an individual’s belief (Begg, Anas, & Farinacci, 1992; Dechene, Stahl, Hansen, & Wanke, 2010; Wang, Brashier, Wing, Marsh, & Cabeza, 2016). For example, DiFonzo, Beckstead, Stupak, and Walders (2016) presented participants with rumors about campus life, such as a professor giving a student good grades to keep them quiet about the professor’s plagiarism. DiFonzo and colleagues varied the number of presentations (from 0-9 times) and found that belief in the statements was logarithmically related to the number of repetitions, such that belief increased with each repetition (albeit in diminishing amounts). In line with this, a recent popular news survey found that 75% of people assumed fake headlines to be true if they were familiar (Silverman, 2016).

The illusory truth effect could be problematic when attempting to correct misinformation, as a correction often repeats the original claim. For example, truthfully stating that playing Mozart to your child will not boost its IQ mentions the two concepts of “Mozart” and “increased IQ”, thereby making the link between the concepts more familiar even though the statement seeks to dispel the Mozart-IQ myth. This inadvertent increase in familiarity may reduce the effectiveness of a correction and may thus contribute to the continued influence effect of misinformation.

3.3.2 Strategic and Automatic Memory Processes

The potential familiarity-related difficulties that arise during the correction of misinformation may be explained from a dual-processing perspective. Dual-process theories of memory assume a dichotomy between automatic memory processes, which include familiarity, and strategic memory processes such as recollection and
output monitoring (cf. Brown & Warburton, 2006; Diana, Yonelinas, & Ranganath, 2007; Rugg & Curran, 2007; Yonelinas, 2002; Yonelinas & Jacoby, 2012; Zimmer & Ecker, 2010). Familiarity is thought to be a fast, context-free automatic process that allows for the rapid recognition of previously encountered information. Recollection, by contrast, is a slower process thought to allow for the retrieval of contextual details, such as the information’s source, its spatiotemporal encoding context, or its veracity. In the case of corrected misinformation, it is often assumed that a “negation tag” is linked to the original statement, for example, “playing Mozart to your child will boost its IQ—NOT TRUE” (Gilbert, Krull, & Malone, 1990). Thus, a corrected statement may require strategic memory processes for veracity to be successfully retrieved, as the negation tag is at risk of being lost if only automatic processes are employed, which may however identify the statement (e.g., the Mozart-IQ link) as familiar.

Regardless of whether statements are correct or have been invalidated, existing memory representations will be activated in response to cues via automatic retrieval to the extent that the information is familiar (cf. Ayers & Reder, 1998). To avoid reliance on familiar but invalid information, strategic memory processes are required to act as a filter of automatically retrieved memory output. However, strategic memory processes take effort and often fail (e.g., Herron & Rugg, 2003), and thus people can rely upon invalid but automatically retrieved information in their judgments (Ecker et al., 2011; Koutstaal & Schacter, 1997; Reyna & Lloyd, 1997; Roediger, Watson, McDermott, & Gallo, 2001). A post-correction misinformation effect is therefore likely to occur when misinformation has been automatically activated but strategic memory processes have failed (Ecker, Lewandowsky, & Tang, 2010). Familiarity can thus hinder the remediating effect of a correction when the repetition of misinformation in the course of its correction boosts an invalid item’s
familiarity such that it outweighs the correction’s strategic-retrieval dependent corrective effect.

### 3.3.3 The Familiarity Backfire Effect

Some reports suggest that the familiarity boost associated with a correction can be so detrimental that it causes a *familiarity backfire effect*, such that a correction can ironically increase an individual’s belief in the very misconception the correction is aiming to rectify (J. Cook & Lewandowsky, 2011; Lewandowsky et al., 2012).

An unpublished manuscript by Skurnik, et al. (2007) is frequently cited when discussing the familiarity backfire effect (e.g., Berinsky, 2015; J. Cook, Bedford, & Mandia, 2014; Gemberling & Cramer, 2014; Lilienfeld, Marshall, Todd, & Shane, 2015; Peter & Koch, 2016; Schultz, 2012). Skurnik et al. presented participants with a flyer presenting a number of flu-vaccine related claims. The flyer was either presented in a “myths vs. facts” format, which affirmed the factual statements and refuted the incorrect, or the flyer only affirmed the facts but did not mention the myths (or, in a control condition, there was no flyer at all). Immediately after participants read the “myth vs. facts” flyer, they were able to reliably distinguish between myths and facts, yet after a delay of 30 minutes, participants misidentified 15% of the myths as facts (compared to only 2% of the facts being classified as false). Also, after a 30-minute delay, participants in the myths vs. facts condition had less favorable attitudes towards the vaccine than participants who had never seen the flyer at all.

---

3 The term familiarity backfire effect has been used somewhat inconsistently. The term is sometimes used simply when myths are misremembered as facts, without a control condition or baseline comparison (cf. Peter & Koch, 2016). However, we argue it should only pertain to cases where a correction inadvertently *increases* myth belief relative to a no-correction or pre-correction baseline. Our definition is in accordance with other backfire effects such as the *worldview backfire effect* (Lewandowsky et al., 2012).
Skurnik et al.’s (2007) finding that people had a less favorable attitude towards vaccines than those who did not view the flyer may reflect a familiarity backfire effect. However, given that Skurnik et al. only focused upon one contentious issue, there is an alternative explanation—namely a worldview backfire effect. This backfire effect ensues when a correction challenges a person’s belief system and the person becomes motivated to defend their worldview, resulting in an increased belief in the inaccurate information, relative to a situation where the correction was never presented (Lewandowsky et al., 2012; Trevors, Muis, Pekrun, Sinatra, & Winne, 2016). This effect is a known risk when debating contentious issues and can contribute to belief polarisation (Hart & Nisbet, 2012; Nyhan & Reifler, 2010).

Nyhan and colleagues (Nyhan, Reifler, Richey, & Freed, 2014; Nyhan & Reifler, 2015) recently demonstrated that corrections of vaccine-related misconceptions can backfire in people concerned about vaccination safety and/or opposed to vaccinations. Thus, when attempting to measure the effects of familiarity, it would be beneficial to not exclusively focus upon politicised information (e.g., Berinsky, 2015) or contentious topics such as vaccination to avoid confounding the effects of familiarity and worldview.

Regarding the misidentification of myths as facts in the Skurnik et al. (2007) study, results from the comparison between the “myths vs. facts” condition and the no-flyer control condition are not available, so it is unclear whether myth belief was greater after presenting corrections than after not presenting any information at all. However, misremembering myths as facts could in general reflect an interesting hurdle in belief updating, even if there is no “true” backfire effect. Considering the benefits of clear communication, in particular in the context of myth debunking, it is important to better understand the factors that may contribute to this differential
forgetting of myths relative to facts (or more precisely, differential forgetting regarding the veracity of myths relative to facts).

Theoretically, if people are presented with explanations affirming facts or refuting myths, belief in facts may be sustained over time, whereas myth items appear to be “forgotten”, simply because automatic and strategic memory processes operate in concert for facts but stand in opposition for myths (cf. Brainerd & Reyna, 2008; Jacoby, 1991; Toth, 1996). For fact items, regardless of whether automatic processes or strategic memory processes are employed, both would lead a participant to conclude that the item is true. By contrast, if a participant is unable to correctly recall the correction of a myth due to the forgetting that primarily affects strategic memory processes, the familiarity of the myth—boosted by its repetition during the correction—could lead to the myth being inaccurately accepted as true.

3.3.4 Factors Likely to Influence the Correction of Information: Detail, Time, and Age

It follows from the dual-process notion that the relative impact of familiarity on corrections could potentially be influenced by factors that are known to affect strategic memory processes, including (1) the amount of detail presented in the corrective explanation, (2) the retention interval between encoding and a retrieval attempt, and (3) the age of the participant.

Regarding the correction’s level of detail, providing sufficiently detailed explanations as to why a piece of misinformation is false—in other words, providing a detailed refutation rather than a sparse “that-is-not-true” retraction (cf. Guzzetti, 2000)—might counteract familiarity’s influence. For example, where a simple retraction would merely state that “listening to Mozart will not boost your child’s IQ”, a detailed refutation would also explain why (e.g., by highlighting that scientific findings were misrepresented in a popular yet unscientific book, and that the original
study neither tested infants nor measured IQ; Campbell, 1997; Pietschnig, Voracek, & Formann, 2010; Rauscher, Shaw, & Ky, 1993). Thus, refutations directly address the misconception and explain the reasons why the misinformation is false and/or where the misconception originated. Refutations have been found to promote belief change over long periods of time (Diakidoy, Mouskounti, Fella, & Ioannides, 2016; Guzzetti, Snyder, Glass, & Gamas, 1993); it is assumed that refutations are more effective than retractions because they encourage the detection of inconsistencies between a person’s inaccurate beliefs and the corrective information, and because they provide richer informational detail that can later support recollection of the correction (Guzzetti, 2000).

Regarding the retention interval, failure of strategic processes is particularly likely when there is some delay between encoding and attempted retrieval, as strategic recollection of details diminishes with time, whereas familiarity stays relatively constant (Knowlton & Squire, 1995). Thus, false acceptance of myths based on their familiarity seems particular likely at longer retention intervals.

Regarding age, older adults have less efficient strategic memory processes than young adults, whereas automatic processing such as familiarity-detection remains relatively age-invariant (e.g., Prull, Dawes, Martin, Rosenberg, & Light, 2006). In particular, older adults seem to become less efficient at binding item and context information (Naveh-Benjamin, 2000); therefore, the mnemonic link between a statement and its veracity could be weaker in older adults. This is in line with the finding that source memory—memory for where or how information was acquired—is particularly susceptible to the effects of ageing (e.g., Glisky, Rubin, & Davidson, 2001). Consistent with this notion, Skurnik et al. (2005) found that older adults were particularly likely to misremember myths as facts after repeated retractions (compared to single retractions) after a three-day retention interval (but not after 30-
minutes, and not in younger adults as per the Skurnik et al. 2007 study). However, it is difficult to draw firm conclusions from the Skurnik et al. (2005) study for several reasons: (i) there was no control group where corrections were not presented at all or pre-manipulation belief ratings were measured; (ii) no cognitive screening task was given to participants, potentially reducing the generalisability of findings; and (iii) health claims were used that were arbitrarily labeled as valid or invalid without explanation, even though all claims were actually true—thus some corrections were misleading, and distrust in the corrections may have contributed to the results, as it is well established that source credibility is an influential factor in the persuasiveness of messages (Eagly & Chaiken, 1993; Guillory & Geraci, 2013).

In summary, factors such as the correction’s level of explanatory detail, retention interval, and participant age are likely to play a role in determining the success of a correction, but their specific importance is unclear and findings have been inconsistent. By manipulating and comparing these factors, the present research aimed to clarify if and under what conditions familiarity is most problematic. Experiment 1 tested young adults, Experiment 2 tested older adults.

3.4 Experiment 1

This study presented an undergraduate population with both incorrect and correct claims (i.e., myths and facts), then corrected the false claims in a way that boosted their familiarity. To this end, participants were presented with a range of statements of unclear veracity that were subsequently labeled as true or false. People’s belief in the statements was measured both before the true/false explanation and in a post-manipulation test phase to yield a measure of belief change. To avoid the problems associated with posttest-only designs (Morris, 2008), we used a pretest-
posttest design so that each individual could be used as their own control (Hunter & Schmidt, 2004).

Level of explanatory detail and study-test retention interval were manipulated to identify the parameters of corrections that promote successful discounting of misinformation. The experiment used a $2 \times 2 \times 3$ within-between design, with within-subjects factors type of item (myth vs. fact) and type of explanation (the veracity of each statement was explained either briefly or in some detail), and the between-subjects factor retention interval (immediate, 30-minute, or one-week). In some studies, continued influence effects were found primarily in more indirect measures of belief that require participants to use the misinformation in reasoning (cf. Johnson & Seifert, 1994). Therefore, inference questions were also administered at test, serving as a more indirect measure of belief that could help avoid issues related to social desirability.

We hypothesised that (1) detailed explanations would lead to stronger belief change than brief explanations for both myths and facts, and (2) belief change would be more sustained over time after fact affirmation compared to myth retraction, as false familiarity-based acceptance of myths would seem particular likely at longer retention intervals. We did not expect a backfire effect, as there are no clear demonstrations of a true familiarity backfire effect in the peer-reviewed literature. However, we hypothesised that one was theoretically most likely to occur with a brief retraction after a one-week delay.

3.4.1 Method

Participants. A power analysis (conducted with GPower3; Faul, Erdfelder, Lang, & Buchner, 2007) suggested that 78 participants were required in order to detect a small-to-medium effect (effect size $f = .15$) with $\alpha = .05$, $1 - \beta = .80$, and a moderate correlation between repeated measures of $r = .50$. Participants were 100
undergraduate students from the University of Western Australia who volunteered after reading an ethically-approved information sheet. They received course credit for participation. Two participants did not complete the study, and five participants over the age of 30 were excluded as age outliers. This conformed to the age split of prior research (Skurnik et al., 2005) as well as the outlier labeling rule threshold (i.e., 2.2 times the interquartile range above the third quartile [Hoaglin & Iglewicz, 1987]), which was 29.8 years of age. The final pool included $N = 93$ participants, with 19 males and 74 females between 16-28 years of age, and a mean age of $M = 19.11$ ($SD = 2.53$).

**Stimuli.** There were 20 myths and 20 facts, each with a corresponding brief explanation, a detailed explanation, and two inference questions. An example myth / fact and the corresponding explanations and example inference questions are given in Table 1 (see Appendix A for the complete list of items, explanations, and inference questions). Brief explanations simply stated whether the item was a myth or a fact with no further clarification. They explicitly repeated the initial statement twice (once in the original and once in a negated format if the item was a myth). Thus participants encountered the initial statement three times altogether: once when being initially rated, and twice in the explanation.

Detailed explanations also provided the myth/fact label but in addition included three or four sentences of further information; myth retractions did not provide a causal alternative to the myth but rather explained why the myth was wrong and/or where it originated from. Detailed explanations explicitly repeated the initial statement only once, but elements of the statement were repeated in the additional information.
### Table 1

**Example of a myth and fact, corresponding explanations, and inference questions**

<table>
<thead>
<tr>
<th>Myth</th>
<th>Liars sometimes give themselves away by physical ‘tells’ such as looking to the right or not looking you in the eye.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brief retraction</td>
<td>Liars sometimes give themselves away by physical ‘tells’ such as looking to the right or not looking you in the eye.</td>
</tr>
<tr>
<td><strong>MYTH</strong></td>
<td>Liars do not give themselves away by physical ‘tells’ such as looking to the right or not looking you in the eye.</td>
</tr>
<tr>
<td>Detailed refutation</td>
<td>Liars sometimes give themselves away by physical ‘tells’ such as looking to the right or not looking you in the eye.</td>
</tr>
<tr>
<td><strong>MYTH</strong></td>
<td>Physical signals which are often assumed to be the ‘tells’ of a liar, are in fact signs of emotional discomfort in general. When a person is being interviewed or is accused of a crime, a non-liar is equally likely to express these signals. A meta-analysis of over 100 studies found no consistent physical cues when a person was lying. The experimenters stated that “there are no behaviours that always occur when people are lying and never occur when they are telling the truth”.</td>
</tr>
<tr>
<td>Myth inference question</td>
<td>What percentage of lies can FBI detectives catch just by looking at physical tells? (0-100%)</td>
</tr>
<tr>
<td>Fact</td>
<td>Humans can regrow the tips of fingers and toes after they have been amputated.</td>
</tr>
<tr>
<td>Brief affirmation</td>
<td>Humans can regrow the tips of fingers and toes after they have been amputated.</td>
</tr>
<tr>
<td><strong>FACT</strong></td>
<td>Humans can regrow the tips of fingers and toes after they have been amputated.</td>
</tr>
<tr>
<td>Detailed affirmation</td>
<td>Humans can regrow the tips of fingers and toes after they have been amputated.</td>
</tr>
<tr>
<td><strong>FACT</strong></td>
<td>Astonishingly, humans have a very amphibian-like trait of being able to regenerate. Unfortunately, this is limited to the very tips of our fingers and toes. A study in 1970 found that if the individual was under the age of 10, they had a limited capability to even regrow bone. There are stem cells at the base of each nail, which aid ordinary nail growth as well as the ability to rebuild the digit tip after amputation. Interestingly, a regenerated finger will sometimes lack a fingerprint.</td>
</tr>
<tr>
<td>Fact inference question</td>
<td>What proportion of fingers will regenerate after the tip has been amputated? (0-100 %)</td>
</tr>
</tbody>
</table>

Inference questions were rated on an 11-point scale, with the specific scale-value range dependent on the item; for example, some items were rated on a 0-10 scale, others were rated on a 0-20 % scale with 2 % increments.
Two pilot studies were conducted to select stimuli from a list of 80 items (55 myths and 25 facts) that was initially compiled by selecting various items from websites such as New Scientist, Scientific American, and myth busting programs such as QI. Each item was researched to the best of our ability, and where possible evidence from the peer-reviewed literature was sought out. The aim of the first pilot study was to select a pool of items that were common and mid-range believable, to allow for either reduction or increase in belief following retractions or affirmations, respectively.

The second pilot study was run to ensure that the inference questions were in fact indirect measures of belief (i.e., that they correlated with the associated explicit belief measures; e.g., to ensure the inference question ‘What percentage of lies can FBI detectives catch just by looking at physical tells’ is sufficiently measuring an individual’s belief that it is possible for liars to give themselves away by physical tells).

**Pilot Study 1.** The aim of the first pilot study was to select an item pool of myths that were common and at least mid-range believable. Thirty-one undergraduate students from the University of Western Australia took part. Participants indicated for 55 myth and 20 fact items (1) if they had heard of the item before (i.e., familiarity) and (2) the extent to which they believed each item (i.e., believability). Familiarity was measured on a five point scale ranging from “Definitely not heard before” to “Definitely heard before”. Myths were removed from the stimulus set if they scored below a mean of 3.5 out of 5. Believability was measured on a 1-10 scale ranging from “Not at all believable” to “Very much so”. Myths were removed from the stimulus set if they scored below a mean of 4.5; one additional item with a mean greater than 9.0 was also removed (to avoid any ceiling effects reducing the likelihood of a familiarity backfire effect).
After Pilot Study 1, there were 37 myths remaining. The mean familiarity score of the myths was $M = 4.46$ out of 5 ($SD = .35$). The mean familiarity rating of the facts was $M = 3.39$ ($SD = 1.04$). The believability mean for the myths was $M = 6.24$ ($SD = 1.12$) and the mean for the facts was $M = 5.34$ ($SD = 1.89$). Pre-manipulation familiarity and belief ratings were positively correlated, $r = .79$, showing that the more familiar items were believed more strongly.

**Pilot Study 2.** The second pilot study was run to ensure that the inference questions were in fact an indirect measure of belief in the initial claims. Participants were 100 individuals who volunteered via Crowdflower (http://www.crowdflower.com), a crowdsourcing website where contributors perform tasks and are paid for their services. Participants were paid $1.80. Five fact items were added to the set after Pilot Study 1 in order to boost their number in comparison to myths. Participants rated how much they believed in the 38 myths and 25 facts, and responded to two corresponding inference questions per claim.

Participants were excluded if they reported their English skills to be only “fair” (0 on a 4-point scale ranging from “fair” to “native speaker”; 5 individuals), if they took less than 15 minutes to complete the task (23 individuals) or more than 85 minutes (3 individuals; mean completion time was $M = 34.88$ minutes, $SD = 35.36$ minutes). The data were also screened for inconsistent response patterns suggestive of participants not paying attention, but no participants were excluded. A total of $N = 75$ participants were included in the analysis. Spearman’s correlations were calculated for each item and the two corresponding inference questions. Items were excluded from the stimulus set if both inference questions did not significantly correlate with belief in the corresponding claim (with $p < .05$; $r$ ranging from $.23$ – $.81$); this resulted in exclusion of 19 items—14 myths and 5 facts—leaving 24 myths and 20 facts.
In a final step, the four remaining myths with the lowest belief ratings and correlations between inference questions and belief ratings were removed. The final stimulus set thus comprised 20 myths and 20 facts, each with two corresponding inference questions.

**Procedure.** Participants were seated individually in testing booths and the experiment was administered by Qualtrics survey software. Participants were presented the 40 items in randomised order, and they indicated on a 0-10 scale the extent to which they believed each item using a computer mouse. Directly after each item was rated, participants received either a brief or a detailed explanation, which were randomly counter-balanced. In the immediate test condition (i.e., no retention interval), the test phase began immediately after all items had been rated and retracted or affirmed. The test phase involved a block of 80 inference questions (two per item) in random order, followed by a block of 40 direct belief ratings in random order. Participants in the 30-minute retention interval group completed an unrelated filler task before the test, and participants in the one-week group completed the test phase a week later—this test was administered in an online format in order to keep participation rates high. The test phase was identical regardless of retention interval.

**3.4.2 Results**

**Belief ratings.** Both pre-manipulation facts and myths attracted mid-range initial belief ratings, as expected, $M_{\text{facts}} = 5.69$, $SD_{\text{facts}} = .79$; $M_{\text{myths}} = 6.03$, $SD_{\text{myths}} = .97$. A within-subjects ANOVA comparing the pre-manipulation fact and myth belief ratings showed that participants initially believed the myths slightly more than the facts, $F(1,92) = 9.61; p = .003; MSE = .58; \eta_p^2 = .10$.

After participants read the affirmations/corrections, participants’ belief for facts increased, and belief for myths decreased, as shown in Figure 1. This belief change was sustained temporarily for both myths and facts, yet after a one-week
period belief for myths regressed. As post-manipulation belief levels remained below pre-manipulation belief levels, no true backfire effect was elicited.\(^4\)

A 2 × 2 × 3 within-between ANOVA (with factors type of item, type of explanation, and retention interval) was performed on the post-manipulation belief ratings. For this and all further statistical analyses, belief ratings and inference scores for myths were reverse-coded. This was to simplify the analysis and allow the type of explanation (brief vs. detailed) to register as a main effect rather than an interaction. The figures and discussion of the data trends are presented in the original untransformed format to facilitate interpretation.

\[\text{Figure 1. Post-manipulation belief ratings over time in Experiment 1. Dotted lines indicate the pre-manipulation belief ratings’ mean.}\]

\(^4\) Nyhan et al. (2014) found that corrective information regarding the flu vaccine reduced participants’ intent to vaccinate, but only in participants with high levels of concern about vaccine side effects. To address the assumption that backfire effects may only occur when correcting strong belief in the original misconception, the analysis was replicated using each participant’s 30% most strongly believed myths and 30% least believed facts. There was no backfire effect observed for this subset of materials—i.e. myths that participants correctly assumed to be false, and facts that participants correctly assumed to be true. At one week, myth belief was not statistically different from pre-manipulation levels, \(p > .05\).
The analysis revealed three significant main effects. The main effect of type of item (myth vs. fact), $F(1,90) = 27.57; p < .001; MSE = 1.68; \eta^2_p = .23$, indicated that overall fact belief ratings were closer to the ceiling than myth belief ratings were to the floor (Figure 1). The main effect of type of explanation (brief vs. detailed), $F(1,90) = 15.38; p < .001; MSE = .74; \eta^2_p = .15$, indicated that detailed explanations were slightly better at eliciting belief change than brief explanations. The main effect of retention interval, $F(2,90) = 4.78; p = .011; MSE = 5.40; \eta^2_p = .10$, indicated that the extent of belief change differed over time. This was qualified by a significant interaction of type of item and retention interval, $F(2,90) = 8.65; p < .001; MSE = 1.68; \eta^2_p = .19$, indicating that the change in belief over time was different for facts and myths, with fact belief remaining stable across intervals and myth belief rebounding over time (all other effects, $F < 1$).

Next, we ran a $2 \times 2 \times 2$ within-between ANOVA (factors type of item, type of explanation, and retention interval) restricted to the 30-minute and one-week retention intervals to clarify specifically whether the difference between fact and (reverse-coded) myth ratings was greater after a week than 30 minutes, or in other words, whether belief change was more stable over time for myths versus facts. The interaction between type of item and retention interval was significant, $F(1,61) = 13.90; p < .001; MSE = 1.89; \eta^2_p = .19$, indicating that belief ratings were stable for facts from 30 minutes to one-week, whereas belief ratings for myths increased during this time period.

Even on an individual level, the items showed a consistent pattern: the retracted myths were more likely to show regression towards their pre-manipulation levels, whereas beliefs in affirmed fact items were relatively sustained over time.
Only one myth item showed a numerically larger belief rating a week after correction compared to pre-manipulation belief levels.\(^5\)

**Inference ratings.** Even if participants were successfully discounting misinformation in the direct belief ratings, they could still be using misinformation in their reasoning. To address this question, we analysed participants’ mean inference scores. All inference scores were significantly correlated at the \(p < .05\) level with the respective belief ratings: myth-brief, \(r = .80\); myth-detailed, \(r = .78\); fact-brief, \(r = .70\); and fact-detailed, \(r = .67\). This indicates that inference questions supplied a valid indirect measure of belief.

A 2 × 2 × 3 within-between ANOVA was performed on the inference scores (with factors type of item, type of explanation, and retention interval). The results mimicked the pattern obtained with the post-manipulation belief scores, as Figure 2 illustrates. There were main effects of type of item, \(F(1,90) = 6.08; p = .016; MSE = 2.01; \eta_p^2 = .06\), and type of explanation, \(F(1,90) = 17.29; p < .001; MSE = .53; \eta_p^2 = .16\), as well as retention interval, \(F(2,90) = 6.17; p = .003; MSE = 3.96; \eta_p^2 = .12\). There was an interaction between type of item and retention interval, \(F(2,90) = 3.54; p = .033; MSE = 2.01; \eta_p^2 = .07\), indicating that the stability of scores across time differed for myths and facts. There was also a marginally significant interaction of type of explanation and retention interval, \(F(2,90) = 2.44; p = .093; MSE = .53; \eta_p^2 = .05\), suggesting that detailed explanations are particularly beneficial over time (all other effects, \(F < 1.72, p > .19\)).

---

\(^5\) This exception was ‘cancer screening is greatly beneficial’ in the brief explanation condition, which had a mean pre-manipulation belief rating of 5.04, which rose to 5.33 after one week.
Figure 2. Post-manipulation inference scores in Experiment 1 over time.

Analogous to the belief ratings analysis, a 2 × 2 × 2 within-between ANOVA (with factors type of item, type of explanation, and retention interval) was run testing specifically whether inference scores were less stable over time for myths versus facts in the 30-minute to one-week interval. The type of item by retention interval interaction was significant, $F(1,61) = 5.83; p = .019; MSE = 2.24; \eta^2_p = .09$, demonstrating that inference scores increased over a one-week period for myths in comparison to facts.

Returning to the omnibus 2 × 2 × 3 analysis, there was also a marginal interaction between type of explanation and retention interval, $F(2,90) = 2.44; p = .09; MSE = .53; \eta^2_p = .05$, suggesting that inference scores were more stable over time after detailed explanations compared to brief explanations. To corroborate this notion, an interaction contrast was run contrasting brief against detailed explanations and the pooled immediate and 30-minute intervals against the one-week interval (assigning lambda weights of 1, -1, 1, -1, to myth-brief, myth-detailed, fact-brief and fact-detailed, and 1, 1, -2 to the immediate, 30-minute, and one-week retention intervals, respectively). The contrast was significant, $F(1,90) = 4.44; p = .038; MSE = .53$, indicating that a detailed explanation had its greatest benefit after a long delay.
A detailed discussion of the Experiment 1 results will be deferred until the Experiment 2 data are presented.

### 3.5 Experiment 2

Experiment 1 showed that belief change was more sustained after fact affirmation compared to myth retraction. Experiment 2 was a conceptual replication of Experiment 1 but tested older adults. As we noted at the outset, it is possible that older adults are more strongly susceptible to the effects of familiarity, as older adults have less efficient strategic memory processes than young adults, whereas automatic processing is relatively age-invariant (Prull et al., 2006). While it is difficult to pinpoint the exact age at which recollection begins to decline, a study by Bender, Naveh-Benjamin, and Raz (2010) suggested a marked decline around the age of 40, and many studies investigating age-related differences in familiarity and recollection have used an older adult population with a mean age in the 60s (e.g. Aizpurua, Garcia-Bajos, & Migueles, 2009; Bastin & Van der Linden, 2003) or 70s (Anderson et al, 2008; Fernandes & Manios, 2012; Prull et al., 2006).

#### 3.5.1 Method

Experiment 2 was identical to Experiment 1, with two changes: (1) it was conducted with an older adult population; (2) an additional three-week retention interval condition was added to maximise the chances of eliciting the familiarity backfire effect, given the temporal stability of familiarity in contrast to the temporal volatility of recollection.

**Participants.** Participants were 124 older adults over the age of 50, who volunteered after reading an ethically-approved information sheet. Participants were recruited by advertising through the University of Western Australia website, Western Australian radio, and flyers around Perth. Participants were paid A$15 for
their participation. Participants were screened using the Montreal Cognitive Assessment (MoCA); thirteen participants were excluded as they scored below the normal range of 26 to 30 (Nasreddine et al., 2005). An additional two participants did not complete the task. Our final sample thus included $N = 109$ participants, with 39 males and 70 females between 50 and 87 years of age ($M = 64.37, SD = 8.91$).

**Procedure.** The procedure replicated Experiment 1, although prior to the study participants received the MoCA. One-week and three-week surveys were completed in an online format in order to keep participation rates high; two participants in the delayed conditions opted to receive paper copies of the survey. These were mailed back to the researchers once they had been completed.

**3.5.2 Results**

**Belief ratings.** A within-subjects ANOVA was performed on the pre-manipulation myth and fact belief ratings, which uncovered no significant differences between conditions, $M_{\text{facts}} = 6.10, SD_{\text{facts}} = 1.01; M_{\text{myths}} = 5.92, SD_{\text{myths}} = 1.04$. This indicates that prior to reading the explanations, participants believed myths and facts equally.

After participants read the explanations, the belief for facts increased and belief for myths declined, as can be seen in Figure 3. In striking similarity to Experiment 1, belief for facts was sustained over a one-week period, whereas belief for myths regressed between 30 minutes and one-week. Between week 1 and week 3, belief scores for both facts and myths regressed to a similar extent. As post-
manipulation myth belief levels remained below pre-manipulation belief levels, no true backfire effect was elicited.\textsuperscript{6}

\textbf{Figure 3.} Post-manipulation belief ratings in older adults in Experiment 2. Dotted lines indicate the pre-manipulation belief ratings’ mean.

For all further analyses, belief ratings and inference score ratings for myths were reverse-coded, as in Experiment 1. A $2 \times 2 \times 4$ within-between ANOVA on belief ratings was run, with within-subjects factors type of item (myth vs. fact) and type of explanation (veracity explained either briefly or in some detail), and the between-subjects factor retention interval (immediate, 30-minute, one-week, or three-weeks).

The analysis revealed three significant main effects. The main effect of type of item (myth vs. fact), $F(1,105) = 30.39; p < .001; MSE = 3.04; \eta_p^2 = .22$, indicated that fact ratings were closer to the ceiling than myth ratings were to the floor. The

\textsuperscript{6}To address the assumption that backfire effects may only occur when correcting strong belief in the original misconception, the analysis was replicated using each participants’ 30 % most strongly believed myths and 30 % least believed facts. The trend in was replicated, and no backfire effect was elicited.
main effect of type of explanation (brief vs. detailed), $F(1,105) = 14.91; p < .001; MSE = 1.08; \eta^2 = .12$, indicated that detailed explanations were better at eliciting belief change than brief explanations, and the main effect of retention interval, $F(3,105) = 11.56; p < .001; MSE = 5.36; \eta^2 = .25$, indicated that belief change differed over time. A significant interaction of type of item and retention interval, $F(3,105) = 4.37; p = .006; MSE = 3.04; \eta^2 = .11$, indicated that the change in belief over time was different for facts and myths, with fact belief remaining largely stable across intervals and myth belief increasing over time. The interaction of type of item and type of explanation was also significant, $F(1,105) = 4.75; p = .031; MSE = 1.11; \eta^2 = .04$, indicating that detailed explanations were slightly more effective for facts than for myths. Lastly, the interaction of type of explanation and retention interval indicated that detailed explanations promoted belief change better than brief explanations particularly after a long delay, $F(3,105) = 3.83; p = .012; MSE = 4.11; \eta^2 = .10$. The remaining interaction of type of item, type of explanation, and retention interval remained non-significant, $F < 1$.

In subsequent contrast analyses, we focused first on the 30-minute and 1-week retention intervals (analogous to Experiment 1). An interaction contrast between type of item (myth vs. fact) and retention interval (30-minute vs. one-week) demonstrated that the belief difference between myths and facts was greater after one week than 30 minutes, $F(1,52) = 8.11; p = .006; MSE = 2.49; \eta^2 = .13$. Thus, fact belief remained stable over time whereas myth belief increased over the period of one week.

Focusing on retention intervals of 1 and 3 weeks, the analogous type of item by retention interval contrast was not significant, $p > .05$, while a contrast comparing week-1 and week-3 ratings collapsing all conditions across item and explanation levels was significant, $F(1,52) = 7.08; p = .010; MSE = 4.79$, indicating that from
week 1 to week 3, fact and myth belief ratings regressed equivalently. In other words, item validity, in general, was being forgotten between a one and three-week period.

As post-manipulation myth belief significantly correlated at the $p < .05$ level with age for both brief retractions, $r = .21$, and detailed retractions, $r = .28$, a final set of belief-rating analyses looked at age at a finer level of granularity. Specifically, to further address the assumption that myth-belief updating deteriorates with age, a median-split analysis comparing participants aged 50-64 (“middle-aged” participants) with those 65 and older (“old” participants) was conducted (see Figure 4). Investigating type of explanation (brief vs. detailed), age (middle-aged vs. old), and retention interval (immediate vs. 30-minute vs. one-week), a $2 \times 2 \times 4$ within-between ANOVA on myth beliefs was performed. This analysis yielded a main effect of age, $F(1,101) = 8.73; p = .004; MSE = 5.29, \eta_p^2 = .08$, indicating that old participants were less likely to show sustained myth-belief change than middle-aged participants, and a main effect of retention interval $F(3,101) = 12.73; p < .001; MSE = 5.29, \eta_p^2 = .27$, indicating that belief changed over time (all other effects, $F < 2.77, p > .07$). Moreover, a $2 \times 2 \times 2$ within-between ANOVA (with factors type of explanation, retention interval, and age) focusing on one-week and three-week retention intervals revealed a type of explanation by age interaction, $F(1,50) = 4.07; p = .049; MSE = 1.38; \eta_p^2 = .08$, indicating that after longer delays, detailed retractions led to more sustained belief change in comparison to brief retractions for middle-aged but not old participants.7

---

7 A $2 \times 2 \times 3$ ANOVA also including the young adults from Experiment 1 (with factors type of explanation [brief and detailed], retention interval [30 minutes and one-week], and age [young adults, middle aged, and old], on post-explanation myth scores likewise revealed a main effect of age, $F(1,111) = 3.43; p = .036; MSE = 4.87; \eta_p^2 = .06$. An interaction contrast revealed that young adults
For the sake of completeness, a $2 \times 2 \times 4$ within-between ANOVA on fact beliefs was performed (with factors type of item, type of explanation, and retention interval). The analysis yielded significant main effects of type of explanation, $F(1,101) = 19.20; p < .001; MSE = 1.03; \eta^2_p = .16$, showing that detailed explanations were more effective than brief ones, age, $F(1,101) = 5.60; p = .020; MSE = 2.61; \eta^2_p = .05$, indicating that old participants showed less belief change than middle-aged participants, and retention interval, $F(3,101) = 5.00; p = .003; MSE = 2.61; \eta^2_p = .13$, showing that belief changed over time (all other effects, $F < 1.23, p > .30$). In an analysis confined to the immediate, 30-minute, and one-week conditions, the retention interval effect was non-significant, $p > .05$, demonstrating that for the duration of one week, fact belief was sustained.

and middle aged participants were equivalently better at reducing their belief in misconceptions than older adults, $p = .015$. Unlike the above analysis, there is no interaction of explanation and age—presumably because we could not include the week 3 ratings in the analysis—however, a planned comparison contrasting old adults against pooled young and middle aged adults, and brief explanations against detailed explanations, approached significance, $F(1,111) = 3.03; p = .084; MSE = 1.15$. 

---

Figure 4. Post-manipulation myth belief ratings presented in an age-based median split in the older adult sample.
Inference ratings. Returning to the analysis of the full sample, inference scores are presented in Figure 5. A 2 × 2 × 4 within-between ANOVA (with factors type of item, type of explanation, and retention interval) on the inference scores revealed main effects of type of item, $F(1,105) = 13.47; p < .001; \text{MSE} = 2.00; \eta^2_p = .11$, type of explanation, $F(1,105) = 14.89; p < .001; \text{MSE} = .85; \eta^2_p = .12$, and retention interval, $F(3,105) = 6.43; p < .001; \text{MSE} = 4.24; \eta^2_p = .16$, as well as an interaction between type of item and retention interval, $F(3,105) = 3.75; p = .013; \text{MSE} = 2.00; \eta^2_p = .10$, suggesting that the stability of scores across time differed for facts and myths. An interaction contrast, analogous to Experiment 1, testing whether inference scores were more stable for facts versus myths in the 30-minute to one-week interval, was significant, $F(1,52) = 9.01; p = .004; \text{MSE} = 1.74; \eta^2_p = .15$.

Figure 5. Post-manipulation inference scores in an older adult population in Experiment 2.

There was also an interaction of type of explanation and retention interval $F(3,105) = 3.38; p = .021; \text{MSE} = .85; \eta^2_p = .09$, indicating that inference scores were more stable across time after detailed explanations compared to brief explanations. To corroborate this notion, an interaction contrast was run contrasting brief against detailed explanations and the pooled immediate and 30-minute intervals.
against the pooled one-week and three-week intervals. The contrast was significant, $F(1,105) = 10.07; p = .002; MSE = .85$, indicating that a detailed explanation had its greatest benefit after a long delay (all other effects, $F < 1.30, p > .25$).

### 3.6 General Discussion

The present research aimed to determine the parameters of differential forgetting of myth and fact veracity over time, in order to clarify if and under what conditions familiarity may contribute to false acceptance of corrected myths as true. Dual-process accounts of continued influence effects of misinformation (e.g., Ecker et al., 2010) suggest that post-correction reliance on misinformation can be based on automatic memory processes (i.e., myth familiarity) in the absence of strategic retrieval and control processes. Hence familiarity-based acceptance of corrected falsehoods could be a mechanism underlying continued influence effects of misinformation. To investigate this, we presented participants with both myths and facts, obtained a pre-manipulation belief rating, then corrected the former and affirmed the latter. We manipulated factors known to affect strategic memory processes, thus varying the relative impact of familiarity. Specifically, we manipulated the explanations’ level of detail and retention interval, and contrasted age groups, and we measured how these factors affected people’s post-explanation beliefs and inferences.

While some studies have shown a continued influence effect after a brief retention interval (e.g. Ecker et al., 2011; Johnson & Seifert, 1994), our corrections (and affirmations) were found to be relatively effective in the short-term. This difference may be due to the fact that, unlike the typical continued-influence paradigm, we retracted simple statements rather than causal relationships regarding an event, which may be particularly resistant to correction. The short-term efficacy
of the explanations was more apparent for direct belief ratings (e.g., see Figure 1), whereas our inference measure (e.g., see Figure 2) closely resembled the typical result pattern found in continued-influence studies, which often also use inference questions to assess misinformation effects.

3.6.1 Differential Forgetting of Myths and Facts Over Time

Across both experiments, we found a striking asymmetry in that belief change was more sustained after fact affirmation compared to myth retraction—retractions thus seemingly have an “expiration date”. This asymmetry could be partially explained by familiarity. In the case of an affirmed fact, it does not matter if an individual relies on the recollection of the affirmation or on the boosted familiarity of the factual statement—familiarity and recollection operate in unison and lead to the individual assuming the item to be true. However, in the case of a retracted myth, recollection of the retraction will support the statement’s correct rejection, whereas the myth’s boosted familiarity will foster its false acceptance as true, as familiarity and recollection stand in opposition (Jacoby, 1991).

Our inference results mirrored the trend obtained with the belief ratings, demonstrating that familiarity effects can extend to inferential reasoning and potentially decision making. It is even possible that the act of responding to inference questions can contribute to increased familiarity of the misconception, in that the information is subjectively re-experienced during memory retrieval following exposure to the inference question, once again leading to a potentially increased perception of validity (Ozubko & Fugelsang, 2011).

3.6.2 Age and Level of Detail

Overall, the pattern of belief change over time—and in particular the asymmetry between facts and myths—was similar in young and older participants. Even young adults’ recollection fades over time, leading to an increased reliance
upon familiarity in judging the veracity of information (Gilbert et al., 1990). However, “old” participants aged 65 and over were found to be comparatively worse than those aged 50-64 (“middle-aged” participants) at sustaining their post-correction belief that myths are inaccurate. This supports the notion that older adults have less efficient strategic memory processes and thus less effective retrieval of the link between an item and contextual details (Naveh-Benjamin, 2000; Prull et al., 2006).

As the mnemonic link between a statement and its veracity is weaker in older adults (Glisky et al., 2001), they seem particularly susceptible to the “re-believing” of myths. Although there was also a significant difference in fact belief between the “middle-aged” and “old” groups, this reflected the fact that the old participants were less likely to initially update their belief immediately after the affirmation. This differed from myth belief where belief change immediately after a correction was substantial yet followed by relatively steep forgetting as time progressed.

Detailed refutations seemed to somewhat mitigate the negative impact of familiarity in both younger and middle-aged adults. This is supported by parts of the educational literature, which highlight the benefits of detailed refutations (Tippett, 2010). Refutations may encourage participants to detect inconsistencies between their own inaccurate beliefs and the corrective information, leading to a facilitation of belief change even over long delays (Bedford & J. Cook, 2013; Guzzetti, 2000; Kowalski & Taylor, 2009). The benefit of directly addressing misconceptions could additionally be explained by detailed explanations fostering scepticism regarding the initial misinformation or its source (cf. Lewandowsky, Stritzke, Freund, Oberauer, & Krueger, 2013; Lewandowsky, Stritzke, Oberauer, & Morales, 2005). However, as much of this research stems from the educational literature, it has mostly used undergraduates or school-age participants (Guzzetti et al., 1993). The current study
found that for “old” adults over the age of 65, correcting myths using detailed refutations was as ineffective as brief retractions.

### 3.6.3 The Familiarity Backfire Effect

The present research provides evidence for familiarity causing an increase in post-correction myth belief after a delay; this meshes well with previous studies that similarly reported that myths are often “misremembered” as facts over time (Peter & Koch, 2016; Skurnik et al., 2005; Skurnik et al., 2007). However, we found no evidence for the existence of a true familiarity-based backfire effect. As in these previous studies, the corrections did help participants update their beliefs in the right direction—that is, myth beliefs were reduced by the corrections. Corrections repeating the myth were simply less effective (compared to fact affirmations) rather than backfiring.

The lack of a familiarity backfire effect conforms to a range of theoretical proposals which suggest that repeating misinformation when correcting could even facilitate belief updating. Stadtler, Scharrer, Brummernhenrich, and Bromme (2013) as well as Putnam, Wahlheim, and Jacoby (2014) proposed that the detection of conflict—which is arguably made more salient through repetition of the misinformation during its retraction—is beneficial for updating. Reconsolidation theory likewise argues that reminders of to-be-corrected information will labilise its memory representation, thereby facilitating updating (Hardt, Einarsson, & Nader, 2010). Finally, Kendeou, Walsh, Smith, and O’Brien, (2014) argued that outdated and new information must be co-activated for knowledge revision to occur. This is consistent with a study by Pashler, Kang, and Mozer (2013), who found that repeating the original misinformation prior to learning new information enhanced memory for the new information when tested one week later.
This implies that future research still faces a conundrum: while the present findings suggest that false acceptance of corrected myths as true is at least partially driven by familiarity, it seems that corrections that do not repeat the myth may be even less effective than corrections that do repeat the myth (e.g., Ecker, Hogan, & Lewandowsky, 2017; Wilkes & Leatherbarrow, 1988). In other words, if a myth is not repeated when corrected, the associated lack of salience, conflict detection, and/or myth/correction co-activation may be even more detrimental to belief updating than the boost of the myth’s familiarity.

### 3.6.4 Potential Limitations and Future Directions

Obtaining belief measures prior to the experimental manipulation could be considered a limitation as it may have influenced how the corrective explanations were processed. However, in our opinion it is likely that a person’s belief is spontaneously cued when a statement of unclear veracity (e.g., a potentially dubious news headline) is encountered, or when a correction is presented by itself (e.g., if one is told that listening to Mozart does not increase IQ, it seems likely that one would consider whether or not one believes the original claim). Thus, asking for an explicit expression of belief prior to a correction will not necessarily have a strong impact on how the correction is processed. In our view, from a methodological perspective, the advantages of a pretest-posttest design outweigh the disadvantages. “Posttest-only with control” designs as used by Skurnik et al. (2005)—where one group received the correction and another group received no correction—can be considered quasi-experimental as the treatment and control groups cannot be adequately compared at baseline (Morris, 2008). This potentially reduces internal validity because the differences at posttest may be artificially inflated (T. D. Cook & Campbell, 1979; Morris & DeShon, 2002).
The artificial nature of the task could be seen as another limitation, as participants evaluated a long series of statements. However, people often process a large number of news headlines in a short period of time (e.g., when skimming a newspaper or scanning one’s social media feed), arguably assessing or at least monitoring the truth/belief status of each. Thus, we argue that people routinely deliberate belief prior to correction (i.e., in an experimental context, before a post-correction belief rating), even with large numbers of statements.

We have interpreted our finding that myths are more likely than facts to be misremembered after a delay as an effect of familiarity when strategic memory is limited. The present research focused on factors that influence strategic memory processes; future research could test the proposed relationship between familiarity and misinformation effects more directly, for example by correcting statements that are familiar to some participants but not others. Previous research has found that misinformation effects are particularly strong if the misinformation is repeatedly presented before a correction (Ecker et al., 2011, also see Weaver, Garcia, Schwarz, & Miller, 2007), in line with the familiarity notion.

Moreover, future research could apply alternative testing procedures to further investigate the mechanisms underlying the effects reported here. For example, if myth acceptance is familiarity-driven, one might expect corrected myths to be accepted as true particularly in tasks requiring true/false categorisation of statements (which may be more recognition-based) rather than in tasks that have a stronger recall component.

3.6.5 Practical Applications

The applied goal of this research was to provide empirically-based advice on how to correct misconceptions. The present data suggest the following: First, corrections should include details as to why the misinformation is incorrect, as
detailed refutations are more effective than brief retractions, particularly with younger participants. Thus the misinformation should be explicitly retracted and paired with a comprehensive rebuttal.

Second, even the efficacy of detailed refutations of familiar misconceptions will lessen over time, and important corrections may need to be provided repeatedly, despite the potential risks of further boosting the myth’s familiarity (also see Ecker et al., 2011a). While this recommendation seems somewhat ironic in the context of the boosted-familiarity notion, boosting the more volatile recollection of the correction to offset myth familiarity may be necessary to achieve enduring belief change.

Third, explicitly mentioning a familiar misconception within a retraction will not typically backfire in the true sense of the word (this qualifies earlier recommendations; e.g., J. Cook & Lewandowsky, 2011; Lewandowsky et al., 2012). Repeating the myth when retracting it may be crucial for belief updating because it increases the correction’s salience and fosters conflict detection and co-activation of myth and correction (Kendeou et al., 2014; Putnam et al., 2014; Stadtler et al., 2013). However, given the aforementioned trade-off between the harm from boosting myth-familiarity and the benefit from boosting recollection of the correction (e.g. the association of the myth and its “negation-tag”), theoretically there may be circumstances where the harm outweighs the benefit. Moreover, it may also be problematic to circulate corrections if individuals have not previously encountered the relevant misconception, as this may potentially make the misinformation familiar to new audiences (Schwarz, Newman, & Leach, 2016). It follows that, after correcting a myth, the focus should be placed upon factual information as much as possible in order to avoid boosting myth familiarity more than necessary (cf. Ecker et al., 2010; Johnson & Seifert, 1994; Lewandowsky et al., 2012; Seifert, 2002).
3.7 References


Science & Education


doi:10.1017/S0007123415000186


http://doi.org/10.1037//0278-7393.26.5.1170


http://doi.org/10.1007/s11109-010-9112-2


http://doi.org/10.1542/peds.2013-2365


Chapter 4: The Efficacy of Diverse Explanation Formats
4.1 Foreword

Chapter 3 found that even in circumstances that should be most conducive to the familiarity backfire effect (i.e., in older adults, after a long retention interval, when the correction provided little detail as to why the misinformation is incorrect), it was not elicited. However, familiarity was indeed shown to play a role for the continued influence effect: conditions where automatic memory processes were more likely to be relied upon led to corrections being less effective compared to conditions where strategic memory processes were more likely to be employed. Therefore, it could be potentially beneficial to avoid repeating misinformation within the retraction. Although Chapter 3 found that corrections are unlikely to backfire even if they repeat the myth, it did not directly test whether removing the misconception from the correction itself leads to better updating.

Chapter 4 first and foremost aims to investigate whether corrections omitting the retraction are more effective. It also addresses a broader research question: Does presenting the same correction in different ways change the efficacy of belief updating over time? There are many different ways to correct misconceptions and affirm factual information, and it is possible that certain explanation formats are better than others at promoting belief change (Cook, Lewandowsky, & Ecker, 2017; Lewandowsky et al., 2012). If one particular explanation format is better than others, this will have practical benefits for fact checkers and other communication professionals. It will also be highly relevant for teasing apart theoretical frameworks for why belief in misinformation persists.

Previous research has compared different methods of correcting misinformation in the broader sense. For example, refutational text (presenting the misconception prior to the explanation) has been compared with class discussions.
and demonstrations (Guzzetti, 2000). Guzzetti (1993) performed a meta-analysis including numerous studies that had measured belief revision over time, but these experiments were largely performed with primary school children. Chapter 4 presents empirical research in adults that measured belief change from numerous different correction formats over both a one-week and three-week retention interval. We based our materials on the common “myths vs. facts” format, manipulating it in various ways. Specifically, across different conditions, we reversed the components of the explanatory text (i.e., presenting the explanation prior to the original claim or vice versa), we corrected myths without affirming facts, or presented fact-focused corrections that did not repeat the misinformation.
4.2 Abstract

This study aimed to explore whether there was an optimal explanation frame that is best for correcting misinformation over time. Participants were presented with a range of myths and facts in different formats, which differed in the way the statements and their myth/fact status were explained. The formats were compared against the common myths vs. facts format where the statement is provided first, followed by a true/false label and an explanation. Conditions were designed to test theoretical notions of why misinformation persists and how to facilitate belief updating. For example, (1) the notion that corrections should not repeat the myth to avoid boosting myth familiarity was tested with a “facts-only” format where all myths were re-framed as factual statements; (2) the notion that presenting explanations regarding both myths and facts may subsequently result in source confusion was tested with a “myths-only” format where myths were corrected but factual items were not affirmed; (3) the notion that the explanation should be provided prior to the misconception to avoid the need for backtracking and editing an existing representation was tested with a reverse “facts-first” format. Participants were either asked to rate their belief in the items after a retention interval of one week or three weeks. The results indicated that belief change was largely independent of the specific format of explanation used.
4.3 Introduction

Misinformation continues to influence an individual’s reasoning and memory even after a clear correction has been elicited (Johnson & Seifert, 1994; Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012). Thus, finding an optimum correction method could be highly beneficial. Aside from having clear practical ramifications, if some correction types are more effective than others, this could help tease apart various theoretical notions of why people continue to believe in or are influenced by inaccurate information. This paper investigates the relative efficacy of numerous correction formats and the degree to which these different methods promote belief change over time.

A common method used in public information campaigns is what we will refer to as the “standard” explanation format. The standard explanation format is where first a claim is presented, followed by a true or false label, and subsequently there is an explanation as to why the information is true/false (Guzzetti, Snyder, Glass, & Gamas, 1993). This form of retracting false information is known as a refutation in the educational literature, and has been demonstrated to promote greater sustained belief change in comparison to other educational strategies such as only providing valid factual content (i.e., only the correction), and presenting corrective information without a description of the misconception itself (Guzzetti, 2000). Kendeou, Walsh, Smith, and O’Brien (2014) suggested that the reason for the refutation’s relative success may be that co-activation of the misconception and new

---

8 To clarify, where Chapter 3 showed that factual affirmations were more effective in comparison to misinformation retractions, the educational literature rather compares the presentation of the corrective information alone in comparison to presenting the misinformation prior to the corrective information.
corrective information is essential for knowledge revision to occur (also see Kendeou, Butterfuss, van Boekel, & O’Brien, 2017).

4.3.1 The Avoidance of Familiarity

While directly addressing misconceptions in this sequential manner appears to be ideal for producing sustained belief change, research by Skurnik, Yoon, and Schwarz (2007) provides evidence to the contrary. This unpublished manuscript posited that presenting retractions in this standard format might instead contribute to the rejection of the correction and the false acceptance of myths. The authors presented participants with information about the flu vaccine in two versions of an informative flyer. The flyer was either presented in the standard “myths vs. facts” explanation format, or in a “facts-only” format that affirmed the facts but did not mention the myths at all. The former was a real-world flyer used by the US Centre for Disease Control and Prevention (CDC), and the latter was a variation of this.

Both flyers immediately increased participants’ intentions to be vaccinated; however, after 30 minutes the increased vaccination intention only persisted for the “facts-only” flyer. Participants who read the “myths vs. facts” flyer had less favourable vaccination attitudes than participants who had never seen the flyer at all. In addition, participants who read the “myths vs. facts” flyer correctly distinguished the myths from the facts immediately after the flyer was presented, yet 30 minutes later they misidentified 15% of the myths as facts, in comparison to only 2% of the facts being misidentified as myths. (The percentage of facts mistaken as myths in the “facts-only” flyer condition was not reported.) Skurnik et al. attributed this outcome to the increased familiarity caused by the repetition of the inaccurate information when retracting the myth, thus recommending that mentioning misconceptions directly in a correction should be avoided.
It is theoretically possible that repeating the original misconception within the retraction could hinder its corrective impact. It is known that people are more likely to believe information when it is familiar (Begg, Anas, & Farinacci, 1992; DiFonzo, Beckstead, Stupak, & Walders, 2016). Based on this notion and the findings of Skurnik et al. (2007; also see Peter & Koch, 2016; Skurnik, Yoon, Park, & Schwarz, 2005), it has been suggested that a familiarity boost associated with a myth-repeating correction can cause a familiarity backfire effect, where belief in the misconception ironically becomes stronger after a correction is presented (Cook & Lewandowsky, 2011; Lewandowsky et al., 2012). While Swire, Ecker, and Lewandowsky (2017; i.e., Chapter 3) found no evidence that correcting misinformation led to an actual backfire effect, they concluded that familiarity was still a contributing factor to the persistence of misinformation after a correction. Thus, there is sufficient motivation to explore alternative correction formats. For instance, framing a retraction as a factual affirmation could still be more effective than the standard retraction format. In other words, rather than stating that “the claim that people only use 10% of their brain is false”, one could focus on the true statement that “people use 100% of their brain”. This method does not mention the original myth, therefore avoiding increased familiarity, while still correcting the underlying misconception.

4.3.2 Source Confusion

An alternative explanation for the efficacy of Skurnik et al.’s (2007) “facts-only” flyer—other than reduced familiarity from avoiding repetition of the misconceptions—is that participants may have experienced less confusion. Not only did participants have fewer items to remember (compared to the standard myths vs. facts format), but the claims for which they received explanations were all true. It is possible that presenting all items with the same valence can help participants avoid a form of retrieval failure known as source confusion, where people confuse or
misattribute the contextual details of a memory (Johnson, Hashtroudi, & Lindsay, 1993; Schacter & Dodson, 2001). Thus, presenting only the facts, or alternatively, just the myths associated with a topic, might be better at promoting sustained belief change. This makes intuitive sense: Participants will be able to think back to the encoding phase knowing that all the claims encountered in that encoding context were either true or false.

4.3.3 Explanation Prior to the Misconception

Aside from investigating the potential impact of source confusion and familiarity, the effects of rearranging the correction components could have additional theoretical ramifications. Specifically, rearranging the standard format (i.e., myth—false-label—retraction) to the reverse (i.e., retraction—false-label—myth) would mean that the individual could avoid backtracking and editing an existing representation. Participants would already have the information as to why the misinformation is incorrect before they receive it. As the misconception is only presented in contrast to the fact, participants are more cognitively prepared and potentially more likely to encode the misinformation strategically and integrate it with the fact (Ecker, Lewandowsky, & Tang, 2010; Kendeou & O’Brien, 2014).

Moreover, if presenting the explanation prior to the misconception is better (or worse) than the standard format, this could be seen as evidence for a “query order” effect (Appelt, Hardisty, & Weber, 2011). Query theory suggests that people consider options actively and sequentially, and therefore the order of the presentation and consideration of components can impact judgments and decision-making (Weber et al., 2007). This implies that the component with the greatest informational value should be placed at the most impactful position in the order. Alas, it is unclear which component carries the greatest informational value—although one could argue it is likely to be the retraction—and it is also unclear what the most impactful position is.
Literature regarding primacy effects (and query theory itself) suggests that the most important information should come first (e.g., Farrell & Lewandowsky, 2002; Page & Norris, 1998), as presenting information earlier allows for more rehearsal (Tan & Ward, 2008) and increases temporal distinctiveness due to an absence of proactively interfering information (Brown, Neath, & Chater, 2007; Ecker, Tay, & Brown, 2015). For example, in impression formation, more emphasis tends to be placed on information received about a person early rather than later (e.g., Dreben, Fiske, & Hastie, 1979). By contrast, literature on recency effects suggests the most important information should come last, as recently acquired representations are stronger in memory (e.g., Baddeley & Hitch, 1993) and more temporally distinct because of lack of retroactive interference (Brown et al., 2007; Ecker, Brown, & Lewandowsky, 2015). For example, Ecker, Lewandowsky, Cheung, and Maybery (2015) found that when people were presented with multiple causes of an event, the more recent cause tended to have the strongest influence on memory and reasoning.

**4.3.4 The Current Experiment**

This paper investigated the efficacy of different explanation formats, in an endeavour to investigate whether one format is superior to others in order to reduce the continued influence effect. In addition, we hoped to assess various theoretical frameworks regarding why misinformation often persists. In an encoding phase, participants were presented with a range of facts and myths from various topics, and were then presented with affirmations and retractions; they were then asked to rate their beliefs in the presented claims in a test phase after a varying retention delay.

We used five different explanation formats: (1) the *standard* “myths vs. facts” format, where all myth and fact claims were first presented followed by a “true/false” label and then an explanation (i.e., a retraction or affirmation); (2) a *reverse* format, where the explanation and the “true/false” label were provided before
the original claims; (3) a facts-only format, where all myths were re-framed as factual statements, thus avoiding myth repetition; (4) a myths-only format, where the factual items were omitted to avoid source confusion; and (5) a no-explanation control condition, which involved no encoding phase and only belief ratings at test (for an illustration of the components included in each condition, see Table 2).

Thus, the experiment used a $2 \times 2 \times 5$ mixed design, with within-subjects factors type of item (myth vs. fact) and type of explanation (standard vs. reverse vs. facts-only vs. myths-only vs. control), and the between-subjects factor retention interval (one week vs. three weeks between encoding and test). We were primarily interested in the efficacy of myth retractions, but also present the data from fact affirmations.

Table 2

<table>
<thead>
<tr>
<th>Format</th>
<th>Items presented</th>
<th>Order of components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>3 myths</td>
<td>(i) claim</td>
</tr>
<tr>
<td></td>
<td>3 facts</td>
<td>(ii) true / false label</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(iii) affirmation / retraction</td>
</tr>
<tr>
<td>Reverse order</td>
<td>3 myths</td>
<td>(i) affirmation / retraction</td>
</tr>
<tr>
<td></td>
<td>3 facts</td>
<td>(ii) true / false label</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(iii) claim</td>
</tr>
<tr>
<td>Facts-only</td>
<td>3 myths (framed as facts)</td>
<td>(i) claim</td>
</tr>
<tr>
<td></td>
<td>3 facts</td>
<td>(ii) true / false label</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(iii) affirmation / retraction</td>
</tr>
<tr>
<td>Myths-only</td>
<td>3 myths</td>
<td>(i) claim</td>
</tr>
<tr>
<td></td>
<td>0 facts</td>
<td>(ii) false label</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(iii) retraction</td>
</tr>
<tr>
<td>No explanation control</td>
<td>0 myths</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>0 facts</td>
<td></td>
</tr>
</tbody>
</table>

If the facts-only format was found to be superior to the standard format in promoting belief change, it would be evidence that repeating the myth, thus boosting
its familiarity, is indeed detrimental to its correction. On the other hand, if the myths-only format promoted more sustained belief change, it would suggest that source confusion is a key hurdle for effective communication of myths and facts. Finally, if the reverse format was better (or worse) than the standard format, this would be preliminary evidence that a query order effect is occurring during the processing of corrections and affirmations. It could also help identify whether presenting information without preceding or concurrent emphasis on its myth status is beneficial or detrimental.

4.4 Method

Participants. Participants were 123 undergraduate students from the University of Western Australia, who received course credit for participation. Three participants who did not complete the study and 10 participants who completed the study more than two days after the scheduled date were excluded from the following analysis. The final pool thus included $N = 110$ participants, with 33 males and 77 females between 17-30 years of age, and a mean age of $M = 19.74$ ($SD = 2.88$).

Stimuli. There were five sets of items, each consisting of three myths and three facts. Each set was concerned with a different topic: the brain, alcohol, animals, hypnotism, and the flu. Stimuli from the “flu” topic were taken directly from the Skurnik et al. (2007) experiment. An example myth in the standard format, reverse order, and facts-only format can be found in Table 3.

Belief scores were rated on an 11-point (0-10) scale ranging from “Definitely True” to “Definitely False”. For every item, there was also an inference question designed to gain a less direct measure of belief. These were included (following ample precedent, see Ecker et al., 2011) because people sometimes rely on misinformation in their inferential reasoning even when they exhibit successful
discounting in direct belief ratings. The inference questions were rated on an 11-point scale, with the specific scale-value range varying from item to item (i.e., some were on a 0-10 scale, while others were on a 0-20 % scale with 2 % increments, etc.). The full list of stimuli may be viewed in Appendix B.

Table 3

Example of standard format myth retraction, reverse order retraction and facts-only frame.

<table>
<thead>
<tr>
<th></th>
<th>Most people only use between 10 and 50% of their brains.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard retraction</td>
<td>Most people only use between 10 and 50% of their brains - MYTH</td>
</tr>
<tr>
<td></td>
<td>We use all of our brain: Brain imaging techniques have demonstrated that our whole brain is active, at least to some extent, all of the time. Specific areas of the brain will become more active depending on the demands of the task. The brain, like all other organs, has been shaped by natural selection. It would be extremely costly and inefficient for our body to produce material that it was not going to use.</td>
</tr>
</tbody>
</table>

| Reverse order retraction | We use all of our brain: Brain imaging techniques have demonstrated that our whole brain is active, at least to some extent, all of the time. Specific areas of the brain will become more active depending on the demands of the task. The brain, like all other organs, has been shaped by natural selection. It would be extremely costly and inefficient for our body to produce material that it was not going to use. |
|                         | It is a MYTH that people only use between 10 and 50% of their brains |

| Facts-only frame | People use 100% of our brains. |
|                 | People use 100% of our brains - FACT |
|                 | We use all of our brain: Brain imaging techniques have demonstrated that our whole brain is active, at least to some extent, all of the time. Specific areas of the brain will become more active depending on the demands of the task. The brain, like all other organs, has been shaped by natural selection. It would be extremely costly and inefficient for our body to produce material that it was not going to use. |
**Procedure.** The encoding phase took place in the laboratory, and all materials were presented on a computer screen. Participants were presented with four out of five sets of items—the set that was not presented was allocated to the control condition. All sets were presented in a random order, and the items within the set were also presented in random order. All items of each experimental set were corrected/affirmed using one of the four explanation formats. The assignment of sets to conditions was counterbalanced across participants. The test phase followed either one or three weeks after the encoding phase. The test involved a block of inference questions (one per item, in random order), and a block of direct belief ratings. The test was administered in an online format in order to keep participation rates high.

4.5 Results

**Belief scores.** Mean belief scores across conditions are shown in Figure 6. People were slightly more likely to correctly label myths to be false and facts to be true after one week than after three weeks. It appears that the correction format made very little difference to the belief change achieved.9

A 2 \( \times \) 5 within-between ANOVA with factors retention interval (one week vs. three weeks) and type of explanation (standard vs. reverse vs. facts-only vs. myths-only vs. control) was performed on the myth belief scores, revealing two main effects. The main effect of retention interval, \( F(1,85) = 7.49; p = .008; MSE = 10.81; \eta^2_p = .08 \), indicated that myth belief increased between one and three weeks. The

9 Due to an error with randomisation, the “alcohol” topic was not presented to participants in the standard format. On average, in the remaining conditions, the inclusion of the alcohol topic made the myth statements more believable by 0.02 and 0.21 for one and three-week conditions respectively, and the factual statements less believable by 0.12 and 0.28 after one week and three weeks respectively. Thus, the absence of the alcohol items in the following analysis may have led to a slight artificial inflation of the apparent effectiveness of the standard format to reduce misconceptions and increase fact belief, compared to the other explanation formats.
main effect of type of explanation, \( F(4,340) = 22.57; p < .001; MSE = 3.57; \) \( \eta^2_p = .21, \) indicated that the explanation formats differed. From Figure 6 it can be seen that all explanation formats led to sustained belief change in comparison to the control condition where no retraction was presented.

![Figure 6](image)

*Figure 6.* Belief scores in Experiment 3 over time by explanation formats. Dotted lines illustrate control condition.

We performed a number of multiple comparisons to confirm that (1) all formats differed from control, and that (2) the standard format was not different from the other correction formats. Table 4 shows these comparisons, collapsed over both retention intervals. After controlling for multiple comparisons using the Holm-Bonferroni method (Holm, 1979), we confirmed that all correction formats indeed differed from the control. However, all alternative correction formats (i.e. reverse, facts-only, and myths-only) had equal efficacy to the standard format.
Table 4

F and p values of myth condition belief ratings planned comparisons

<table>
<thead>
<tr>
<th></th>
<th>Standard</th>
<th>Reverse</th>
<th>Facts-only</th>
<th>Myths-only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reverse</td>
<td>$F = 4.26$</td>
<td>$p = .042$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facts-only</td>
<td>$F = 2.67$</td>
<td>$p = .11$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myths-only</td>
<td>$F = .91$</td>
<td>$p = .34$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>$F = 55.49$; $p &lt; .001^*$</td>
<td>$F = 31.96; p &lt; .001^*$</td>
<td>$F = 33.40; p &lt; .001^*$</td>
<td>$F = 51.23; p &lt; .001^*$</td>
</tr>
</tbody>
</table>

Note. All df1 = 1, df2 = 85; * indicates significance after applying the Holm-Bonferroni adjustment for multiple (n = 7) comparisons.

For the fact beliefs, a $2 \times 5$ within-between ANOVA was also performed, revealing two main effects. The main effect of retention interval, $F(1,85) = 4.78; p = .032$; $MSE = 5.52$; $\eta_p^2 = .08$, indicated that fact belief decreased between one and three weeks, and a main effect of type of explanation $F(4,340) = 90.53; p < .001$; $MSE = 3.10$; $\eta_p^2 = .52$, indicated that fact belief differed across explanation formats. In an equivalent manner to the myth items, planned comparisons were performed, as shown in Table 5. After the Holm-Bonferroni adjustment was applied, all affirmation formats were found to be significantly different from control.

The myths-only format (where we presented participants only with three myths and no factual affirmations) was the only affirmation format to differ from the standard format. As can be seen from Figure 6, the myths-only condition was even significantly lower than control. As participants were not presented with any affirmations, they seemingly assumed that any information presented regarding the relevant topic was false.
Table 5

F and p values of fact condition belief ratings planned comparisons

<table>
<thead>
<tr>
<th></th>
<th>Standard</th>
<th>Reverse</th>
<th>Facts-only</th>
<th>Myths-only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reverse</td>
<td>$F = .79$</td>
<td>$p = .38$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facts-only</td>
<td>$F = .01$</td>
<td>$p = .92$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myths-only</td>
<td>$F = 288.98$</td>
<td>$p &lt; .001^*$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>$F = 66.76$</td>
<td>$p &lt; .001^*$</td>
<td>$F = 70.93$</td>
<td>$p &lt; .001^*$</td>
</tr>
</tbody>
</table>

Note. All df1 = 1, df2 = 85. * indicates significance after applying the Holm-Bonferroni adjustment for multiple (n = 7) comparisons.

Inference scores. Mean inference scores can be seen in Figure 7. First, a 2 × 5 within-between ANOVA with factors retention interval and type of explanation was performed on participants’ mean myth inference scores. There was a marginal main effect of retention interval, $F(1,85) = 3.40; p = .069; MSE = 6.53; \eta_p^2 = .04$, indicating that belief increased slightly over time, and a main effect of explanation type, $F(4,340) = 13.73; p < .001; MSE = 2.94; \eta_p^2 = .14$.

Figure 7. Inference scores in Experiment 3 over time by explanation formats. Dotted lines illustrate control conditions.
In an equivalent manner to the belief ratings, planned comparisons were performed on the myth inference scores. The comparisons, collapsed over retention interval, are shown in Table 6. After controlling for multiple comparisons with a Holm-Bonferroni adjustment, we confirmed that all correction formats differed from control, yet all alternative correction formats were equivalent to the standard format.

Table 6

<table>
<thead>
<tr>
<th></th>
<th>Standard</th>
<th>Reverse</th>
<th>Facts-only</th>
<th>Myths-only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reverse</td>
<td>$F = 2.02$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$p = .16$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facts-only</td>
<td>$F = .27$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$p = .60$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myths-only</td>
<td>$F = .95$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$p = .33$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>$F = 32.56$</td>
<td>$F = 12.74$;</td>
<td>$F = 43.39$</td>
<td>$F = 44.35$</td>
</tr>
<tr>
<td></td>
<td>$p &lt; .001^*$</td>
<td>$p &lt; .001^*$</td>
<td>$p &lt; .001^*$</td>
<td>$p &lt; .001^*$</td>
</tr>
</tbody>
</table>

Note. All df1 = 1, df2 = 85. * indicates significance after applying the Holm-Bonferroni adjustment for multiple ($n = 7$) comparisons.

Next, a $2 \times 5$ within-between ANOVA was performed on the fact inference scores. There was a main effect of retention interval, $F(1,85) = 4.64; p = .034$; $MSE = 4.06; \eta_p^2 = .05$, indicating that fact belief decreased between one and three weeks, and a main effect of type of explanation, $F(4,340) = 17.01; p < .001$; $MSE = 4.13; \eta_p^2 = .17$.

Planned comparisons of the fact inference scores collapsed over retention interval can be seen in Table 7. A set of Holm-Bonferroni contrasts revealed that all conditions differed from control, apart from the myths-only format (where no factual affirmations were presented). Additionally, the reverse and facts-only format both seemed less effective than the standard format. That is, people were slightly better at
using the factual information in their reasoning when the affirmation was presented in the standard format, in comparison to the facts-only and reverse formats.

Table 7

*F and p values of fact condition inference score planned comparisons*

<table>
<thead>
<tr>
<th>Condition</th>
<th>Standard</th>
<th>Reverse</th>
<th>Facts-only</th>
<th>Myths-only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reverse</td>
<td>$F = 6.88$</td>
<td></td>
<td>$F = 5.30$</td>
<td>$F = 45.69$</td>
</tr>
<tr>
<td></td>
<td>$p = .010^*$</td>
<td></td>
<td>$p = .024^*$</td>
<td>$p &lt; .001^*$</td>
</tr>
<tr>
<td>Facts-only</td>
<td>$F = 5.30$</td>
<td></td>
<td>$F = 10.69^*$</td>
<td>$F = 11.45$</td>
</tr>
<tr>
<td></td>
<td>$p = .024^*$</td>
<td></td>
<td>$p &lt; .001^*$</td>
<td>$p = .001^*$</td>
</tr>
<tr>
<td>Myths-only</td>
<td>$F = 45.69$</td>
<td></td>
<td>$F = 11.45$</td>
<td>$F = 2.93$</td>
</tr>
<tr>
<td></td>
<td>$p &lt; .001^*$</td>
<td></td>
<td>$p = .001^*$</td>
<td>$p = .09$</td>
</tr>
</tbody>
</table>

*Note.* All df1 = 1, df2 = 85. * indicates significance after applying the Holm-Bonferroni adjustment for multiple $(n = 7)$ comparisons.

4.6 Discussion

This paper aimed to investigate the relative efficacy of various explanation formats to promote belief change over time, with a particular focus on reducing myth beliefs. The results indicated that belief change was largely independent of the specific format of explanation used. While this study did not succeed in finding an “optimal” correction method, the lack of difference between explanation conditions remains informative, and provides indications for theoretical frameworks regarding why belief in misinformation persists.

4.6.1 Avoiding the Misconception may not be Necessary

Had avoiding the original misconception within the retraction helped individuals update their myth beliefs, this would have supported the notion that false acceptance of myths is driven by the correction boosting the myth’s familiarity
(Skurnik et al., 2007). However, the facts-only condition, where myths were reframed as facts, was found to be equally as effective as the standard format that explicitly mentioned the misinformation. This suggests that either the misconception’s familiarity is activated regardless of the frame, or alternatively that familiarity is not as detrimental as once thought, particularly if paired with a valid correction. Thus, although familiarity contributes to the continued influence effect (Swire et al., 2017; i.e. Chapter 3), there is no particular need to avoid the misconception when retracting it. This is particularly apt advice when considering the inference-score data from the present experiment showed that the facts-only condition was less effective than the standard format.

The current findings are inconsistent with the proposal by Skurnik et al. (2007) and the associated notion of a familiarity backfire effect (Cook & Lewandowsky, 2011; Lewandowsky et al., 2012). However, the data are in line with a recent report by Ecker, Hogan, and Lewandowsky (2017), which suggested that the repetition of a misconception is not detrimental to a retraction’s corrective impact (also see Ecker et al., 2011; Wilkes & Leatherbarrow, 1988). Even if a correction may make a myth more familiar, it seems that this potential negative effect can be offset by the enhanced salience that repetition of the myth within the correction may provide (Ecker et al., 2017). The current findings are also consistent with a study by Rich and Zaragoza (2016), that found corrections were more effective if the misinformation was explicit rather than only implied prior to correction. This current study thus contributes to a growing body of literature that suggest presenting the misconception alongside a correction is not harmful to belief updating. This supports the notion that co-activation of the myth and the corrective information may be required for belief revision to occur (Guzzetti et al., 1993; Kendeou et al., 2014, 2017).
4.6.2 Presenting Both Myths and Facts may be Beneficial

Presenting participants with only the myths did not reduce belief more than the standard format, and additionally had the unfortunate side effect of leading participants to reject some true items as false. This suggests that if people only hear what is not true about a topic, they may become more sceptical about other claims relating to that topic even when those are true. To avoid this cynicism, it is thus worth presenting a balance of information as is intrinsically executed in the standard myths vs. facts format. While presenting “balanced” arguments may not always be appropriate and can at times be misleading (e.g., false-balance media coverage; see Cook, Lewandowsky, & Ecker, 2017; Dixon & Clarke, 2013), truthfully explaining both the facts and the fiction in an educational setting might well give people a more nuanced view of a subject.

4.6.3 Reverse Order is Equivalent to the Standard Format

We presumed that clarifying the fact/myth status before presenting the claim may be a more effective strategy than the standard format. However, contrary to these expectations, there was no order effect found for myth belief or inference measures. The reverse-order format performed equivalently to the standard format, suggesting the order is not as consequential as assumed. There was some evidence to suggest that the standard format might be better than the reverse-order format at facilitating updating for factual inferences. This could reflect a query order effect—for factual reasoning it may be beneficial to clarify the statement prior to affirming that it is true, perhaps in order to facilitate understanding of the explanation. However, the effect was small and this explanation remains speculative.

In conclusion, the current findings largely suggest that the explanation format is not a strong determinant of belief change. However, this should be interpreted with caution due to the experiment’s limitations. The principal limitation is that there was
a randomisation error resulting in one topic to be absent from the standard format condition. It is thus necessary to replicate this experiment prior to making generalisable recommendations for communication.
4.7 References


Chapter 5:  Processing Political Misinformation—Comprehending the Trump Phenomenon
5.1 Foreword

Where Chapter 3 discussed memory processes underlying the continued influence effect, and Chapter 4 endeavoured to find an ideal correction method, Chapter 5 advances to ideological biases. The experiments reported in Chapter 5 utilised political stimuli to investigate the cognitive processing of true and false information when the information stems from a polarising source. Where previous misinformation literature regarding source credibility has often utilised fictitious scenarios and sources (e.g., misconduct of fictional politicians, fictional experts; Guillory & Geraci, 2013; Ecker & Antonio, 2016), the following studies used real-world information and sources. By holding the content of the information constant, and only manipulating who stated the information, we were able to obtain a clean measure of the source’s influence on information processing.

The US political landscape in 2015 and 2016 lent itself to our intended field of study. We thus capitalised upon the extreme polarisation of the (then) presidential candidate Donald Trump. Chapter 5 investigates the perceived disconnect between Donald Trump’s high polling and the amount of misinformation that he was eliciting. Donald Trump’s supporters seemingly either believed that his assertions were accurate or they were aware that he was dispersing misinformation, but it did not factor into their voting preferences.

Chapter 5 thus addresses three open questions: (1) Do people update their misinformed beliefs post-correction even if they support the source of the misinformation and are thus motivated to reject the correction? (2) How does the trustworthiness of the correction source affect the efficacy of a correction under these conditions? (3) If people do successfully update their beliefs, does that impact on their voting preferences? We thus measured intended voting intentions alongside
belief updating of numerous pieces of misinformation, over a one week period of time.
5.2 Abstract

This study investigated the cognitive processing of true and false political information. Specifically, it examined the impact of source credibility on the assessment of veracity when information comes from a polarising source (Experiment 4), and effectiveness of explanations when they come from one's own political party or an opposition party (Experiment 5). These experiments were conducted prior to the 2016 Presidential election. Participants rated their belief in factual and incorrect statements that President Trump made on the campaign trail; facts were subsequently affirmed and misinformation retracted. Participants then re-rated their belief immediately or after a delay. Experiment 4 found that (i) if information was attributed to Trump, Republican supporters of Trump believed it more than if it was presented without attribution, whereas the opposite was true for Democrats and (ii) although Trump supporters reduced their belief in misinformation items following a correction, they did not change their voting preferences. Experiment 5 revealed that the explanation's source had relatively little impact, and belief updating was more influenced by perceived credibility of the individual initially purporting the information. These findings suggest that people use political figures as a heuristic to guide evaluation of what is true or false, yet do not necessarily insist on veracity as a prerequisite for supporting political candidates.
5.3 Introduction

Individuals from opposing sides of the political spectrum often disagree over what is fact and what is fiction. While both conservatives and liberals aim to be well informed, even empirical information that seems straightforward can lead to discord (Hochschild & Einstein, 2015). For example, people perceive unemployment, inflation and crime rates to be lower when their preferred party is in power (Bartels, 2002). Partisanship clearly influences the way people process information, but the exact cognitive mechanisms that underlie these differences are still being debated (Berinsky, 2015; Kahan, 2013; Lewandowsky, Stritzke, Freund, Oberauer, & Krueger, 2013). In this study, we focus on source credibility. Individuals have limited time and cognitive resources to comprehend complex topics such as policy or current affairs, and may therefore use the perceived credibility of political figures as a heuristic to guide their evaluation of what is true or false. For instance, Republicans and Democrats are likely to assess the veracity of a statement differently depending on whether it comes from a favoured politician (Housholder & LaMarre, 2014).

To study how individuals evaluate whether political information is true or false, we first examined the impact of source credibility on the initial assessment of information veracity. To this end, we used statements from perhaps the most polarising political figure of recent times, President Trump. As these experiments were conducted prior to his election and inauguration, we henceforth refer to the him as ‘Donald Trump’, or ‘Trump’. Second, we investigated the impact of source credibility on the corrective effect of retracting misinformation and affirming factual statements.
5.3.1 The Continued Influence Effect

False information continues to influence memory and reasoning even after credible corrections; this has been termed the continued influence effect of misinformation (Ecker, Lewandowsky, Swire, & Chang, 2011; Johnson & Seifert, 1994; Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012). Once information is assumed to be true, this conviction is subsequently difficult to change. The continued influence effect occurs even with non-politicised misinformation and at least in part reflects the inherent difficulty of updating one's mental model of an event or a causality (Ecker, Lewandowsky, & Cheung, 2015; Ecker, Lewandowsky & Tang, 2010). However, ongoing reliance on corrected misinformation becomes an even greater problem when the misinformation conforms to a person's pre-existing belief and supports their ideological worldviews, whereas the correction runs counter (Bolsen, Druckman, & Cook, 2014). Once an individual feels personally connected to information, their ideology and values influence how that information is processed (Kunda, 1990; Wells, Reedy, Gastil, & Lee, 2009); this is known as motivated reasoning or motivated cognition.

5.3.2 Motivated Cognition

There is an extensive literature on motivated cognition that suggests individuals are more critical when evaluating information that is counter to their beliefs than belief-congruent information, and conclusions that people reach are likely to be consistent with their prior assumptions about how the world functions (Gains, Kuklinsky, Quirk, Peyton, & Verkuilen, 2007; McGuire, 1972; Taber & Lodge, 2006). For example, a classic study by Lord, Ross, and Lepper (1979) found that both supporters and opposers of capital punishment rated studies regarding the death penalty as more convincing when the studies confirmed their existing views. In addition, after receiving mixed evidence comprising both supportive and critical
findings, participants' attitudes further diverged—those who initially opposed the death penalty reported opposing it even more, and the reverse occurred for those in support of the death penalty. This illustrates how an individual's worldview can dictate how new information is assessed, legitimising the preservation of the person's ideological belief system (Jost & Krochik, 2014).

In the real world, information sometimes turns out to be incorrect and therefore may be subject to revision. Once people have decided that they believe some particular information to be true, they may encounter a correction that challenges their conviction. The extent to which people take heed and change their beliefs based on such corrections may depend on motivated cognition. Specifically, if a correction runs counter to a person's beliefs and worldview, they may be more likely to ignore it, and cling to the original misinformation. For example, when incorrect information arising from a Democratic politician's statement is retracted, Democrats—and particularly those who support the politician—may resist the correction more than their Republican counterparts who have a vested interest in the political figure being incorrect. At worst, a potential outcome of the attempt to correct contentious misinformation is a worldview backfire effect. This occurs when an individual feels motivated to defend their belief system, and ironically reports a stronger belief in the original misconception after receiving a retraction. For example, worldview backfire effects have been documented with attempts to promote vaccine safety (Nyhan & Reifler, 2015), as well as attempts to correct misconceptions regarding anthropogenic climate change or the existence of weapons of mass destruction (WMDs) in Iraq immediately prior to the invasion of 2003 (Hart & Nisbet, 2012).

This phenomenon might be especially pronounced among certain individuals. A recent debate in the literature is concerned with the question of whether
conservatives are generally more prone to motivated cognition and worldview backfire effects. One school of thought assumes that personality characteristics associated with conservative ideology present a specific susceptibility for motivated cognition. For example, Jost, Glasser, Kruglansky and Sulloway (2003) suggested that psychological variables such as dogmatism (that is, intolerance of ambiguity, avoidance of complexity and a need for closure) are predictive of conservatism and increase the likelihood that an individual engages in ‘black-or-white’ assessments of information. This tendency to readily decide on information veracity with subsequent resistance to change could lead to greater rejection of factual information for those on the political right relative to moderate and liberal segments of the population (Nam, Jost, & Bravel, 2013).

By contrast, Kahan (2013) posits that identity-protective motivated cognition occurs equally at both ends of the political spectrum, arguing that conservatives and liberals perform comparably on a measure of information-processing dispositions associated with cognitive biases. Individuals who scored higher on ‘cognitive reflection’—a disposition to engage in effortful processing (Frederick, 2005)—were more likely to demonstrate motivated cognition, regardless of partisanship. While the rejection of scientific evidence seems to be primarily associated with conservative ideology (Mooney, 2012), the observed asymmetry may not reflect fundamental differences in cognition; rather, it may just be the case that the contested scientific findings happen to challenge primarily the worldview of conservatives rather than liberals (Lewandowsky & Oberauer, 2016). In support of this, Nisbet, Cooper, & Garrett (2015) found that liberal participants react in a manner equivalent to conservatives if they encounter liberal-dissonant science messages, for example regarding the efficacy of nuclear power.
In contrast to these backfire effects, Kahan (2015) reported no partisan difference for scientific rejection among issues that do not challenge worldviews, such as cell-phone radiation or exposure to high-voltage powerlines. Additionally, Kuklinski, Quirk, Jerit, Schwieder and Rich (2000) found that while strong partisans held the least accurate beliefs regarding welfare policy (e.g. the proportion of the federal budget that welfare absorbs), and the highest confidence that these beliefs were accurate, they were not more inclined to reject factual information once corrections were presented. It is therefore possible that party-line differences in the willingness to engage in belief revision are not as pervasive as some research has suggested; there is some evidence that if strong partisans receive quality information, they may be able to interpret it in a similar fashion and update their beliefs to the same extent (Bullock, 2009; see also Ecker, Lewandowsky, Fenton, & Martin, 2014).

5.3.3 Source Credibility

In addition to motivated reasoning, when people are evaluating whether information is fact or fiction, the source of the information matters a great deal. In general, high-credibility sources are more persuasive and promote greater attitude change than low credibility sources (Eagly & Chaiken, 1993). Additionally, given that attitude homophily—i.e. the extent to which a person perceives similarities between the way they think and another person does—is a key determinant of perceived source credibility, candidate support has substantial impact when estimating the credibility of preferred versus non-preferred political candidates (Houshholder & LaMarre, 2014). Two key components of source credibility are (i) expertise—the extent to which the source is able to give accurate information—and (ii) trustworthiness—the extent to which the source is willing to provide information that the source itself assumes to be correct (Pornpitakpan, 2004).
When it comes to the efficacy of correcting inaccurate information, it appears that the latter is more important than the former—it is more important that the source of the correction is perceived to be trustworthy than having expertise (Ecker & Antonio 2016; Guillery & Geraci, 2013; McGinnes & Ward, 1980). This finding suggests that the most effective way to reduce misconceptions is to attribute the correction to a source that the person finds a trustworthy source of information, such as a member of the political party the individual identifies with. On the other hand, there is contrasting evidence suggesting that an unlikely source—for example, a Republican correcting another Republican—could be more effective at reducing misconceptions than a source that is expected to provide the corrective information. Thus, a Democrat's belief in misinformation originating from a Republican source may be more strongly reduced by a correction that also comes from a Republican source, rather than a Democrat source (Berinsky, 2015).

Even if people are able to change their beliefs immediately after a correction, belief change may be fleeting (Swire, Ecker, & Lewandowsky, 2017). In this case, worldview and an individual's trust in the veracity of the source may influence the rate of forgetting, and could thus lead to ‘motivated forgetting’ (Anderson & Hanslmayr, 2014). For example, if misinformation arising from a Democratic politician's statement is retracted, Democrats who support the politician may initially update their belief, but conveniently forget the correction at an accelerated pace over time, thus eventually reverting to their pre-existing beliefs.

Finally, even if it is possible to correct people's misconceptions, it is unclear whether or not such corrections affect candidate support. If an individual acknowledges that a number of a politician's statements are untrue, they should reduce their support to the extent that truthfulness is a desirable trait of a political figure. However, Redlawsk (2002) found that participants increased their support for
candidates whom they endorsed when provided with negative information about the candidate. Likewise, Meffert, Chung, Joiner, Waks, and Garst (2006) found that participants spent more time reading negative stories about candidates they preferred, yet this led to a more positive outlook of the candidate. This shows that candidate support ratings are also subject to worldview backfire effects, and it is therefore possible that highlighting misinformation that candidates have disseminated may not result in any loss in support, and could ironically lead to increased support.

5.3.4 The Case of Donald Trump

It is clear that individuals view the world through a partisan filter; however, the extent to which citizens use partisan cues such as political figures to evaluate the veracity of information and corrections requires further exploration. Donald Trump is an interesting case study for misinformation research, as bipartisan fact-checking media outlets have found that Donald Trump has been particularly prone to inaccuracies (‘Post truth”, 2016; Lippman, Samuelsohn, & Arndorf, 2016), and for much of the presidential campaign was a divisive figure even among Republicans (Grenier, 2016).

While voters are well aware that they encounter politically motivated misinformation during election campaigns, they find it difficult to pinpoint the accuracy of specific messages and are therefore misinformed on a wide array of prominent issues (Ramsay, Kull, Lewis, & Subias, 2010). Donald Trump's popularity, despite the amount of misinformation he distributed, can be explained by either the notion that (i) people believe that his assertions are true (partially because they see Donald Trump as a trustworthy source of information) and they avoid or resist the many corrections available in the public sphere (partially based on motivated cognition), or alternatively (ii) the public is aware that Donald Trump is spreading misinformation, but does not insist on veracity as a prerequisite for their
support of a candidate. In this study, we explored these possibilities through several means. First, we tested whether the public believes misinformation spread by a polarising source, and whether such information can be effectively corrected. We also explored whether a change in belief leads to a shift in voting preferences (i.e. after a credible correction, did people reduce their belief in misinformation yet continued to support Donald Trump?).

Specifically, Experiment 4 investigated (i) whether belief in both misinformation and factual information differs depending on whether or not the information is associated with a polarising source (i.e. Donald Trump); (ii) whether the impact of corrections/affirmations differs when support for the polarising source of the original information is taken into account; and (iii) whether belief change is sustained over time. Experiment 5 tested whether the impact of corrective/affirmative explanations is moderated by partisanship (i.e. stating that a correction/affirmation stems from a Democratic, Republican or non-partisan source).

5.4 Experiment 4

Experiment 4 was conducted in November 2015 prior to the Iowa caucus, when 13 other candidates apart from Donald Trump were still viable options (these candidates were Jeb Bush, Ben Carson, Chris Christie, Ted Cruz, Carly Fiorina, Jim Gilmore, Lindsay Graham, Mike Huckabee, John Kasich, George Pataki, Rand Paul, Marco Rubio and Rick Santorum). The experiment featured actual statements made by Donald Trump on the campaign trail in 2015. Some of these statements were inaccurate and others were factual. When these statements were presented to participants, they were either explicitly attributed to Trump or presented without attribution. The objectively false statements were then corrected, and the true statements were affirmed, with a brief explanation. Participants rated their belief in
the statements both before and after the corrective/affirmative explanation; the second rating was either immediate or following a one-week delay.

To tease apart partisanship from candidate advocacy, we separated Republican participants into those who supported Trump and those who did not. This step is somewhat rare in studies of political cognition, but given the polarising nature of Trump's candidacy within the Republican party at the time of the study, we felt it was inappropriate to mix these two groups. The study thus used a 2 × 2 × 2 × 3 design—type of item (misinformation versus fact) was a within-subjects factor, and the between-subjects factors were the source of information (Trump versus unattributed), study-test retention interval (immediate versus delayed) and Trump support (Democrat versus Republican non-supporters versus Republican supporters). See Figure 8 for a schematic diagram of the experimental design. Our prime dependent variable was participants' belief in the inaccurate and factual statements measured on an 11-point scale, as well as participants' self-reported support for Donald Trump.

We hypothesised that participants would use Donald Trump as a cue to evaluate information veracity: we expected that Republican Trump supporters would increase belief in both misinformation and factual statements if they were attributed to Donald Trump, and Democrats and Republican non-supporters would decrease their belief. We also hypothesised that explanations would have a limited effect and would be less sustained over time when they ran counter to participants' expectations arising from their affiliation (i.e. when Republican supporters encountered corrections of Trump's misinformation or Democrats and Republican non-supporters encountered affirmations of Trump's true statements). Lastly, we hypothesised that voting preferences would increase or not change, even if participants reduced belief in misinformation (or increased belief in facts) attributed to Trump.
5.4.1 Method

**Participants.** Participants were 2023 US residents recruited through Amazon.com's Mechanical Turk. Republican participants who had recently taken part in previous studies from the Massachusetts Institute of Technology's Political Experiments Research Laboratory were invited to participate. We adopted this oversampling strategy due to the relative scarcity of Republicans within the Mechanical Turk population. Participants were paid 85 cents and an additional 50 cents in the one-week delayed condition. They were excluded from the analysis if they did not complete all parts of the study \(n = 247\)^10 The final sample included \(N = 1776\) participants, with 884 males and 892 females in the age range of 19–78 years, with a mean age of \(M = 35.73\) (s.d. = 11.41).

---

10 Of the excluded participants, 94% were in the one-week retention interval. A Pearson \(\chi^2\)-test indicated that neither Trump support, \(\chi^2 = (3, N = 2023) = 1.92, p = .589\), nor source, \(\chi^2 = (1, N = 2023) = .28, p = .592\), differed between participants who were included and those who were excluded.
Stimuli. Four inaccurate statements and four factual statements made by Donald Trump on the campaign trail prior to 1 October 2015 were compiled by the authors. The Trump condition explicitly stated that Donald Trump was the propagator of the information, while the unattributed condition presented the information without specifying its source. Corrections and affirmations of equal length (i.e. two to three sentences) were created; each explanation explicitly referenced a reputable source. Sources were chosen to be non-partisan (e.g. the ‘Danish Epidemiology Science Centre’ or the ‘US Bureau of Labor Statistics’). An example misinformation item with its corresponding correction can be found in Table 8 (see Appendix C for the complete list of items). Explanations consisted of four segments: (i) the participant was reminded of the initial item; (ii) the veracity was presented; (iii) information was given as to why the statement was true or false and (iv) the participant was given a reminder of their initial belief rating.

Procedure. After reviewing a University of Western Australia and Massachusetts Institute of Technology approved consent form, participants took the survey through Qualtrics.com. They were first presented with general demographic and political-ideology questions. Participants who did not identify with a party, but indicated that they leaned towards a particular party were classified as partisans (Keith et al., 1992). This was followed by questions regarding the likelihood of voting for candidates in the 2016 Presidential campaign.
Table 8

*Examples of Trump and unattributed misinformation with corresponding corrections.*

<table>
<thead>
<tr>
<th>Misinformation</th>
<th>Correction</th>
<th>This is false.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Trump</strong></td>
<td>Donald Trump said that vaccines cause autism.</td>
<td>There is strong consensus in the scientific community that vaccines are not linked to autism. For example, one study by the Danish Epidemiology Science Centre tracked all children born in Denmark from 1991 to 1998 and concluded that there was no increase in the rate of autism for vaccinated as opposed to non-vaccinated children.</td>
</tr>
<tr>
<td>On a scale between 0-10, do you believe Trump’s statement to be true?</td>
<td>This statement x out of 10 (0 = definitely false, 10 = definitely true).</td>
<td>You previously rated this statement x out of 10 (0 = definitely false, 10 = definitely true).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Unattributed</strong></th>
<th>Vaccines cause autism.</th>
<th><strong>This is false.</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>On a scale between 0-10, do you believe this statement to be true?</td>
<td>This statement x out of 10 (0 = definitely false, 10 = definitely true).</td>
<td>You previously rated this statement x out of 10 (0 = definitely false, 10 = definitely true).</td>
</tr>
</tbody>
</table>

The candidates were Donald Trump, Ben Carson, Hillary Clinton and Bernie Sanders, who were the front-runners at the time the survey was conducted.

Participants' feelings towards the candidates were also measured using the ‘candidate-feelings thermometers’ employed in the American National Elections Study. These entail asking participants to rate how favourably and warm they feel
towards the person; ratings between 0 and 50 degrees are taken to indicate they do not feel particularly warm, and ratings between 50 and 100 are taken to indicate they do feel favourably and warm towards the candidate.

Participants were presented with the eight statements in a randomised order; participants received either all statements attributed to Donald Trump or alternatively presented without source specification. After rating the extent to which they believed each item on a 0–10 scale, participants received an explanation for each item as to whether it was true or false.11 Participants then moved directly to the test phase if they were in the immediate post-test group. This involved re-rating belief in all eight statements in random order, as well as re-rating candidate support and feelings towards the candidates. In the delayed post-test condition, participants were instead re-contacted after one week and given the opportunity to complete the test phase.

5.4.2 Results

Of the 1776 participants, 1015 identified as Democrats and 535 identified as Republicans. The 226 participants who had no political affiliation were omitted from the following analyses. Of the Republicans, 323 were classified as Trump supporters as they scored 5 or more (out of 10) on the likelihood to vote for Trump measure, and the 212 participants who scored less than 5 were classified Trump non-supporters. There were 99 Democrats who supported Trump—all main effects and interactions of the following analyses were replicated if these participants were omitted from the analyses.

11 As all items were presented within-subjects, it could be a concern that participants receiving multiple pieces of corrective information are more vulnerable to social desirability biases. However, a one-way ANOVA on the pre-explanation belief ratings confirmed that their presentation order did not have a significant influence on belief, \( F(7,12425) = 1.38; \ p = .210 \). Post-explanation belief was likewise not affected by presentation order, \( F(7,12425) = 1.61; \ p = .127 \).
First, Trump support groups were compared on demographic measures. A one-way ANOVA indicated that age was different between groups, $F(2,1547) = 26.03; p < .001; MSE = 128; \eta^2_p = .03$. Democrats are younger than both Republican groups, $F(1,1547) = 46.82; p < .001; MSE = 128$. Next, a one-way ANOVA indicated that education was different between groups, $F(2,1547) = 12.29; p < .001; MSE = 1.48; \eta^2_p = .01$. Planned comparisons revealed that Republican non-supporters were significantly more educated than Democrats, $F(1,1547) = 4.51; p = .034; MSE = 1.48$, yet Democrats were significantly more educated than Trump supporters, $F(1,1547) = 8.82; p = .003; MSE = 1.48$. Finally, a Pearson chi-squared test revealed there were no gender differences between groups, $X^2 = (3, N = 1776) = 2.24, p = .489$. The following analyses remained statistically significant when controlling for education and age using factorial ANCOVAs (unless indicated otherwise).

**Pre-explanation belief scores.** Pre-explanation belief scores partitioned by Trump support are shown in Figure 9. The left-hand side of the figure shows the misinformation and the right-hand side shows the facts. We further split the sample into those respondents who received statements without source attribution and those who received statements attributed to Trump. For both misinformation and factual statements, Trump attribution was associated with lower belief in the statements among Democrats and greater belief among Republican supporters of Trump. Among Republican non-supporters, a Trump attribution did not affect belief in the misinformation, but did reduce belief in factual statements.
Figure 9. Pre-explanation Democratic and Republican belief in statements associated with Trump or presented unattributed. Error bars denote 95% confidence intervals.

A 2 × 3 factorial ANOVA was performed on the misinformation pre-explanation belief scores. The analysis revealed two significant main effects. The main effect of type of source (unattributed vs. Trump), $F(1,1544) = 6.12; p = .013; MSE = 2.60; \eta^2_p = .004$, indicated that Trump attribution influenced belief. The main effect of Trump support (Democrats vs. Republican non-supporters vs. Republican supporters), $F(2,1544) = 116.94; p < .001; MSE = 2.60; \eta^2_p = .13$, indicated that beliefs of the three groups differed. These main effects were qualified by an interaction between source and Trump support, $F(2,1544) = 28.84; p < .001; MSE = 2.60; \eta^2_p = .04$, reflecting that Trump attribution led to decreased belief for Democrats but increased belief for Trump supporters. Additionally, a planned comparison confirmed that for Republican non-supporters, misinformation belief was not affected by Trump attribution, $p = .575$.

Next, we performed a 2 × 3 factorial ANOVA on the pre-explanation belief scores for the factual statements. The analysis revealed main effects of both type of source, $F(1,1544) = 15.96; p < .001; MSE = 2.25; \eta^2_p = .01$, and Trump support, $F(2,1544) = 34.50; p < .001; MSE = 2.25; \eta^2_p = .04$, as well as an interaction of source and Trump support, $F(2,1544) = 25.50; p < .001; MSE = 2.25; \eta^2_p = .03$. An
interaction contrast confirmed that for factual statements, Republican non-supporters believed in the facts less when the information was associated with Trump rather than unattributed, whereas the Republican supporters expressed greater belief in statements made by Trump, $F(1,1544) = 8.03; p = .005; MSE = 2.25$. A planned comparison revealed that Democrats believed the statements less if attributed to Trump, $F(1,1544) = 119.61; p < .001; MSE = 2.25$. Thus, Trump support influenced the perceived truth of the information.

**Post-explanation belief scores.** The general trend and the full trajectory of belief change over time is shown in Figure 10. The left side of the figure shows the unattributed condition, and the right side shows the Trump attributed condition. Immediately after the corrections/affirmations, both Democrats and Republicans showed a substantial amount of belief change, which generally diminished over the course of one week for both misinformation and facts. We found no evidence for backfire effects, as post-explanation belief scores in misinformation remained below pre-explanation levels.

![Figure 10. Belief in Trump and unattributed misinformation and facts over time, across Trump support groups and source conditions. Rep = Republican, Misinfo = Misinformation. Dotted lines show misinformation items. Error bars denote 95% confidence intervals.](image)

To simplify the data, we computed total accuracy scores by subtracting participants’ misinformation scores from their fact scores. On this measure, the
higher the score, the more likely participants were to accurately assume misinformation to be false and factual information to be true. These accuracy scores across conditions are shown in Figure 11. A $2 \times 2 \times 3$ factorial ANOVA involving the source, retention interval, and Trump support factors was performed on the post-correction accuracy scores. The analysis revealed three significant main effects. The main effect of source, $F(1,1538) = 15.42; p < .001; MSE = 6.93; \eta^2_p = .01$, indicated that Trump attribution was associated with less accurate post-correction beliefs. The main effect of retention interval, $F(1,1538) = 183.44; p < .001; MSE = 6.93; \eta^2_p = .11$, indicated that belief accuracy dropped over the course of a week, and the main effect of Trump support, $F(2,1538) = 9.34; p < .001; MSE = 6.93; \eta^2_p = .01$, indicated that belief accuracy differed by Trump support, with Republican Trump supporters showing the lowest scores overall.

These main effects were qualified by a significant interaction of source and retention interval, $F(1,1538) = 3.94; p = .047; MSE = 6.93; \eta^2_p = .003$, indicating that

Figure 11. Accuracy scores—misinformation scores subtracted from fact scores—across Trump support and source. Error bars denote 95% confidence intervals.
the influence of Trump attribution changed over time. From Figure 11, we can see that in the immediate condition, Trump attribution does not have a strong influence; over the course of a week, however, participants from all groups seemed to forget the corrective/affirmative explanations at an accelerated rate when the original information was associated with Donald Trump. This was confirmed with a significant planned comparison focusing on the one-week delayed condition collapsed over Trump support, contrasting the Trump against unattributed conditions, \( F(1,1538) = 15.13; p < .001; MSE = 6.93 \). In other words, if the original information came from Donald Trump, after a one-week delay participants had less accurate beliefs, regardless of their affiliation or initial support for Trump.

If the post-explanation misinformation and items are analysed separately, we see similar trends (the full analyses can be found in Appendix D). The most prominent differences to the above accuracy score analyses are that (1) misinformation items do not show an interaction of source and retention interval, indicating that unlike the fact scores (where Trump attribution led to less accurate beliefs particularly over time), Trump attribution led to less accurate belief over both time periods, and (2) fact items additionally show an interaction of Trump support and retention interval, \( F(2,1538) = 3.44; p = .032; MSE = 2.38; \eta_p^2 = .004 \). While Democrats over both time periods are worse at updating their belief in the facts if information is attributed to Trump, Republicans immediately update their belief equally in the Trump and unattributed conditions, yet after one week belief in the Trump information reduces below the unattributed condition, \( F(1,1538) = 5.08; p = .0243; MSE = 2.38 \).

---

12 This interaction was only marginally significant when controlling for age, \( p = .060 \).
To illustrate why accuracy is reduced after one week due to Trump attribution, Figure 12 shows the Trump condition subtracted from the unattributed condition—observations above zero indicate that the attribution of a statement to Trump encourages participants to believe the information; values below zero indicate that the attribution of statements to Trump made participants less likely to believe in the information. The left side of the figure shows the misinformation, and the right shows the facts. The distance from zero indicates the impact that the Trump attribution is having upon the belief scores.

Figure 12 highlights the fact that initially, before they receive the correction, participants use their support for Donald Trump as a heuristic for whether information is true or false (i.e., independent of actual veracity, Republican supporters believe Trump information more, Democrats believe Trump information less, and Republican non-supporters are not affected much). However, after one week—regardless of partisanship and level of Trump support—people tend to assume Trump’s facts are incorrect, and Trump’s misinformation is accurate.
Likelihood to vote and feelings thermometer ratings. Attributing the information to Trump did not influence participants’ intention to vote nor their feelings towards the political figure. Figure 13 shows the full trajectory of participants’ likelihood to vote for Donald Trump, both prior to and after the corrective/affirmative explanations. To simplify the analysis, the post-explanation scores were subtracted from the pre-explanation scores to create change indices for both the likelihood-to-vote and feelings thermometer ratings.

A 2 × 2 × 3 factorial ANOVA on the likelihood-to-vote change index revealed two main effects. The main effect of Trump support, $F(2,1537) = 13.39; p < .001; MSE = 1.35; \eta_p^2 = .02$, indicated that Republicans were more likely to change their voting preferences than were Democrats. For example, Republican non-supporters increased their support for Trump by 0.22 on the 10-point scale in the immediate condition and by 0.35 after one week, yet Democrats only increased their support by 0.07 in the immediate condition, and decreased their support by 0.01 after one week. The main effect of retention interval, $F(1,1538) = 8.00; p = .005; MSE = 1.35; \eta_p^2 = .005$, indicated that change was greater after one week than immediately after the explanations.

These main effects were qualified by an interaction of retention interval and Trump support, $F(2,1537) = 9.06 p < .001; MSE = 1.35; \eta_p^2 = .01$, indicating that change in voting preferences differed between Trump support groups over time. Republican supporters slightly reduced their likelihood of voting for Trump and Republican non-supporters slightly increased their likelihood. As there was no main effect or interaction of source, it can be assumed that these differences can be attributed to natural fluctuation of voting preferences over time rather than the explicit association of information to Donald Trump. The likelihood-to-vote trend
was mimicked by the feelings-thermometer ratings (see Appendix E for the figure and analysis).

![Figure 13. Likelihood-to-vote ratings over time between Trump support and source. Error bars denote 95% confidence intervals. Rep = Republican, Misinfo = Misinformation.](image)

Finally, 48 pairwise correlations were calculated for Democrats, Republican supporters, and Republican non-supporters to investigate whether belief change in misinformation or factual statements was associated with (1) a change in likelihood to vote or (2) feelings towards Trump over time for each retention interval and source. Using a Bonferroni adjusted alpha level of .001, two correlations were significant, revealing that the more Democrats reduced their belief in Trump-attributed misinformation, the more they reduced their feelings and likelihood of voting for Trump one week post-explanation ($r = .36$ and $r = .33$, respectively). This could reflect the fact that Democrats who reduce misconceptions attributed to Trump view him less favorably after one week, or alternatively, that those who do not like Trump report that they believe him less after one week. The absence of significant correlations for the remaining Democratic and Republican groups indicated that their intentions to vote and feelings towards Trump were independent of belief change.
5.4.3 Discussion

Experiment 4 revealed several notable findings. First, when initially evaluating the veracity of both misinformation and factual statements, Republican supporters of Trump believed the information more when it was attributed to Trump, whereas the opposite occurred for Democrats. Republicans who did not support Trump also believed less in facts associated with Trump (but not to the same extent as Democrats), while their belief in the misinformation was not affected by information source. Overall, the Trump attribution did indeed colour people's assessment of information veracity, dictating how valid they perceived it to be.

Second, there was a large bipartisan shift in belief post-explanation, indicating that all members of the political spectrum are capable of substantial belief change when sound non-partisan explanations are presented. However, after a one-week delay, participants partially ‘re-believed’ in the misinformation and partially forgot that factual information was true. Thus, even if individuals update their beliefs temporarily, explanations regarding both fact and fiction seemingly have an expiration date (cf. Swire, Ecker, & Lewandowsky 2017). People revert to their original assumptions, highlighting that once inaccurate information is in the public sphere, it is difficult to permanently correct, and reservations regarding factual information are likewise challenging to permanently overcome.

From the pre-explanation belief scores, we know that Republican Trump supporters were predisposed to assume that information attributed to Trump was correct, regardless of its actual veracity. One week after the explanations, this bias continued for the misinformation items, but for factual items participants became less likely to think that Trump's statements were true. Similarly, Democrats—who are predisposed to assume that information attributed to Trump is incorrect—continued to exhibit bias for factual items, but were more likely to think Trump's
misinformation was true. It thus seems as if participants on both sides of the spectrum took into account their Trump-related biases but overcorrected for them: Republican supporters overcorrected by assuming that Trump's facts were false, and Democrats overcorrected by assuming that Trump's misinformation was true.

Third, Republican Trump supporters showed the highest level of post-explanation belief in misinformation in both Trump and unattributed conditions. This may reflect that only so much belief revision is possible (as their pre-explanation misinformation belief was also at a higher level), or alternatively that Republican Trump supporters were less inclined to believe our corrections.

Fourth, it was noteworthy that if the original information came from Donald Trump, after an explanation participants were less able to accurately label what was fact or fiction in comparison to the unattributed condition, regardless of their support for Trump. This was particularly the case for fact items after a delay, where even the Republican groups were less likely to think that the true information was indeed accurate if attributed to Trump.

Finally, while Republican supporters did update their beliefs when presented with corrections of misinformation, they did not change their voting intentions nor feelings towards Trump when the misinformation was attributed to the political figure. The degree that Republican supporters updated their belief that Trump's misinformation was false was not significantly correlated with a change in voting intentions nor feelings towards Trump. This suggests that the public, or at least Trump supporters, are not overly concerned with a candidate disseminating misinformation and seem to be looking to qualities other than veracity.

To test how processing of corrective/affirmative explanations is moderated by explanation source, we ran Experiment 5.
5.5 Experiment 5

Experiment 5 was conducted in July 2016. As in Experiment 4, participants were presented with inaccurate statements and factual statements that Donald Trump mentioned on the campaign trail in 2015, and the objectively false statements were corrected and the true statements affirmed. However, unlike Experiment 4, all statements were attributed to Trump. The other predominant difference between the two experiments was that we varied the nature of the explanations regarding the veracity of the information. In Experiment 5, the same explanations came from different partisan sources. Specifically, we randomised the attribution of the explanation to follow one of three forms: (i) ‘According to Democrats’, (ii) ‘According to Republicans’ or (iii) ‘According to a non-partisan fact-checking website’. Participants rated their belief in the statements both before and immediately after the explanation (though not one week later). The study thus used a $2 \times 3 \times 3$ design, with the within-subjects factors type of item (misinformation versus fact) and explanation source (Democrat versus Republican versus non-partisan), and a between-subjects factor of Trump support (Democrat versus Republican non-supporters versus Republican supporters). See figure 14 for a schematic of this design. Our prime dependent variables were participants' belief in the statements, as well as participants' self-reported support for Donald Trump.
Two potential outcomes were that (i) partisanship-congruent explanations would be more effective than partisanship-incongruent explanations due to greater support and trust in the source (e.g. Democrats being more influenced by a Democratic explanation; Swire, Ecker, & Lewandowsky, 2017; Guillory & Geraci, 2013; McGinnes & Ward, 1980) or (ii) a Democratic source would be more effective for all participants at affirming Trump's factual statements, and a Republican correction would be more effective at retracting Trump's misinformation, due to the surprise of an unlikely source presenting the explanation (Berinsky, 2015).

### 5.5.1 Method

**Participants.** Participants were 1019 US residents recruited through Survey Sampling International of Shelton, Connecticut. An over-18 population was recruited, targeting the census population for education, gender, age, geography and income, resulting in a diverse national sample. Participants were excluded from the analysis if they did not complete all parts of the study ($n = 59$). The final sample

---

13 A Pearson χ2-test indicated that there was no difference in Trump support between participants who were included and those who were excluded, $X^2 = (3, N = 983) = 5.42, p = 0.144.$
included $N = 960$ participants, with 456 males and 504 females. The age range was 19–86 years with a mean age of $M = 41.89$ (s.d. = 17.96).

**Stimuli.** As stimuli, we used six of the eight statements from Experiment 4: three inaccurate statements and three factual statements. The corrective/affirmative explanations were pseudo-randomly determined. Specifically, each item was attributed to one of the three different explanation sources (Republican, Democrat and non-partisan) in a counterbalanced manner, but we ensured that participants received all explanation sources during the experiment. This resulted in each participant seeing each of the respective explanation sources for one misinformation and one factual statement.

**Procedure.** Participants first rated their likelihood to vote for Donald Trump, and were then presented with all six statements in a randomised order. Participants rated the extent to which they believed each item to be true on a 0–10 scale, prior to receiving an explanation for each item as to whether it was true or false (with explanations coming from the three different sources). The test phase involved re-rating belief in all six statements in random order as well as re-rating Trump support immediately after all explanations were presented.

### 5.5.2 Results

Of the 960 participants, 514 identified as Democrats. Of the 286 Republicans, 186 were Trump supporters and 100 were Trump non-supporters. The 160 participants who had no political affiliation were omitted from the following analyses. There were 81 Democrats who supported Trump—all main effects and interactions of the following analyses were replicated if they were omitted from the analyses. A one-way ANOVA revealed there was a main effect of age, $F(2,797) = 4.88; p = .008; MSE = 328.70; \eta_p^2 = .01$, reflecting the fact that Republican non-supporters were younger than Republican supporters and Democrats.
The following analyses remained statistically significant when controlling for age using repeated measures general linear models. There were no gender differences between groups, \( p = .121 \), nor education differences, \( p = .346 \).

**Pre-explanation belief scores.** Pre-explanation belief scores by Trump support are shown in Figure 15. In a clear replication of Experiment 4, the Trump attribution led all participants to support the information to the extent that they supported Trump.

![Figure 15. Pre-explanation Democratic and Republican belief in statements associated with Trump. Error bars denote 95% confidence intervals.](image)

A 2 × 3 repeated measures ANOVA was performed on the pre-explanation belief scores. The analysis revealed two significant main effects. The main effect of type of item (misinformation vs. fact), \( F(1,797) = 322.37 \ p < .001; \ MSE = 2.13; \ \eta^2_p = .29 \), indicated that misinformation was believed less than facts. The main effect of Trump support (Democrats vs. Republican non-supporters vs. Republican supporters), \( F(2,797) = 114.49; \ p < .001; \ MSE = 8.27; \ \eta^2_p = .22 \), indicated that pre-explanation belief scores differed by Trump support. Republican supporters clearly believed Trump statements more than the other two groups; a planned comparison
also indicated that Republican non-supporters believed the information more than Democrats, $F(1,797) = 6.40; p = .012; MSE = 8.27.$

**Post-explanation belief scores.** The general trend and the full trajectory of pre and post-explanation belief change over time is shown in Figure 16. Immediately after the corrections/affirmations, both Democrats and Republicans showed a substantial amount of belief change—belief in misinformation reduced and belief in factual information increased. Partisanship and Trump support were far better predictors of the extent of belief updating than the explanation source.

![Figure 16. Belief in Trump misinformation and facts after partisan explanations. Rep = Republican, Misinfo = Misinformation. Dotted lines show misinformation items. Error bars denote 95% confidence intervals.](image)

A 3 × 3 repeated measures ANOVA involving explanation source (Democrat vs. Republican vs. non-partisan) and Trump support (Democrat vs. Republican supporter vs. Republican non-supporters) was performed on the post-explanation misinformation belief scores. The analysis revealed a main effect of Trump support, $F(2,797) = 19.15; p < .001; MSE = 20.72; \eta_p^2 = .05$, indicating that groups differed in their belief, with Republican supporters believing in the misinformation more than Republican non-supporters and Democrats. There was also a main effect of explanation source, $F(2,1594) = 6.01; p = .003; MSE = 4.81; \eta_p^2 = .007$, showing
that a Republican correction reduced belief to a greater extent than the Democratic or non-partisan corrections. However, it must be noted that this is a small effect size, and should be interpreted with caution.

To explore the observed trend that post-correction misinformation belief seemed to depend on the correction source in Republican non-supporters more so than in Democrats and Republican supporters, we ran an interaction contrast. Contrasting Republican non-supporters against the pooled Democrats and Republican supporters, and the Republican correction against the pooled Democrat and non-partisan corrections, revealed a significant effect, $F(1, 797) = 4.79; p = .029; \text{MSE} = 4.68$. It appears that misinformation belief was lowest after a Republican correction (vs. Democrat/non-partisan correction) in Republican non-supporters, $F(1, 797) = 9.69; p = .002$, whereas there were no effects of correction source on post-correction misinformation belief in Democrats or Republican supporters [all $F(1, 797) < 1.27; p > .257$]. However, as these were post-hoc analyses of a marginal effect, they too should be interpreted cautiously.

A $3 \times 3$ mixed ANOVA restricted to the post-affirmation fact belief scores revealed a main effect of Trump support, $F(2, 797) = 19.96; p < .001; \text{MSE} = 12.70; \eta_p^2 = .05$, indicating that Republican supporters were more accurate for fact belief than Republican non-supporters and Democrats.

**Likelihood to vote.** Figure 17 shows the full trajectory of participants’ likelihood to vote for Donald Trump, both prior to and after the corrective/affirmative explanation. Explanations regarding Trump statements did not greatly influence participants’ intention to vote. As in Experiment 4, to simplify the analysis, post-explanation scores were subtracted from the pre-explanation scores to create a vote change index.
Figure 17. Likelihood of voting for Trump across Trump support groups. Error bars denote 95% confidence intervals.

A one-way ANOVA on the likelihood to vote for Trump change index revealed a main effect of Trump support, $F(2,797) = 8.23; p < .001; MSE = 1.68; \eta^2_p = .02$, indicating that change differed between groups. Whilst Democrats and Republican supporters hardly shifted their rating (by +0.05 and -0.21, respectively), Republican non-supporters increased their likelihood to vote for Trump significantly more (by +0.44), $F(1,797) = 7.72; p = .006; MSE = 1.67$.

Analogous to Experiment 4, pairwise correlations were calculated for all Trump support groups to investigate whether belief change in misinformation or factual statements was associated a a likelihood to vote for Trump. As in Experiment 4, intentions to vote and feelings towards Trump were largely independent of belief change. However, using a Bonferroni-adjusted alpha level of .008, three correlations were significant: Democrats were shown to (1) reduce their likelihood to vote for Trump the more they reduced their belief in Trump-attributed misinformation ($r = .13$), as well as (2) increase their likelihood to vote for Trump the more they increased their belief in Trump-attributed facts ($r = .18$). Somewhat ironically, (3) Republican Trump supporters reduced their likelihood to vote for Trump when they increased their belief in the Trump-attributed fact items ($r = -.24$).
5.5.3 Discussion

Experiment 5 primarily investigated whether partisanship-congruent explanations were more effective than partisanship-incongruent explanations, or whether an unexpected explanation source would be more effective. Pre-explanation findings of Experiment 4 were replicated, as Republican supporters believed in the Trump-attributed misinformation and factual information to a greater extent than both the Republican non-supporters and Democrats.

Post-explanation, we found that the partisanship congruence of explanations did not have as large an impact as hypothesised, and post-explanation belief was rather dictated by the group membership of the individual (i.e. whether the participant was a Democrat, Republican non-supporter or Republican supporter). However, Republican non-supporters were somewhat more likely to update their misinformed beliefs if a correction was attributed to a Republican source. It is possible that a respected explanation source is particularly helpful when the initial information is from a source that is not respected, although this effect did not extend to true statements.

Finally, the increase in the Republican non-supporters' post-explanation likelihood-to-vote ratings could reflect a backfire effect—it is plausible that Republican non-supporters do not wish to be nudged by explanations that could be perceived as liberal, thus leading them to further support the Republican figure. However, as Experiment 5 did not have an unattributed control condition for comparison (as Experiment 4 did), it is uncertain whether or not this shift was due to the Trump attribution of the corrections.
5.6 General Discussion

The present research aimed to determine whether belief in misinformation and factual information depended on whether or not it stemmed from a politically polarising source, and whether it could be successfully corrected or affirmed. To this end, we presented participants with both inaccurate and factual statements made by Donald Trump on the campaign trail. Experiment 4 allowed us to investigate whether people use their support in political figures as a heuristic to guide their assessment of the veracity of this information, and Experiment 5 addressed whether partisanship-congruent explanations were more effective than partisanship-incongruent explanations. By keeping the content of the initial information and explanations stable across conditions, we were able to provide an accurate measure of a source's impact upon information processing.

5.6.1 Pre-explanation Belief Scores

We found that participants' opinion of Donald Trump influenced their assessment of information, that is, how valid they perceived it to be. The graded nature of information belief when it was attributed to Trump in comparison to the unattributed condition (i.e. Democrats decreasing, Republican supporters increasing, Republican non-supporters not affected as much) fits well with the graded intention to vote for Donald Trump, as revealed in Experiment 4. These findings are consistent with the findings from the literature regarding source credibility (Eagly & Chaiken, 1993). Given that attitude homophily is a crucial component of source credibility (Housholder & LaMarre, 2014), coupled with the notion that higher source credibility results in an increased perception of information credibility (Greer, 2003), it is reasonable that political figures such as Donald Trump act as a heuristic when evaluating the veracity of information.
Democrats showing lower levels of belief when information is attributed to Trump could reflect rational updating that takes the experienced base rates into account. However, this could also be an occasion where Democrats demonstrate equal biases to those of Republicans. While Republicans increased belief in inaccurate information if it came from a source they regard as trustworthy, Democrats indicated lower fact belief if the information came from a source they did not regard as trustworthy. Some of the true items used in this study are more aligned with traditional liberal ideology (e.g. that the USA spent $2 trillion on the war in Iraq), indicating that this effect holds even when processing factual information that could be considered worldview-congruent. This supports Kahan’s (2013) stance that biases such as motivated cognition could occur at both ends of the political spectrum, while running counter to the notion that people who hold right-wing ideology are more susceptible to motivated cognition in general. Our paper therefore contributes to mounting literature that all individuals—regardless of partisanship—are biased by their own worldview, rather than there being fundamental differences in cognition between people with differing political values (Blais et al., 2010; Ecker, Lewandowsky, Fenton, & Martin, 2014; Kahan, 2015; Nisbet, Cooper, & Garett, 2015; Parker-Stephan, 2013; Weeks, 2015).

5.6.2 Post-explanation Trump Attribution

Intriguingly, even when Trump statements were followed by credible explanations that ought to induce sustained knowledge revision and belief updating, in all groups there was a greater level of inaccuracy in comparison to the unattributed condition. This was particularly the case with regard to factual statements over the long term.

Republicans and Democrats seemed to take into account their Trump-related biases and overcorrected for them one week after the explanations: Republican
supporters by assuming that Trump's facts were false and Democrats by assuming that Trump's misinformation was true. There is precedent for such meta-cognitive effects in the political information-processing literature. Overcorrection has been seen to occur for mood-related biases when people assume their feelings are affecting their judgement and attempt to correct for their influence (Ottati & Isabel, 1996). For example, Isbell and Wyer (1999) found that participants rated political figures less favourably when participants were happy than when they were not, in an attempt to adjust for what they perceived to be an irrelevant affective influence. This overcorrection for biases appears to also influence the judgement of veracity when it comes to correcting misinformation and affirming factual information that stems from a polarising source.

It is important to highlight that Trump attribution has a relatively small effect size in comparison with the common effects of the retention interval in the post-explanation analyses. The consistency in belief updating and forgetting over time perhaps reflects that partisan effects are not as consequential as more general cognitive consequences such as the reversion to original assumptions over time.

5.6.3 Explanation Source

Different explanation sources did not have as large an impact as hypothesised. It is noteworthy in itself that the explanation source did not have as large an impact as the support of the person purporting the initial information. While Berinsky (2015) found that corrections from an unlikely source aided belief updating, this was when the to-be-corrected information was specifically counter to the traditional stances of a political party, for example, when Republicans debunked rumours regarding health care. It is possible that our amalgamation of items was not sufficiently in opposition to the core values of the Republican party to replicate these results. While it seemed that Republican non-supporters reduced their
misinformation belief most following a Republican correction, it is necessary to replicate these results due to the post hoc nature of the analysis.

5.6.4 Worldview Backfire Effects

There was no evidence for a worldview backfire effect in either experiment, as post-explanation misinformation belief scores remained below pre-explanation levels. In 2005, Nyhan and Reifler (2010) found a backfire effect in conservatives when trying to correct the belief that weapons of mass destruction were found in Iraq. Yet in 2006, this effect was not replicated—the correction led conservatives to appropriately update their belief. The authors argued that, between 2005 and 2006, conservatives came to place less importance on the war, suggesting that backfire effects may only occur when an issue is strongly and currently connected with an individual's political identity. In the present case, perhaps not all four pieces of misinformation resonated strongly enough with Republicans to create a notable backfire effect. The present pattern—obtained using a variety of real-world items rather than relying on only one contentious topic (as previous studies have; Hart & Nisbet, 2012; Nyhan & Reifler, 2013; 2015)—suggests that worldview backfire effects are not the norm and may only occur under very specific circumstances.

5.6.5 Voting Preferences

While it is possible that the observed changes in voting preferences between pre- and post-explanation are due to the presentations of the corrections and affirmations, it appears that the negative political ramifications of disseminating misinformation are limited. Belief change in Trump-attributed misinformation remained uncorrelated with a change in voting intentions and feelings towards Trump. Many individuals, and indeed political scientists, did not predict the success of Donald Trump (Dugan, 2015; Mayer, 2015). This study contributes one further piece of the puzzle as to why his success was sustained: spreading misinformation
did not hinder his candidacy, and even if misinformation was exposed, this did not reduce voting preferences or positive feelings. This could reflect that, to a certain extent, people expect politicians to make inaccurate statements (Skinner & Evans, 2015), thus they are not overly concerned when this expectation is met. Moreover, in the context of this study, providing an equal number of misinformation and factual items could have both reduced and boosted candidate support. Although people's opinions of a political candidate should ideally not increase if they hear the candidate made a factual statement—this should be an expectation rather than an added benefit—the equal presentation of misinformation and facts could explain the null effect. An avenue for future research would be to vary the proportion of true and false statements from the political figure that are provided to participants.

Understanding Donald Trump's popularity, despite the degree of misinformation he has distributed (Lippman, Samuelsohn, & Arnsdorf, 2016; “Post-truth”, 2016), is an interesting case study of American politics. However, it is uncertain to what extent the findings of the current experiments are in fact a ‘Trump phenomenon’. While he is perhaps a good candidate for the study of misinformation, political misinformation is common in the political arena (Horchschild & Einstein, 2015). To test whether the present findings are generalisable beyond Donald Trump, this experiment should be replicated with a Democratic and a different Republican political figure. Another potential barrier to generalisability is that the participants from Experiment 4 were Mechanical Turk workers. However, several studies have found that this population yields high-quality data, comparable to other convenience samples such as university students (Berinsky, Huber, & Lenz, 2012; Mullnix, Leeper, Druckman, & Freese, 2015), and Experiment 5 replicated Experiment 4's data trends in a more diverse sample.
There are many possible explanations for why Americans voted for Donald Trump in the primary and the general election: factors such as his perceived business acumen, his economic or immigration policies, or perhaps the fact that he was not a career politician increased his appeal (Newport & Saad, 2016a; Newport & Saad, 2016b). We cannot speak to these possibilities. This study illustrates that something other than veracity accounted for his success, as supporters did not change their voting intentions even if they altered their beliefs about the truth of his statements. If spreading falsehoods does not discredit character, it is perhaps not surprising that many individuals rallied behind him on election day (“Trump wins presidency”, 2016, Trump reaches delegate number”, 2016). According to Ramsay et al. (2010), 91% of voters said that information in campaigns sometimes seemed misleading or false, yet struggled to pinpoint exactly what is fact and what is fiction. The real-world consequences of this study suggest that politicians can seemingly spread misinformation without dramatic negative consequences of losing supporters—the results of the 2016 Presidential election are consistent with this interpretation. It thus appears that it is possible to appeal through the art of rhetoric and demagoguery rather than necessitating cohesive arguments constructed of logic and fact.
5.7 References


http://doi.org/10.1017/S0007123415000186


http://doi.org/10.1111/j.1460-2466.2006.00003.x


http://doi.org/http://dx.doi.ezproxy.library.uwa.edu.au/10.1371/journal.pone.0059837


http://doi.org/10.1017/s11109-010-9112-2


Chapter 6:  Misinformation and its Correction: Cognitive Mechanisms and Recommendations for Mass Communication
In 2007, a man in the United Kingdom posted a photograph on his website of a “mummified fairy” which he created as an April Fools’ prank. After receiving 20,000 visitors to the site in one day, he explicitly revealed that he had fabricated the scenario, yet many accused him of covering up the truth and vehemently insisted that the fairy was real (“Fairy fool”, 2007). This anecdote highlights a valid concern to mass communicators: regardless of how ridiculous information seems, once it is in the public sphere, it can take on a life of its own and may never be fully retractable.

It has become a societal norm that the media and the internet provide vast quantities of information, placing the onus on the individual to sort fact from fiction. However, individuals have limited time, cognitive resources, or motivation to understand complex topics such as scientific findings or political developments, and misconceptions are commonplace. Unfortunately, once inaccurate beliefs are formed, they are remarkably difficult to eradicate (Ecker, Lewandowsky, Swire, & D. Chang, 2011a). Even after people receive clear and credible corrections, misinformation continues to influence their reasoning: in cognitive psychology, this is known as the *continued influence effect of misinformation* (H. Johnson & Seifert, 1994; Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012). The mummified fairy is a benign example, but the ramifications can be serious. Belief in misinformation can adversely impact decision making, and the continued influence effect has real-world implications in areas as disparate as education, health, and the economy.

One prominent example is the misconception that the measles, mumps, rubella (MMR) vaccine causes autism. This falsehood has been repeatedly—and convincingly—retracted by the media and scientific community over a number of years since the original myth was disseminated in a fraudulent article. Despite these debunking efforts, the myth has led to a drop in vaccination rates, and an increase in vaccine-preventable disease (Poland & Spier, 2010). The economic burden of 16
measles outbreaks in the US in 2011 alone has been estimated somewhere between $2.7 million and $5.3 million (Ortega-Sanchez, Vijayaghavan, Barskey, & Wallace, 2014). Thus, developing evidence-based recommendations on how to adequately communicate corrections and minimise reliance upon inaccurate information is not only important for individual decision making but also has ramifications for society as a whole.

The most important recommendation for both traditional mass media such as newspaper and television, as well as more recent technologies such as Twitter—which have essentially transformed ordinary citizens into would-be journalists—is to take greater care to ensure that information is correct to begin with. However, this is not always realistic due to the fast pace of modern information consumption and dissemination, and the fact that ordinary citizens are not bound by rules of journalistic integrity. Social media is thus an ideal breeding ground for the propagation and transition of misinformation, which can be exemplified by its role in rumors surrounding the Boston Marathon bombing in 2013. For example, a well-intentioned Reddit thread was created to help find the perpetrators, yet the accusation of an innocent and deceased Brown University student subsequently went viral (Guzman, 2013). Information shared through social media is usually disseminated without fact-checking, based merely on its potential to elicit emotional responses or support a personally motivated argument (Peters, Kashima, & Clark, 2009).

This chapter focuses on cognitive mechanisms and theories accounting for the continued influence of misinformation. In particular, we will discuss what drives belief in inaccurate information, why certain individuals are predisposed to refrain from belief change even in the face of good corrective evidence, and how corrections can be designed to maximise impact. We therefore provide six practical recommendations based upon our current knowledge of cognitive processes. We first
discuss theoretical accounts for the continued influence effect such as mental models, dual processing theory, the necessity of co-activation of misinformation and new information, and the impact of the information’s source. We then discuss individual predispositions to the continued influence effect, in particular a person’s worldview and scepticism.

6.1 Mental Models

When people initially encounter information, a situation model of integrated memory representations is built, and this model is continuously updated as new information becomes available and relevant (Bower & Morrow, 1990). If the required changes are small, they can be integrated into the situational model incrementally (Bailey & Zacks, 2015), yet if a larger change is required, a “global” update is necessary, which involves discarding the old mental model and creating a new one (Kurby & Zacks, 2012). However, even if there are sufficient cognitive resources to notice a difference between one’s mental model and the current environment, people are often quite inadequate at assimilating new information or mapping it onto existing memory representations (van Oostendorp, 2014). It is possible that the continued influence effect occurs when people update incrementally when in fact a global update is called for. Reliance on inaccurate information is less likely in instances when there is an alternative to replace the inaccurate information in a person’s mental model, as a readily available alternative explanation facilitates global updating (Ecker, Lewandowsky, Cheung, & Maybery, 2015).

A classic paradigm for studying the continued influence effect involves presenting participants with fictitious scenarios involving the retraction of an event cause. One common example is a narrative where negligent storage of gas cylinders is initially held responsible for starting a warehouse fire, yet their presence is
retracted shortly thereafter (H. Johnson & Seifert, 1994; Wilkes & Leatherbarrow, 1988). If participants are explicitly queried about the gas cylinders, they typically acknowledge the gap in their understanding (i.e., a gap in their mental event model) created by the earlier retraction, and correctly state that there were none. However, when answering inferential reasoning questions regarding the event—such as “what was the cause of the explosions?”—participants often still rely upon the outdated information. This indicates that people prefer to have an inaccurate over an incomplete event model, which can lead to reliance upon discredited information even after an explicit correction (Ecker, Lewandowsky, & Apai, 2011b).

6.2 Recommendation 1: Providing Factual Alternatives

One of the most effective methods of correcting misinformation is to provide an alternative factual cause or explanation to facilitate “switching out” the inaccurate information in an individual’s initial situation model. For example, if people are told that it was not gas cylinders that caused the warehouse fire, but that there was evidence of arson, people are dramatically less likely to rely upon the original inaccurate information (H. Johnson & Seifert, 1994; Ecker, Lewandowsky, & Tang, 2010). The alternative explanation effectively plugs the model gap left by the retraction. The alternative should ideally have the same explanatory relevance as the misinformation it replaces, and it is important that it is plausible—in fact, if the new information is more plausible and easy-to-understand than the original, updating is even more efficient (Baadte & Dutke, 2012).

In the real world, providing an alternative explanation to ameliorate reliance upon inaccurate information can be problematic, as often there is no available substitute—sometimes all that can be said about a piece of misinformation is that it is not true. For example, if a person is accused of a crime, they might simply
turn out to be “not guilty” without an alternative suspect being readily available. The lack of adequate alternatives can have profound ramifications. For example, the ongoing rumors regarding missing Malaysian Airlines flight MH370, which disappeared over the Indian Ocean in 2014, have proven difficult to retract: In the absence of unequivocal evidence regarding what happened to the plane, traditional and social media was rife with speculations that the plane was hijacked by terrorists or a suicidal pilot (e.g., Quest, 2016). Arguably, belief in the hijacking speculation has been difficult to shift because a convincing factual alternative has not been available.

6.3 Dual Process Theory: Strategic and Automatic Memory Processes

The notion that retractions create gaps in mental models is useful to understand the basic phenomenon that is the continued influence effect. Invalidated information is not simply deleted from memory—memory does not work like a whiteboard and retractions do not simply erase misinformation. To explain why corrected misinformation is used during reasoning, some theorists have focused on the memory processes governing information retrieval, where a common assumption is that there are two separates types of memory retrieval, strategic and automatic (Yonelinas, 2002).

Strategic memory processes are effortful and allow for the controlled recollection of the information’s contextual details. Similar to the meta-data of a computer file, contextual details include information about the information itself. This includes qualities such as the information’s spatiotemporal context of encoding, source, and veracity (Frithsen & Miller, 2014). A person’s ability to use strategic memory processes efficiently will depend upon factors such as effort, motivation, the period of time since encoding, and age (e.g., Herron & Rugg, 2003). In contrast,
automatic processes are fast and relatively acontextual, and serve to quickly provide an indication of memory strength or familiarity with an item or notion (Zimmer & Ecker, 2010).

Automatic retrieval processes can contribute to misinformation effects in two ways. Firstly, the evaluation of a statement’s veracity is influenced by its familiarity; this is problematic as information can be accepted as true just because it seems familiar. When increased familiarity gives the illusion that information is valid, this is known as the *illusory truth effect* (e.g., Begg, Anas, & Farinacci, 1992). Secondly, when questioned about an event or otherwise cued, retracted misinformation can be automatically retrieved from memory without any accompanying contextual details, and potentially without recalling that the information has been retracted (cf. Ayers & Reder, 1998; Ecker et al., 2010). To illustrate, it has been argued that once misinformation has been encoded and then retracted, a “negation tag” is linked to the original memory representation (e.g., “Flight MH370 was hijacked—*NOT TRUE*”; cf. Gilbert, Krull, & Malone, 1990). When queried about the topic, fast automatic memory processes might simply retrieve the familiar claim, while strategic memory processes are required to retrieve the negation tag and dismiss the familiar statement as untrue. If strategic memory processes are not engaged, familiar claims are thus likely to be judged as true even after plausible retractions (Dechene, Stahl, Hansen, & Wanke, 2010).

**6.4 Recommendation 2: Boosting Retrieval of the Retraction, Not Familiarity of the Myth**

The extent to which people engage their strategic memory processes can be actively encouraged, and this can reduce misinformation effects. Ecker et al. (2010) found that presenting participants with a pre-exposure warning detailing the continued influence effect greatly reduced reliance on misinformation, and was as
effective as providing a factual alternative. The authors argued that warnings not only allowed individuals to more effectively tag misinformation as false when encoding its retraction, but also boosted later recall of the retraction (or the “negation tag”). The effect of warnings was investigated mainly for theoretical reasons, and providing a pre-exposure misinformation warning will not be a viable option in most real world settings. However, any incentive to engage in strategic memory processes should be useful, such as boosting source-monitoring (Lindsay & M. Johnson, 1989; Poole & Lindsay, 2002).

Enhancing recollection is one way of reducing reliance on misinformation, but circumventing the inflation of a misconception’s familiarity is potentially another way. This involves minimising unnecessary explicit repetition of misinformation. For example, an educational pamphlet using a “myth-busting” format that repeats the myth before indicating that it is false (e.g., “Flight MH370 was hijacked—FALSE”) can boost the familiarity of the misconception, potentially increasing the risk that misconceptions are later mistakenly remembered as being true. This misremembering of myths as facts was demonstrated by Skurnik, Yoon, Park, and Schwarz (2005), as well as Peter and Koch (2016). In both these studies, participants misremembered the originally false statements as true more often than misremembering originally true statements as false. Additionally, Swire, Ecker, and Lewandowsky (2017) found that retracting myths and affirming facts led to comparable belief change initially (i.e., belief reduction for myths, belief increase for facts), but that belief change was less sustained with myths over the course of a one-week period. In other words, misinformation began to be “re-believed” while fact belief remained stable. Thus, where possible, communicators should focus on the facts and explicit repetition of a myth should be minimised if the retraction does not provide adequate information to allow people to revise their understanding.
6.5 Co-activation of Misconception and Corrective Facts

Despite the theoretically motivated suggestion to avoid myth repetition, for practicality, corrections usually do require repetition of the myth—the question then becomes how best to execute this. As discussed previously, presentation of factual alternative information is conducive to successful mental-model revision. Beyond that, several theoretical accounts have proposed that the co-activation of inaccurate knowledge and newly encoded factual information facilitates knowledge revision. Co-activation is believed to increase the likelihood that the individual notices discrepancies between originally-held misconceptions and factual evidence, and that they update their knowledge accordingly (Kendeou & van den Broek, 2007).

After a correction, both the outdated and new information may co-exist in memory, and can both be activated by relevant cues (cf. Ayers & Reder, 1998). Thus, it is crucial for efficient updating and knowledge revision that a sufficient amount and quality of factual information is provided, and ideally, that the correction also explains the reasons as to why the misconception is wrong (Seifert, 2002). Adding adequate detail to the new accurate information can systematically strengthen the correction by slowly decreasing interference from the outdated information (Kendeou, Smith, & O’Brien, 2013). This illustrates how when ample factual information is available, misinformation can be used as an educational tool (Bedford, 2010).

6.6 Recommendation 3: Refutations of Misinformation as an Educational Tool

A refutation involves not only a statement that the misconception is false, but a comprehensive explanation as to why it is incorrect (Hynd, 2001). The efficacy of refutations has primarily been investigated in the field of education, and has often focused on the updating of scientific misconceptions held by students in a classroom.
A meta-analysis of 70 studies by Guzzetti, Snyder, Glass, and Gamas (1993) indicated that corrections are most successful when they include sufficient explanation as to why a misconception is false (and why the facts are true). Other educational strategies aimed at reducing reliance on misinformation such as class discussions, demonstrations, and non-refutational texts (which simply present the correct information without a description of the misconception itself), are often successful in the short term, but not after a delay (Guzzetti, 2000).

It has been argued that the relative success of the refutation at promoting belief change is that, by design, it increases the likelihood of the old and new information being co-activated in memory (Kowalski & Taylor, 2009). It follows that when debunking a myth, its repetition seems acceptable (despite the potential myth-familiarity boost) as long as (1) the repetition serves to highlight a discrepancy between a misconception and factual evidence, thus promoting co-activation, (2) the focus of the intervention can be shifted promptly from the myth to the factual evidence, and (3) the target audience has the necessary resources—in particular in regards to time and motivation—to engage with the provided materials and sees the information source as credible, as would hopefully be the case in a classroom setting.

### 6.7 Retraction Source Credibility

People often do not have the time or inclination to be an expert in all fields, so most knowledge, to a degree, is reliant upon accepting what someone else (or google) claims to be true. Thus, people hold many opinions and beliefs about events and causal relationships without having relevant involvement or expertise. For example, trust in climate scientists is a predictor of whether or not an individual acknowledges that climate change is anthropogenic (Mase, Cho, Prokopy, 2015). In general, high-credibility sources are more persuasive than low-credibility sources.
(Eagly & Chaiken, 1993), and the lower one’s prior knowledge regarding a topic, the more influential source credibility becomes (Jung, Walsh-Childers, & Kim, 2016). The two core factors of source credibility discussed in the literature are (1) expertise—the extent to which the source is capable of providing accurate information, and (2) trustworthiness—the perception that the source is willing to provide information that the source itself believes to be accurate (Pornpitakpan, 2004). A source can independently have varying degrees of these two qualities, for example, a doctor may have a high degree of (perceived) expertise, but if found to be paid by pharmaceutical companies may have relatively low (perceived) trustworthiness.

When it comes to retracting inaccurate information or belief change, intriguingly trustworthiness seems to play a much larger role than expertise (McGinnes & Ward, 1980). For example, Guillory and Geraci (2013) investigated the credibility of retraction source by presenting participants with a story about a politician who was witnessed taking a bribe. This was later retracted by people with varying degrees of trustworthiness and expertise. The authors found that although trustworthiness was integral to the success of the retraction, expertise was not. It should be noted that the way expertise was operationalised in this study was more akin to “involvement in an event” rather than expertise in its perhaps more common meaning (i.e., “possessing relevant knowledge”). However, Ecker and Antonio (2016) replicated Guillory and Geraci’s main finding with a more traditional interpretation of expertise and also found an effect of trustworthiness but not expertise on the effectiveness of retractions.
6.8 Recommendation 4: Building Credibility

The ability to successfully correct misinformation appears to rely more upon the source’s perceived honesty and integrity than its expertise. This means that Leonardo DiCaprio’s 2016 Oscar speech correcting climate-change misconceptions (Goldenberg, 2016) could be more effective than an expert communication. Additionally, Paek, Hove, Jeong, and Kim’s (2011) found that YouTube videos created by peers had more impact in terms of attitude change than videos created by a non-profit organisation. This means that social media can be an effective vehicle for influencing others, and Facebook or Twitter posts may have more influence on friends’ opinions than expert advice.

Ideally, and ethically, science communicators should aim to combine high trustworthiness with high expertise. The quality and accuracy of the presented information will influence how the source itself is perceived—this includes factors such as the information’s presentation, plausibility, and whether or not it is supported by good examples (Jung et al. 2016; Metzger, 2007). In general, perception of a source seems to be an iterative process in that the more quality information is released, the greater the level of perceived credibility. In mass communications in particular, basing claims on evidence, adequately referencing the evidence, and presenting data in an easily accessible way to minimise misinterpretations—and doing this consistently—will build credibility and thus contribute to a greater efficacy of corrections (Gigerenzer, Gaissmaier, Kurz-Milcke, Schwartz, & Woloshin, 2007).

6.9 Worldview

If an individual holds a strong belief that is fundamental to their identity, even the most credible source may not be able to shift it. A person’s ideology often
influences how information is sought out and evaluated, and if the information runs counter to prior beliefs, it is likely to be ignored or more critically appraised (Wells, Reedy, Gastil, & Lee, 2009). This is known as motivated reasoning (Kunda, 1990). Motivated reasoning can be compounded due to the formation of ideological “echo-chambers,” where information is exchanged primarily amongst people with similar viewpoints, such that corrections are less likely to reach the “target” audience (Barbera, Jost, Nagler, Tucker, & Bonneau, 2015). This is fostered by social media, where misinformation tends to circulate quicker than associated corrections (Shin, Jian, Driscoll, & Bar, 2016).

Even if a correction reaches the misinformed target audience, simply providing the correct information is inefficient, as continued reliance on misinformation is likely when the misinformation conforms to a person’s pre-existing belief system, yet the correction does not (Lewandowsky, Stritzke, Oberauer, & Morales, 2005s). Retracting misinformation that runs counter to a person’s worldview can ironically even strengthen the to-be-corrected misinformation, a phenomenon known as the worldview backfire effect; this has been demonstrated when correcting misinformation surrounding contentious issues such as climate change (Hart & Nisbet, 2012), or vaccine safety (Nyhan & Reifler, 2015). Worldview biases are particularly difficult to overcome, as even neutral coverage of an issue can lead to polarisation (Jerit & Barabas, 2012).

6.10 Recommendation 5: Provide Worldview or Self-Affirming Corrections

If a correction is regarding a contentious topic or politically sensitive subject matter, it is beneficial to frame the correction in such a way that it is congruent with the person’s values in order to reduce perceived threat (Kahan, 2010). For example, conservatives are more likely to accept anthropogenic climate science if it is
presented as a business opportunity for the nuclear industry (Feygina, Jost, & Goldsmith, 2010). Additionally, in line with the above-mentioned effects of source credibility, worldview congruence can potentially be conveyed through the appropriate choice of messenger. Callaghan and Schnell (2009) found that attitudes towards gun control were affected not only by the way the information was framed, but also the source of the message. Participants who were presented an argument regarding the impacts of crime and violence were 19% more likely to support gun control measures if the message came from a New York Times journalist than if it was presented without a source. People also seem less defensive regarding counter-attitudinal information when their self-worth is strengthened. For example, Cohen, Aronson, and Steele (2000) demonstrated this effect of self-affirmation: participants who had been instructed to write about a personal quality that made them feel good about themselves were subsequently more likely to respond positively to evidence that challenged their beliefs regarding the death penalty.

6.11 Scepticism

Rather than evidence-denial driven by motivated reasoning, scepticism is the awareness of potential hidden agendas and a desire to accurately understand the evidence at hand (Mayo, 2015). Scepticism can reduce misinformation effects, as it leads to more cognitive resources being allocated to the task of weighing up the veracity of both the misinformation and the correction. For example, people rely less upon misinformation when given the task of fact checking, looking for inconsistencies and correcting inaccuracies as they read a text (Rapp, Hinze, Kohlhepp, & Ryskin, 2014). The increased deliberation over the accuracy of information is often instigated when the information counters an individual’s worldview (Taber & Lodge, 2006). To illustrate, Lewandowsky et al. (2005) found
that a greater degree of scepticism led to better discounting of retracted real-world news reports, and DiFonzo, Beckstead, Stupak, & Walders (2016) found that individuals with greater dispositional scepticism tended to believe inaccurate rumors to a lesser extent. The ability to maintain doubt, question evidence and scrutinise the original data—even when it aligns with one’s worldview—is conducive to avoiding reliance on misinformation, but it is a difficult task. Thus, honing the skill of knowing when to trust evidence, and when not to, can potentially have great benefits.

### 6.12 Recommendation 6: Fostering Scepticism

Scepticism is a quality that can be encouraged and even temporarily induced—for example, negative mood increases scepticism and improves accuracy in detecting deceitful communications (Forgas & East, 2008). There is also a growing movement suggesting that evidence-based evaluation and critical thinking should formally be taught in schools. Schmaltz and Lilienfeld (2014) suggested that activities such as asking students to identify pseudoscience on campus and in the media could highlight the plethora of falsifiable claims in the public sphere. Alternatively, the authors recommended activities where students create their own pseudoscience to demonstrate and experience the ease with which anecdotal evidence or “psychobabble” can be fabricated. Even examining real-world false advertising cases can be educational, for example, investigating the Federal Trade Commission’s verdict to charge Lumosity $2 million for claiming its brain training could protect against cognitive impairment, or Dannon $21 million for claiming their yoghurt can prevent the flu (Lordan, 2010; Rusk, 2016). Lastly, the ability to question causal illusions—the perception that one event caused another, where in fact they are unrelated—can also be taught, and a better understanding about the probability of an
outcome, the probability of a cause, and cause-outcome coincidences can help promote scepticism (Matute et al., 2015).

6.13 Conclusion

Assessing the accuracy of information can be a difficult task. In today’s fast-paced society, mass communication and social media play a key role in the sharing and receiving of current events. In reality, the public do not have time to investigate each claim they encounter in depth; therefore, providing quality information is essential. In the aftermath of Brexit and the 2016 US election, where the political landscape was rife with misinformation and fake news (Barthel, Mitchell, & Holcomb, 2016; McCann & Morgan, 2016), the ability to correct inaccuracies has never seemed more pertinent. The six recommendations provided can serve as guidelines for mass communication as to how best to retract the plethora of misinformation in the public sphere. However, it is important to note that no corrective technique can reduce belief to base level, as if the misinformation was never previously mentioned. In addition, even if people do shift their opinion and acknowledge that information they previously believed to be true is incorrect, they are unlikely to change their voting preferences or feelings towards political candidates (Swire, Berinsky, Lewandowsky, & Ecker, 2017). Given what we know about misinformation and its correction, communicators thus hold a great deal of responsibility to ensure that the information initially released is as accurate as possible.
6.14 References


Chapter 7: Conclusion
7.1 Summary of Empirical Findings

The present thesis contributes to the literature on why belief in misinformation persists even after valid corrections have been presented. It provides insight into three distinct questions: (1) How memory processes contribute to the continued influence effect of misinformation; (2) Whether some correction formats are better at promoting belief change than others; and (3) How people process information when it stems from a polarising source, given their group identity. Each experiment in this thesis was designed to yield widely-applicable conclusions by using real-world misconceptions and factual information. Additionally, the experiments sought to gain an understanding of how belief changes over time. By systematically measuring belief at multiple time points after a correction was elicited, we were able to track how belief shifts as time progresses.

Chapter 3 aimed to clarify if and under what conditions myth familiarity is most problematic. We concluded that familiarity does indeed contribute to the continued influence effect of misinformation. Older adults over the age of 65 were worse at sustaining their belief that misinformation is incorrect, detailed information promoted belief change, and people were more likely to correctly label misinformation as false immediately after it had been corrected in comparison to long delays. However, a familiarity backfire effect failed to be elicited, even under conditions most conducive. This suggests that the myth-familiarity boost associated with a correction is unlikely to outweigh the benefits of a correction. Nevertheless, it remained unclear whether avoiding repeating the misconception within the retraction would be beneficial to belief updating.
Chapter 4 investigated whether different explanation formats promoted different levels of belief change, including whether the avoidance of the initial misconception within the retraction was beneficial. We found that the explanation format was not a strong determinant of belief change. There was no perceived difference between corrections that explicitly mentioned the misconception and those that did not. Thus, it is possible that the associated myth-familiarity boost with the correction is relatively inconsequential when paired with a valid retraction. Based on these findings, it seems necessary to tone down recommendations to avoid myth repetitions in corrective communications (see Cook & Lewandowsky, 2011; Lewandowsky et al., 2012). While it still seems wise to avoid any unnecessary myth repetition, there is little need for concern about repeating myths that serve to provide a clear and salient correction (also see Ecker, Hogan, & Lewandowsky, 2017; Kendeou & O’Brien, 2014; Kendeou et al., 2017; Schwarz, Newman, & Leach, 2016).

Finally, Chapter 5 aimed to examine the impact of source credibility on the assessment of veracity when initial information or the corrective information comes from a polarising source. We found that people were more likely to believe the information if it comes from a figure who is trusted, whereas the opposite was true if they were not. However, it appears that the source of the correction was not as consequential as initially predicted (also see Guillory & Geraci, 2013; van Boekel, Lassonde, O’Brien, & Kendeou, 2017). As voting preferences for the political candidate eliciting the misinformation remained stable even when the belief in the misinformation reduced, this suggests that people do not necessarily insist on veracity as a prerequisite for supporting political candidates.
7.2 Thesis Inconsistencies and Limitations

There are two thesis inconsistencies that warrant further discussion: (1) Chapter 6 recommends that one should “boost retrieval of the retraction, rather than the familiarity of the myth”, in spite of finding no evidence to support this claim in Chapter 4, and (2) factual belief was sustained over the course of one week in the experiments of Chapter 3, but it was not sustained over one week for those in Chapter 5.

Relating to point (1), the Chapter 6 recommendation that the “retrieval of the retraction should be boosted, rather than the familiarity of the myth”, needs to be qualified in light of the Chapter 4 data (which was analysed after Chapter 6 was in press). As Chapter 4 found that avoiding the misconception within the retraction to prompt equal levels of belief change over time, the avoidance of the initial misconception will not necessarily be of benefit. In other words, stating “we use 100% of our brain” will likely be as effective as a correction that repeats the original misconception, for example, “it is not true that we only use 10% of our brain”.

Despite earlier warnings regarding a familiarity backfire effect (Cook & Lewandowsky, 2011; Lewandowsky et al., 2012; Schwarz et al., 2007), there is now growing evidence that the boost of familiarity is not as detrimental to belief updating as previously thought, particular if paired with a believable correction (Ecker, Hogan, & Lewandowsky, 2017; Kendeou, Walsh, Smith, & O’Brien, 2014; Rich, Van Loon, Dunlosky, & Zaragoza, 2017; Swire, Ecker, & Lewandowsky, 2017).

However, it is still a valid recommendation to avoid unnecessary repetition of the misconception. This is because needlessly repeating the misconception could potentially spread the misinformation to new audiences (Schwarz, Newman, & Leach, 2016). Moreover, it is yet to be seen whether the backfire effect occurs in...
other circumstances when strategic memory processes are debilitated. For example, divided attention could theoretically create an occasion where the benefit of the correction is outweighed by the familiarity boost (Jennings & Jacoby, 1993; Troyer & Craik, 2000).

Relating to point (2), factual belief was sustained over a one week period for some experiments in this thesis, and not others. One hypothesis is that the factual information was more novel in experiments from Chapter 3 in comparison to the experiments in Chapter 5. If a person rates an item 3 out of 10 on the belief scale because it sounds like it is unlikely to be correct (but has not been encountered before), this would be very different from a person who scores a 3 out of 10 because they have a preconceived notion that the statement is likely to be incorrect. The low rating of the novel item may just reflect guessing, whereas the same rating for a previously encountered assertion may reflect genuine scepticism.

This discrepancy in stimuli may have led to different processing of affirmations in the experiments reported in Chapters 3 and 5. For example, people may be more likely to have preconceived notions for political stimuli—such as the cost of the Iraq war—over and above curious statements such as “humans can regrow the tips of fingers and toes after they have been amputated”. Unfortunately, the stimuli from Chapter 5 were not pilot tested for a-priori familiarity, so the available data do not allow assessment of this hypothesis. Future research could explicitly compare novel vs. already-encountered items to further pin down the effects of familiarity. This could help confirm whether a contributing factor of the sustained belief change observed in Chapter 3 is that participants did not process a “corrective” affirmation, they were merely informed for the first time that the information was true.
7.3 Concluding Remarks

In today’s complex modern society, fast-paced journalism and social media shape the social narrative. During the course of this thesis, misinformation has sprung into the limelight for being a legitimate and sinister social concern, and the ability of beliefs based on “alternative facts” to sway elections and mobilise voters has become apparent. Unfortunately, the task of correcting misinformation is far from trivial, and we have just begun to scratch the surface of why the continued influence effect occurs. It is likely that inaccurate beliefs do not solely persist due to memory processes, nor solely due to the biases connected to our ideological beliefs. This thesis has touched on both these facets for a comprehensive investigation into the continued influence effect. It is integral to continue this work to further improve our understanding of the underlying mechanisms of the continued influence effect, and to continue the endeavour to develop techniques to facilitate evidence-based updating. In lieu of dire alternatives—such as clamping down on free speech—it would be far superior if people were able to disregard fake news and misinformation simply because they had the tools to do so.
7.4 References


Appendix A: Chapter 3 Stimuli
<table>
<thead>
<tr>
<th>Item number</th>
<th>Items</th>
<th>Brief explanation</th>
<th>Detailed explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myth - 1</td>
<td>The daddy long legs spider is extremely venomous, but fortunately its fangs are too small to pierce human skin.</td>
<td>The daddy long legs spider is not extremely venomous, nor are its fangs are too small to pierce human skin.</td>
<td>While the fangs of the daddy long legs spider are indeed very small (estimated at 0.25mm in length), they are capable of piercing human skin. Additionally, tests have revealed that their venom is no more powerful than most other spiders, and not dangerous to human health. It is possible that the reputation of the daddy long-legs arose because they can kill red back spiders. Perhaps the logic went that spiders that kill red backs must be even more dangerous than the red backs themselves.</td>
</tr>
<tr>
<td>Myth - 2</td>
<td>A tooth partially dissolves if you leave it in cola overnight.</td>
<td>A tooth will not partially dissolve if you leave it in cola overnight.</td>
<td>In 1950, Professor Clive McCay of Cornell University told the US House of Representatives that cola causes tooth decay. He also stated that if a tooth was placed in a glass of cola, it would begin to dissolve. This latter statement is not true, as anyone with a loose tooth can find out for themselves. Over a long period of time, drinking cola is harmful for your teeth due to the extremely high sugar content and phosphoric acid. Unfortunately, you will not be able to see this exemplified overnight with your tooth in a glass of cola.</td>
</tr>
</tbody>
</table>
| Myth - 3    | Liars sometimes give themselves away by physical ‘tells’ such as looking to the right or not looking you in the eye. | Liars do not give themselves away by physical ‘tells’ such as looking to the right or not looking you in the eye. | Physical signals which are often assumed to be the ‘tells’ of a liar, are in fact signs of emotional discomfort in general. When a person is being interviewed or is accused of a crime, a non-liar is equally likely to express these signals. A meta-analysis of over 100 studies found no consistent physical cues when a person was lying. The experimenters stated that “there are no behaviours that always occur when people are lying and never occur when they are telling the truth”.

Appendices
<table>
<thead>
<tr>
<th>Item number</th>
<th>Items</th>
<th>Brief explanation</th>
<th>Detailed explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myth - 4</td>
<td>Females, on average, talk more than males.</td>
<td>Females, on average, do not talk more than males.</td>
<td>There have been a number of studies which have converged upon the conclusion that females do not talk more than males. One researcher recorded college students’ speech between 1998 and 2004. Hundreds of students wore digital devices that recorded 30 seconds of sound every 12.5 minutes. From this data they estimated that both men and women say around 16 000 words per day. Amusingly, the most talkative person in the study (who said an astonishing 47 000 words a day) was in fact a man.</td>
</tr>
<tr>
<td>Myth - 5</td>
<td>Carrots can improve eyesight.</td>
<td>Carrots do not improve eyesight.</td>
<td>Eating carrots will not improve your visual acuity, nor will they make your eyes better if you have less than perfect vision. Even if your eye problems are due to a vitamin A deficiency, both sweet potato and liver contain more vitamin A than carrots. The origin of this myth is the Second World War. The British government encouraged rumours that Captain John Cunningham and his squadron were able to see in the dark because they ate so many carrots, when in fact they were testing out the newly developed airborne radar system.</td>
</tr>
<tr>
<td>Myth - 6</td>
<td>Playing Mozart to your baby will boost its IQ.</td>
<td>Playing Mozart to your baby does not boost its IQ.</td>
<td>This myth originated from a study published in <em>Nature</em> in 1993. Participants temporarily performed better at spatial reasoning tasks after listening to 10 minutes of Mozart’s Sonata for Two Pianos in D Major. IQ was not mentioned in the study, because IQ was never measured. Somehow, the findings have been twisted to apply to infants and IQ in general, and popularised in an unscientific book entitled “The Mozart effect: Tapping the power of music to heal the body, strengthen the mind and unlock the creative spirit”.</td>
</tr>
<tr>
<td>Item number</td>
<td>Items</td>
<td>Brief explanation</td>
<td>Detailed explanation</td>
</tr>
<tr>
<td>-------------</td>
<td>------------------------------------</td>
<td>-------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Myth - 7</td>
<td>Cracking knuckles can cause arthritis.</td>
<td>Cracking knuckles does not cause arthritis.</td>
<td>There is no correlation between cracking your knuckles and arthritis. One interesting example is Dr. Unger, who decided as a child to crack the knuckles in his left hand diligently, while never cracking that of his right hand. After 60 years, he published his findings in the Journal <em>Arthritis and Rheumatism</em>, demonstrating that there was no significant difference between his left and right hand. He was awarded the 2009 Ig-Nobel prize, which is awarded for research that makes you laugh, then think.</td>
</tr>
<tr>
<td>Myth - 8</td>
<td>Using a defibrillator is the best way to restart a stopped heart.</td>
<td>Using a defibrillator is not the best way to restart a stopped heart.</td>
<td>Defibrillation delivers a dose of electrical energy to the heart, and is only used when the patient has a dangerously <em>irregular</em> heart rhythm. Once the heart has actually stopped it cannot be helped by electrical means, and an inappropriate shock would actually hinder the patient’s chances of survival. So, although a popular scenario for television dramas involves a panicked doctor reaching for the defibrillator as a patient’s heart beat flat lines, in real life, re-establishing a heartbeat would never involve a defibrillator.</td>
</tr>
<tr>
<td>Myth - 9</td>
<td>Mercury in vaccines can cause harm.</td>
<td>Mercury in vaccines cannot cause harm.</td>
<td>While methyl mercury builds up in the body, ethyl mercury (the type of mercury within vaccines) is excreted rapidly. In 2006, an expert panel assembled by the World Health Organisation concluded that there was “no evidence of toxicity in infants, children or adults exposed to [mercury] in vaccines”. Keep in mind that toxicity is dependent upon concentration: the minute amount of mercury within vaccines is 10 times below the lowest level calculated to potentially cause harm.</td>
</tr>
<tr>
<td>Item number</td>
<td>Items</td>
<td>Brief explanation</td>
<td>Detailed explanation</td>
</tr>
<tr>
<td>------------</td>
<td>----------------------------------------------------------------------</td>
<td>------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Myth - 10</td>
<td>Giving sugar to children makes them hyperactive.</td>
<td>Giving sugar to children does not make them hyperactive.</td>
<td>The best way to test whether sugar causes hyperactivity in children is using ‘blind studies’. This is where the children, the parents, and the experimenters do not know which child receives sugar, and which receives a placebo. A 1996 review of 12 blind studies found that sugar does not have any influence on a child’s behaviour. When children are in a naturally stimulating environment (such as at a birthday party or on holiday), this leads to their energetic behaviour, rather than the sugary food which happens to be served at these events.</td>
</tr>
<tr>
<td>Myth - 11</td>
<td>Fish oil is good for heart health.</td>
<td>Fish oil is not good for heart health.</td>
<td>In September 2012, researchers re-examined fish oil data from 20 clinical trials involving nearly 70 000 people. They found that compared to placebos, fish oil supplements were not associated with a lower risk of heart attack, stroke, or sudden cardiac death. At high doses, the side effects of fish oil include thinning of the blood, nausea and may also reduce immune system activity, hindering the body’s ability to fight infection. The side effects of small doses over long periods of time remain unknown.</td>
</tr>
<tr>
<td>Myth - 12</td>
<td>Due to epidemic vitamin D deficiency, more time in the sun would reduce health problems.</td>
<td>More time in the sun would not reduce health problems, even with epidemic vitamin D deficiency.</td>
<td>While some sun exposure is vital to good health, in most parts of Australia we only need a very small amount. During summer, adequate vitamin D levels are met by just a few minutes daily exposure to our face, arms and hands before 10am and after 3pm. The harms of sun exposure far outweigh the risks of vitamin D deficits. Our bodies can store vitamin D to last between 30 and 60 days, which means that your body can mobilize its own reserves if you do not receive sufficient amounts temporarily.</td>
</tr>
<tr>
<td>Item number</td>
<td>Items</td>
<td>Brief explanation</td>
<td>Detailed explanation</td>
</tr>
<tr>
<td>-------------</td>
<td>-------</td>
<td>-------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Myth - 13</td>
<td>Cancer screening is greatly beneficial.</td>
<td>Cancer screening is not greatly beneficial.</td>
<td>Cancer screenings generate many false alarms, as well as discover non-progressive cancers. Non-progressive cancers are abnormalities that meet the definition of cancer, but never progress to harm the patient in their lifetime. Any painful treatment received as a result of non-progressive cancers is needless. In the UK, mammography is estimated to save 1,300 lives p.a., yet an estimated 4,000 women suffer from unnecessary treatment and surgery. The survival rate for countries with publicly funded cancer screening, and those without, is the same.</td>
</tr>
<tr>
<td>Myth - 14</td>
<td>Antioxidants are important for health.</td>
<td>Antioxidants are not important for health.</td>
<td>It is clear that fruits and vegetables are good for you, but scientists are still debating why. Vegetables such as Brussels sprouts and broccoli (which have been linked with anti-cancer benefits) may actually generate these benefits through their ability to promote pro-oxidative cellular processes, rather than anti-oxidative ones. A study in 2009 found that physical exercise prevents type-2 diabetes, yet only if the participants did not consume daily doses of antioxidant vitamin C and E supplements.</td>
</tr>
<tr>
<td>Myth - 15</td>
<td>Ostriches hide their head in the sand when frightened.</td>
<td>Ostriches do not hide their head in the sand when frightened.</td>
<td>The myth that ostriches hide their heads in the sand was apparently first reported by the Roman historian Pliny the Elder, who also thought that ostriches could hatch their eggs by looking at them aggressively. No ostrich has ever been observed to bury its head in the sand, and if it did, it would likely suffocate. At times, it may look as though it is burying its head, when in fact they are swallowing sand and pebbles which help grind the food in its stomach, aiding digestion.</td>
</tr>
</tbody>
</table>
### Table 9 (continued)

<table>
<thead>
<tr>
<th>Item number</th>
<th>Items</th>
<th>Brief explanation</th>
<th>Detailed explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myth - 16</td>
<td>Bulls are mostly colour-blind, but can see the colour red vividly.</td>
<td>Bulls are not mostly colour-blind, nor can they see the colour red vividly.</td>
<td>The myth that bulls are infuriated by the colour red has been around since at least 1580, when a best-selling writer at the time, John Lyly, made reference to it. Bulls can see colour, but they only have two types of cones in their eyes, rather than three types like us humans. This means that they can see blues and greens, but ironically, not reds. Matadors hold red capes purely due to tradition; it is the movement of the bullfighter’s cape that causes it to charge.</td>
</tr>
<tr>
<td>Myth - 17</td>
<td>The song ‘ring a ring around a rosie’ is about the plague.</td>
<td>The song ‘ring a ring around a rosie’ is not about the plague.</td>
<td>The first time that the rhyme ‘ring a ring a rosie’ was associated with the plague was in 1961, in the book ‘The Plague and the Fire’. Until then, no-one had made the connection between the plague and the rhyme. There is some evidence to suggest that it dates back to 1790, but first known publication was in 1881, in ‘Mother Goose’. The earlier versions of the rhyme are far less fitting to plague interpretation, for example, some read ‘ring a ring a rosie, a bottle full of posie, all the girls in our town ring for little Josie’.</td>
</tr>
<tr>
<td>Myth - 18</td>
<td>Urine is an effective treatment for a jellyfish sting.</td>
<td>Urine is not an effective treatment for a jellyfish sting.</td>
<td>This myth originated from the television show ‘Friends’. When one of the characters is stung by a jellyfish, the other characters recall a useful fact that they saw on the discovery channel - that urine will remove the pain. Although this worked for the character in the show, in reality there is no scientific evidence to support this remedy. Urinating on the affected area may only aggravate the sting, not to mention making your day even worse than it already is.</td>
</tr>
</tbody>
</table>
## Table 9 (continued)

<table>
<thead>
<tr>
<th>Item number</th>
<th>Items</th>
<th>Brief explanation</th>
<th>Detailed explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myth - 19</td>
<td>Stress causes cancer.</td>
<td>Stress does not cause cancer.</td>
<td>Although stress is often considered to be linked to behaviours which may increase one’s chances of developing cancer (such as smoking, overeating, being physically inactive or drinking excessive alcohol), stress itself cannot cause cancer. People with a low stress lifestyle are equally as likely to be diagnosed with cancer as those with a high stress lifestyle. One Danish study followed 6689 women over 18 years, and found that women with high stress were actually less likely to develop breast cancer than women with low stress.</td>
</tr>
<tr>
<td>Myth - 20</td>
<td>Reading in dim light will eventually ruin your eyesight.</td>
<td>Reading in dim light will not eventually ruin your eyesight.</td>
<td>Our eyes are cleverly designed to adjust themselves in response to varying amount of light in our environment - but what if the light is just too dim? There is no evidence to suggest that reading in suboptimal light causes any lasting eye damage. It may feel uncomfortable after an extended period of time, yet the most serious outcome would be minor eye strain, or eye fatigue, which is temporary. Some people may also get a headache from squinting, yet this too is transient.</td>
</tr>
<tr>
<td>Fact - 1</td>
<td>The colour of a chicken’s egg is related to the chicken’s ear lobe.</td>
<td>The colour of a chicken’s egg is related to the chicken’s ear lobe.</td>
<td>Chicken breeds with white earlobes are likely to lay white eggs, and chickens with red earlobes are likely to lay brown eggs. Chickens’ earlobes are actually quite large, yet are covered with feathers so are rarely visible. The colour of the egg’s shell is caused by pigment deposition during the eggs formation. It is important to note that the earlobes do not have a direct impact upon the colour of the egg, but simply reflect trends in genetics, like people with black hair are more likely to have brown eyes.</td>
</tr>
<tr>
<td>Item number</td>
<td>Items</td>
<td>Brief explanation</td>
<td>Detailed explanation</td>
</tr>
<tr>
<td>-------------</td>
<td>-------</td>
<td>-------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Fact - 2</td>
<td>Stomach acid can dissolve razor blades.</td>
<td>Stomach acid can dissolve razor blades.</td>
<td>A study in 1997 confirmed that our gastric juices can indeed dissolve razor blades. This is possible due to simple chemistry. The lining of our stomach secretes hydrochloric acid, and razor blades are made of steel, which is an alloy of iron. Iron is readily dissolved by hydrochloric acid. The study concluded that the best time for surgery would be 15 hours or so after ingestion. This is because the blade has become fragile, and the retrieval is facilitated as the blade can be broken and removed piecemeal.</td>
</tr>
<tr>
<td>Fact - 3</td>
<td>Caucasians are often born with blue eyes which can subsequently change colour.</td>
<td>Caucasians are often born with blue eyes which can subsequently change colour.</td>
<td>A Caucasian baby is often born with blue eyes, which may potentially change colour as the baby ages. As the eye is exposed to sunlight, the iris begins to produce melanin (a naturally occurring pigment) and the eye colour gradually changes towards its adult colouring. The iris often changes colour within the first year of a child’s life, although may gradually change for up to 6 years. People from Asian or African decent are most likely born with dark eyes, which do not change colour over time.</td>
</tr>
<tr>
<td>Fact - 4</td>
<td>Some sleepwalkers can do amazingly complex tasks.</td>
<td>Some sleepwalkers can do amazingly complex tasks.</td>
<td>Sleepwalking occurs more often in children than adults, and can involve just about anything. Actions can include getting dressing, going on an outing, or cooking a complete meal. In 2010, the New York Times reported that Shirley Koecheler, a 54-year-old woman, had no clue why she had gained 45 kilograms until her husband told her she had been heading to the kitchen and binging on junk food. Sleepwalking is increasingly a difficult legal area, as people using the “I did it in my sleep” defence is on the rise.</td>
</tr>
<tr>
<td>Item number</td>
<td>Items</td>
<td>Brief explanation</td>
<td>Detailed explanation</td>
</tr>
<tr>
<td>-------------</td>
<td>----------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Fact - 5</td>
<td>It is not safe to talk on landline telephones when there is a thunderstorm.</td>
<td>It is not safe to talk on landline telephones when there is a thunderstorm.</td>
<td>The current in a lightning bolt can exceed 100 000 volts. Electrical wires are good transmitters of electricity, so when lightning strikes a house it has the potential to move through the interconnected cables. Usually, the energy is simply absorbed into the ground, but it is possible that the current travels through the landline’s cables and shocks the person on the end of the phone line. It is also wise to unplug electronic goods such as laptops, as the surge of electricity can severely damage them.</td>
</tr>
<tr>
<td>Fact - 6</td>
<td>Some body gestures are universally understood.</td>
<td>Some body gestures are universally understood.</td>
<td>The vast majority of body language is culturally specific. Our ‘OK’ gesture could mean ‘money’ in Japan, ‘zero’ in Finland, or a threatening gesture in some Arab cultures. However, there are a limited number of gestures which are universally used. For example, athletes from all cultures make the same posture when they win: arms raised in a V with the chin raised high. The same is true for athletes who have been blind from birth, suggesting that the victory pose is innate. The defeat posture, hunched shoulders, seems to be universal too.</td>
</tr>
<tr>
<td>Fact - 7</td>
<td>Dogs can smell cancer.</td>
<td>Dogs can smell cancer.</td>
<td>Dogs perform better than state-of-the-art screening tests at sniffing out people with lung and breast cancer. This is not as outrageous as it sounds. Cancer patients have traces of chemicals (like alkanes and benzene derivatives) in their breath, and dogs can detect chemicals in concentrations as small as a few parts per trillion. The University of California did the first double blind study, where dogs correctly detected 99% of lung cancer samples, and made a mistake with 1% healthy controls.</td>
</tr>
</tbody>
</table>
### Table 9 (continued)

<table>
<thead>
<tr>
<th>Item number</th>
<th>Items</th>
<th>Brief explanation</th>
<th>Detailed explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fact - 8</td>
<td>Humans can regrow the tips of fingers and toes after they have been amputated.</td>
<td>Humans can regrow the tips of fingers and toes after they have been amputated.</td>
<td>Astonishingly, humans have a very amphibian-like trait of being able to regenerate. Unfortunately, this is limited to the very tips of our fingers and toes. A study in 1970 found that if the individual was under the age of 10, they had a limited capability to even regrow bone. There are stem cells at the base of each nail, which aid ordinary nail growth as well as the ability to rebuild the digit tip after amputation. Interestingly, a regenerated finger will sometimes lack a fingerprint.</td>
</tr>
<tr>
<td>Fact - 9</td>
<td>We are taller in the morning than in the evening.</td>
<td>We are taller in the morning than in the evening.</td>
<td>We are taller in the mornings in comparison to the evenings due to the compression of our spine over the course of the day. When you are standing or sitting, there is pressure on the intervertebral discs, which causes water to be expelled. At night, when the spine is horizontal, water is reabsorbed by the disks. In 1935, DePukey measured 1216 participants between 5 and 90 years old, and found the average person was 1.5cm shorter in the evening in compared to the morning.</td>
</tr>
<tr>
<td>Fact - 10</td>
<td>An opera singer’s piercing voice can shatter glass.</td>
<td>An opera singer’s piercing voice can shatter glass.</td>
<td>Every piece of glass has a natural resonant frequency, which is the speed at which it will vibrate with a sound wave. Glass goblets, such as wine glasses, are especially resonant due to their shape. If you rub the rim of the glass continuously, the pitch that you hear is its natural frequency. If a person sings this note loud enough and long enough, the glass will shatter. Only well-made glass such as fine crystal is fragile enough to break at volumes that people can produce without amplification.</td>
</tr>
</tbody>
</table>
### Table 9 (continued)

<table>
<thead>
<tr>
<th>Item number</th>
<th>Items</th>
<th>Brief explanation</th>
<th>Detailed explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fact - 11</td>
<td>Too many carrots will turn your skin orange.</td>
<td>Too many carrots will turn your skin orange.</td>
<td>Eating foods high in beta-carotene can indeed cause a discoloration of the skin. The medical term is ‘carotenaemia’, and it is a benign condition which mostly affects young children and vegetarians. It is most easily noticeable on people with a naturally fair complexion, particularly on the palms of their hands or the soles of their feet. Unlike jaundice, it does not cause discoloration of the whites of the eyes. Treatment involves simply not eating carrots for an extended period.</td>
</tr>
<tr>
<td>Fact - 12</td>
<td>The hummingbird is the only bird that can fly backwards.</td>
<td>The hummingbird is the only bird that can fly backwards.</td>
<td>Most birds simply flap their wings up and down. Hummingbirds have a much greater wing flexibility which enables them not only to fly forwards, but also backwards, sideward and straight up in the air. The uniqueness of the hummingbird has fascinated bird scientists for decades, but a new ultra slow camera has been able to capture the birds also rotating their body as they fly. At 55 rotations per second, they are the fastest moving vertebrate (relative to their body size) in the world.</td>
</tr>
<tr>
<td>Fact - 13</td>
<td>Yawning is contagious.</td>
<td>Yawning is contagious.</td>
<td>When human adults are shown videos of people yawning, many begin to yawn themselves. In fact, it is so contagious that simply thinking about it (or reading about it), is sufficient to initiate a yawn. You may be tempted to yawn right now whilst reading this passage. Research has shown that other primates also find yawning contagious. The reason to why we do this is still being debated. One suggestion is that it plays a role in empathy, whilst others think that it may have evolved as a social cue to synchronize sleep amongst a group.</td>
</tr>
<tr>
<td>Item number</td>
<td>Items</td>
<td>Brief explanation</td>
<td>Detailed explanation</td>
</tr>
<tr>
<td>-------------</td>
<td>----------------------------------------------------------------------</td>
<td>------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Fact - 14</td>
<td>The whip can break the sound barrier.</td>
<td>The whip can break the sound barrier.</td>
<td>Whips were invented 7000 years ago in China, but it wasn’t until the invention of high speed photography in 1927 that the ‘crack’ of the whip was seen to be a mini sonic boom, and not the leather hitting the handle. A sonic boom is the sound associated with shock waves created by an object travelling through the air faster than the speed of sound. Physicists say that the whip’s tip travels at 30 times the initial speed of the movement made by the individual holding the whip.</td>
</tr>
<tr>
<td>Fact - 15</td>
<td>Dogs shouldn’t eat chocolate.</td>
<td>Dogs shouldn’t eat chocolate.</td>
<td>Canines can experience dangerous effects from eating chocolate, and in some cases, it could be lethal. Chocolate contains a compound called methylxanthine. Whilst this acts as a relaxant in humans, it is processed more slowly in a dog’s body, and can affect their heart, central nervous system and kidneys. This can result in vomiting, cardiac arrhythmia and seizures. Treatment and whether a dog will need its stomach pumped will depend upon the size of the dog, and the type of chocolate ingested.</td>
</tr>
<tr>
<td>Fact - 16</td>
<td>Laughing regularly helps improve vascular function.</td>
<td>Laughing regularly helps improve vascular function.</td>
<td>It is well known that laughter reduces stress hormones and releases endorphins, yet strangely enough, it also has a positive impact on vascular function. A 2009 study found that people with heart disease were 40% less likely to laugh in a variety of situations compared to people without heart disease. A study in 2010 demonstrated the short-term benefits of laughter by showing participants either a 20-minute clip of a comedy or a documentary. Laughter led to tissue dilation in the inner lining of blood vessels, which increased blood flow.</td>
</tr>
</tbody>
</table>
Table 9 (continued)

<table>
<thead>
<tr>
<th>Item number</th>
<th>Items</th>
<th>Brief explanation</th>
<th>Detailed explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fact - 17</td>
<td>It is possible to hypnotize animals.</td>
<td>It is possible to hypnotize animals.</td>
<td>Interestingly, some animals can seem to be hypnotized – a state more commonly referred to as <em>tonic immobility</em>. It is thought to be an involuntary reflex that an animal enters in response to a threat. For example, many species of shark become immobile if you turn them upside down. Chickens may be ‘hypnotized’ if you draw a line from their beak across the ground, and lobsters can be put into this state if you stroke their head. This plays a role in survival to help the animal not be noticed, or ‘play dead’.</td>
</tr>
<tr>
<td>Fact - 18</td>
<td>The plague still exists in some parts of the world today, including the USA.</td>
<td>The plague still exists in some parts of the world today, including the USA.</td>
<td>Although outbreaks of the plague have declined since the early 20th Century, it still exists in various parts of the world, including the United States of America. The most commonly affected states include New Mexico, Arizona and Colorado. If correct antibiotics are given in time, most people survive. The World Health Organization reports that there are between 1000 and 2000 cases of the plague worldwide each year. There was a particularly large outbreak in Africa in 2003, where 2100 cases were reported and 180 people died.</td>
</tr>
<tr>
<td>Fact - 19</td>
<td>A cockroach can live for over a week without its head.</td>
<td>A cockroach can live for over a week without its head.</td>
<td>A cockroach’s body works very differently to that of a human, which allows it to survive after it has been decapitated. Firstly, a cockroach does not breathe through its mouth, but through ‘spiracles’, which are little holes in its body segment. Secondly, its nervous system is located throughout their body, so do not need their brain to control all bodily functions. Thirdly, they can simply seal off their neck by clotting, like we can to a minor cut. Lastly, they need much less food than a human, but will eventually die due to starvation.</td>
</tr>
</tbody>
</table>
### Table 9 (continued)

<table>
<thead>
<tr>
<th>Item number</th>
<th>Items</th>
<th>Brief explanation</th>
<th>Detailed explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fact - 20</td>
<td>Approximately 95% of the ocean remains unexplored.</td>
<td>Approximately 95% of the ocean remains unexplored.</td>
<td>Although the ocean covers more than 70% of the planet’s surface, 95% of the ocean remains unexplored. This not only means it is unseen by human eyes, but that it is still yet to be mapped. This is partly due to logistics. Not only is the ocean exceedingly vast, but it is up to 11 kilometres deep (as far as we know). In deep sea, there is no light, crushing pressure, and freezing temperatures, and we have only recently developed technology that can withstand this harsh environment.</td>
</tr>
<tr>
<td>Item number</td>
<td>Items</td>
<td>Inference question 1</td>
<td>Inference question 2</td>
</tr>
<tr>
<td>-------------</td>
<td>----------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Myth - 1</td>
<td>The daddy long legs spider is extremely venomous, but fortunately its fangs are too small to pierce human skin.</td>
<td>On a scale of 0-10, how venomous do you think the daddy long legs spider is? (0-10)</td>
<td>Scientists believe it is quite likely that the daddy long legs spider will develop larger fangs due to a genetic mutation: do you agree that health officials should fund projects to develop an anti-venom? (0 Strongly disagree–10 Strongly agree)</td>
</tr>
<tr>
<td>Myth - 2</td>
<td>A tooth partially dissolves if you leave it in cola overnight.</td>
<td>If a tooth is left in cola, what percentage has dissolved in 24 hours? (0-100 %)</td>
<td>What percentage of tooth enamel could be lost in a lifetime by daily consumption of cola (compared to daily consumption of a different soft drink)? (0-100 %)</td>
</tr>
<tr>
<td>Myth - 3</td>
<td>Liars sometimes give themselves away by physical ‘tells’ such as looking to the right or not looking you in the eye.</td>
<td>What percentage of lies can FBI detectives catch just by looking at physical tells? (0-100 %)</td>
<td>How important do you think it would be for police officers, judges, etc. to be trained on lie detection techniques (such as identifying physical tells)? (0 Not important at all–10 Extremely important)</td>
</tr>
<tr>
<td>Myth - 4</td>
<td>Females, on average, talk more than males.</td>
<td>On average, what percentage do females spend talking over and above males? (0-50 %)</td>
<td>Out of 100 couples in marriage counselling, how many have the issue that the man does not talk enough and/or the woman talks too much? (0-100)</td>
</tr>
<tr>
<td>Myth - 5</td>
<td>Carrots can improve eyesight.</td>
<td>If you ate 20 carrots a day for 3 months, how much would your eyesight improve? (0-20 %)</td>
<td>By feeding carrots to a baby regularly you decrease the risk of impaired vision by x% (0-20 %)</td>
</tr>
<tr>
<td>Item number</td>
<td>Items</td>
<td>Inference question 1</td>
<td>Inference question 2</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Myth - 6</td>
<td>Playing Mozart to your baby will boost its IQ.</td>
<td>Listening to Mozart every evening for 3 years will increase a child’s IQ by what percentage? (0-10 %)</td>
<td>If one twin listened to Mozart every night for the first 10 years of their life, and another twin was not exposed to classical music at all, would you expect them to have different IQs? (0 Definitely not – 10 Definitely)</td>
</tr>
<tr>
<td>Myth - 7</td>
<td>Cracking knuckles can cause arthritis.</td>
<td>Should children be warned about the risk of cracking joints in an attempt to lower the rate of arthritis in future generations? (0 Strongly disagree – 10 Strongly agree)</td>
<td>By how much can you increase your risk of arthritis if you crack your knuckles? (0-20 %)</td>
</tr>
<tr>
<td>Myth - 8</td>
<td>Using a defibrillator is the best way to restart a stopped heart.</td>
<td>If a man had no pulse, what is the likelihood that you would witness the paramedic reach for his defibrillator? (0 Extremely unlikely – 10 Extremely likely)</td>
<td>What is the success rate of paramedics using a defibrillator on a stopped heart? (0-100 %)</td>
</tr>
<tr>
<td>Myth - 9</td>
<td>Mercury in vaccines can cause harm.</td>
<td>What percentage of side effects could be avoided by banning mercury in vaccinations? (0-100 %)</td>
<td>What would you be willing to pay extra for mercury-free vaccine? (0-20 %)</td>
</tr>
<tr>
<td>Myth - 10</td>
<td>Giving sugar to children makes them hyperactive.</td>
<td>How much more hyperactive would a five-year-old act after they have been given a high dose of sugar? (0-100 %)</td>
<td>How much could you avoid hyperactivity by not eating sugar? (0-100 %)</td>
</tr>
<tr>
<td>Item number</td>
<td>Items</td>
<td>Inference question 1</td>
<td>Inference question 2</td>
</tr>
<tr>
<td>-------------</td>
<td>----------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Myth - 12</td>
<td>Due to epidemic vitamin D deficiency, more time in the sun would reduce health problems.</td>
<td>Would you encourage your friends to increase the amount of time they spend outside to promote vitamin D production? (0 Definitely not – 10 Definitely)</td>
<td>When buying sun protection products, how much would you be willing to pay extra for products that do not impact as much on vitamin D production? (0-20 %)</td>
</tr>
<tr>
<td>Myth - 11</td>
<td>Fish oil is good for heart health.</td>
<td>If a person has a 50% risk of a heart attack, by how much could the risk be reduced by regular fish oil intake? (0-20%)</td>
<td>Should the government subsidise and promote regular intake of fish oil supplements? (0 Definitely not – 10 Definitely)</td>
</tr>
<tr>
<td>Myth - 13</td>
<td>Cancer screening is greatly beneficial.</td>
<td>How would you rate the long-term benefits associated with widespread cancer screening? (0 Very low – 10 Very high)</td>
<td>Should governments subsidize and promote cancer screenings for all individuals over the age of 40? (0 Definitely not – 10 Definitely)</td>
</tr>
<tr>
<td>Myth - 14</td>
<td>Antioxidants are important for health.</td>
<td>When buying groceries, how much more are you willing to pay for items that are rich in antioxidants? (0-20 %)</td>
<td>How much do you think the risk of disease can be reduced by regular intake of antioxidants? (0-20 %)</td>
</tr>
<tr>
<td>Myth - 15</td>
<td>Ostriches hide their head in the sand when frightened.</td>
<td>If a lion breaks into the ostrich’s enclosure at the zoo, what is the likelihood that the ostrich will hide its head in the sand? (0 Extremely unlikely – 10 Extremely likely)</td>
<td>Should David Attenborough include a segment in his next documentary regarding why ostriches hide their heads in the sand? (0 Strongly disagree – 10 Strongly agree)</td>
</tr>
</tbody>
</table>
Table 10 (continued)

<table>
<thead>
<tr>
<th>Item number</th>
<th>Items</th>
<th>Inference question 1</th>
<th>Inference question 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myth - 16</td>
<td>Bulls are mostly colour-blind, but can see the colour red vividly.</td>
<td>What is the likelihood that a bull would charge at a blue cape? (0 Extremely unlikely – 10 Extremely likely)</td>
<td>By how much would wearing a red jumper increase your risk of being attacked by a bull during the famous Spanish “Running of the Bulls” events? (0-100 %)</td>
</tr>
<tr>
<td>Myth - 17</td>
<td>The song ‘ring a ring a rosie’ is about the plague.</td>
<td>What is the likelihood that 'ring a ring a rosie’ was being sung during the Great Plague of London in 1665? (0 Extremely unlikely – 10 Extremely likely)</td>
<td>What is the likelihood that the line 'a pocket full of posies’ from the nursery rhyme ring a ring a rosie, refers to people carrying sachets of herbs in an attempt to ward off the plague? (0 Extremely unlikely – 10 Extremely likely)</td>
</tr>
<tr>
<td>Myth - 18</td>
<td>Urine is an effective treatment for a jellyfish sting.</td>
<td>If your friend was stung by a jellyfish, how likely are you to recommend urine as a home remedy? (0 Extremely unlikely – 10 Extremely likely)</td>
<td>Should there be more funding for research isolating which exact properties of urine might be responsible for sting relief? (0 Definitely – 10 Definitely not)</td>
</tr>
<tr>
<td>Myth - 19</td>
<td>Stress causes cancer.</td>
<td>How much could cancer risk be increased by extreme stress? (0-20%)</td>
<td>Should cancer prevention programs include a stress prevention management module? (0 Strongly disagree – 10 Strongly agree)</td>
</tr>
<tr>
<td>Myth - 20</td>
<td>Reading in dim light will eventually ruin your eyesight.</td>
<td>If you spend one hour reading in dim light every day for several years, what is the likelihood that you have permanently damaged your eyesight? (0 Extremely unlikely – 10 Extremely likely)</td>
<td>If you saw a family member reading in dim light, how likely is it that you would encourage them to turn on a light? (0 Extremely unlikely – 10 Extremely likely)</td>
</tr>
<tr>
<td>Item number</td>
<td>Items</td>
<td>Inference question 1</td>
<td>Inference question 2</td>
</tr>
<tr>
<td>-------------</td>
<td>----------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Fact - 1</td>
<td>The colour of a chicken’s egg is related to the chicken’s ear lobe.</td>
<td>How likely is it that a chicken with white earlobes lays a white (not brown) egg?</td>
<td>If you bought a box of white eggs, what is the likelihood that the chickens that laid them have white earlobes?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0 Extremely unlikely – 10 Extremely likely)</td>
<td>(0 Extremely unlikely – 10 Extremely likely)</td>
</tr>
<tr>
<td>Fact - 2</td>
<td>Stomach acid can dissolve razor blades.</td>
<td>After 48 hours in the stomach, how much of a razor blade remains?</td>
<td>Teaching teenagers that our stomach acid can dissolve razor blades would be an accurate and entertaining way to inform them about chemistry.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0-100 %)</td>
<td>(0 Strongly disagree – 10 Strongly agree).</td>
</tr>
<tr>
<td>Fact - 3</td>
<td>Caucascians are often born with blue eyes which can subsequently change colour.</td>
<td>What percentage of children that are born with blue eyes, end up with brown eyes?</td>
<td>If a friend of yours is disappointed that her newborn baby has blue eyes, would you inform them that there is a chance that they could change colour?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0-40 %)</td>
<td>(0 Definitely not – 10 Definitely)</td>
</tr>
<tr>
<td>Fact - 4</td>
<td>Some sleepwalkers can do amazingly complex tasks.</td>
<td>How likely do you think it is that a sleepwalker can drive a car in their sleep?</td>
<td>Would you believe your friend if they said they woke up eating a freshly made sandwich?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0 Extremely unlikely – 10 Extremely likely)</td>
<td>(0 Definitely not – 10 Definitely)</td>
</tr>
<tr>
<td>Fact - 5</td>
<td>It is not safe to talk on landline telephones when there is a thunderstorm.</td>
<td>Next time there is a thunderstorm, how likely are you to avoid using a landline phone?</td>
<td>Should there be an education program for primary school children, teaching them about the dangers of speaking on the phone during a thunderstorm?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0 Extremely unlikely – 10 Extremely likely)</td>
<td>(0 Definitely not – 10 Definitely)</td>
</tr>
<tr>
<td>Item number</td>
<td>Items</td>
<td>Inference question 1</td>
<td>Inference question 2</td>
</tr>
<tr>
<td>-------------</td>
<td>-------</td>
<td>----------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Fact - 6</td>
<td>Some body gestures are universally understood.</td>
<td>What is the likelihood that an indigenous person from Ecuador would interpret a smile as being happy? (0 Extremely unlikely – 10 Extremely likely)</td>
<td>Even if two people do not share a language, a basic level of communication is possible due to the universality of some gestures. Do you agree? (0 Definitely not – 10 Definitely)</td>
</tr>
<tr>
<td>Fact - 7</td>
<td>Dogs can smell cancer.</td>
<td>How much extra funding should go towards programs supporting training for cancer sniffing dogs? (0-100 %)</td>
<td>Should dog kennels for cancer sniffing dogs be included in the design of new hospital buildings? (0 Definitely not – 10 Definitely)</td>
</tr>
<tr>
<td>Fact - 8</td>
<td>Humans can regrow the tips of fingers and toes after they have been amputated.</td>
<td>What proportion of fingers will regenerate after the tip has been amputated? (0-100 %)</td>
<td>How much extra funding should go towards investigating why tips of fingers and toes have the ability to regenerate? (0-100 %)</td>
</tr>
<tr>
<td>Fact - 9</td>
<td>We are taller in the morning than in the evening.</td>
<td>What is the likelihood that you are taller at breakfast time than at dinner time? (0 Extremely unlikely – 10 Extremely likely)</td>
<td>If you are a few millimeters too short to go on a rollercoaster in the evening, how likely is it that you will be allowed to ride the following morning? (0 Extremely unlikely – 10 Extremely likely).</td>
</tr>
<tr>
<td>Fact - 10</td>
<td>An opera singer’s piercing voice can shatter glass.</td>
<td>How likely is it that the world’s top opera singer can break a wine glass using only her voice? (0 Extremely unlikely – 10 Extremely likely)</td>
<td>Should catering companies think twice about bringing delicate glassware to the opera? (0 Definitely not – 10 Definitely)</td>
</tr>
<tr>
<td>Item number</td>
<td>Items</td>
<td>Inference question 1</td>
<td>Inference question 2</td>
</tr>
<tr>
<td>-------------</td>
<td>-------</td>
<td>----------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Fact - 11</td>
<td>Too many carrots will turn your skin orange.</td>
<td>If you eat 5 carrots a day, how likely is it that your skin will turn orange?</td>
<td>If you saw a person on the street who was suspiciously orange, what is the likelihood that too many carrots are to blame?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0 Extremely unlikely – 10 Extremely likely)</td>
<td>(0 Extremely unlikely – 10 Extremely likely)</td>
</tr>
<tr>
<td>Fact - 12</td>
<td>The hummingbird is the only bird that can fly backwards.</td>
<td>How many species of bird, apart from the hummingbird, can fly backwards?</td>
<td>Do you think that the hummingbird is unique due to its aerodynamic abilities?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0-20 Species of bird)</td>
<td>(0 Definitely not – 10 Definitely)</td>
</tr>
<tr>
<td>Fact - 13</td>
<td>Yawning is contagious.</td>
<td>If your friend yawns, what is the likelihood that you will yawn too?</td>
<td>If there are 100 people in a room, and the presenter yawns, how many of the audience members yawn?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0 Extremely unlikely – 10 Extremely likely)</td>
<td>(0-100 Number of people)</td>
</tr>
<tr>
<td>Fact - 14</td>
<td>The whip can break the sound barrier.</td>
<td>If we gave you a whip, how likely do you think it is that you would be able to break the sound barrier?</td>
<td>Do you think that the whip breaking the sound barrier would be an accurate and engaging way to teach primary school children about physics?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0 Extremely unlikely – 10 Extremely likely)</td>
<td>(0 Definitely not – 10 Definitely)</td>
</tr>
<tr>
<td>Fact - 15</td>
<td>Dogs shouldn’t eat chocolate.</td>
<td>If your neighbor’s Beagle had eaten a block of chocolate, would you recommend they see a vet?</td>
<td>Should purposefully giving chocolate to dogs be considered cruelty to animals?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0 Definitely not – 10 Definitely)</td>
<td>(0 Definitely not – 10 Definitely)</td>
</tr>
<tr>
<td>Item number</td>
<td>Items</td>
<td>Inference question 1</td>
<td>Inference question 2</td>
</tr>
<tr>
<td>------------</td>
<td>----------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Fact - 16</td>
<td>Laughing regularly helps improve vascular function.</td>
<td>How much can the risk for a heart attack be lowered by laughing regularly?</td>
<td>Should there be laughing seminars held as part of the rehabilitation programs in cardiovascular hospital wards?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0-20 %)</td>
<td>(0 Definitely not – 10 Definitely)</td>
</tr>
<tr>
<td>Fact - 17</td>
<td>It is possible to hypnotize animals.</td>
<td>If we gave you an animal, and instructions on how to hypnotize it, what is the likelihood that it would work?</td>
<td>If you saw a show where an animal was hypnotized, would you assume it to be true hypnosis rather than simply a well trained animal?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0 Extremely unlikely – 10 Extremely likely)</td>
<td>(0 Definitely not – 10 Definitely)</td>
</tr>
<tr>
<td>Fact - 18</td>
<td>The plague still exists in some parts of the world today, including the USA.</td>
<td>There is a global initiative to eradicate polio. Should there be an initiative of equal vigour to eradicate the plague?</td>
<td>What is the likelihood that you would check with your travel doctor that the countries you are visiting next month are free of the plague?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0 Definitely not – 10 Definitely)</td>
<td>(0 Extremely unlikely – 10 Extremely likely)</td>
</tr>
<tr>
<td>Fact - 19</td>
<td>A cockroach can live for over a week without its head.</td>
<td>What proportion of cockroaches live after they have been decapitated?</td>
<td>Do you think that medical research should investigate the resilience of the cockroach's nervous system?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0-100 %)</td>
<td>(0 Definitely not – 10 definitely)</td>
</tr>
<tr>
<td>Fact - 20</td>
<td>Approximately 95% of the ocean remains unexplored.</td>
<td>There should be more funding for exploration of the oceans, over and above the exploration of space.</td>
<td>How likely is it that only a fragment of species inhabiting the oceans are known to mankind?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0 Strongly disagree – 10 Strongly agree).</td>
<td>(0 Extremely unlikely – 10 Extremely likely)</td>
</tr>
</tbody>
</table>
Appendix B: Chapter 4 Stimuli
<table>
<thead>
<tr>
<th>Item</th>
<th>Claim</th>
<th>Explanation</th>
<th>Inference Question</th>
<th>Facts only claim</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Brain Myth - 1</td>
<td>Most people only use between 10 and 50% of their brains.</td>
<td>We use all of our brain: Brain imaging techniques have demonstrated that our whole brain is active, at least to some extent, all of the time. Specific areas of the brain will become more active depending on the demands of the task. The brain, like all other organs, has been shaped by natural selection. It would be extremely costly and inefficient for our body to produce material that it was not going to use.</td>
<td>What percentage of our brain do we use? (0-100%)</td>
<td>People use 100% of their brains.</td>
</tr>
<tr>
<td>The Brain Myth - 2</td>
<td>Logical people are left-brained, and creative people are right-brained.</td>
<td>All brains are remarkably similar: A study in 2013 scanned the brains of over 1000 people, and found that all people use both sides of their brain to the same extent, regardless of how creative or logical they were. Some brain functions occur in one side of the brain—for example, language tends to be on the left—yet this is true for virtually everyone.</td>
<td>If your friend was extremely creative, would you assume her brain was structured differently to an individual who was more logical? (0-10 Definitely not - Definitely)</td>
<td>Even though some people are more creative and other people are more logical—people’s brains are remarkably similar.</td>
</tr>
<tr>
<td>The Brain Myth - 3</td>
<td>Pregnancy affects brain functioning.</td>
<td>Brain functioning is unaffected by pregnancy: A study in the British Journal of Psychiatry assessed the cognitive abilities of over 1000 women, and found no differences between pregnant and non-pregnant women. Many pregnancy guides inform women that they will be abnormally confused or forgetful, so it is possible that normal lapses in memory or thinking are simply misattributed to being pregnant.</td>
<td>If a pregnant friend misplaced her keys, would you attribute it to her being pregnant? (0-10 Definitely not - Definitely)</td>
<td>Brain functioning is unaffected by pregnancy.</td>
</tr>
<tr>
<td>Item</td>
<td>Claim</td>
<td>Explanation</td>
<td>Inference Question</td>
<td>Facts only claim</td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
<td>-------------</td>
<td>--------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>The Brain Fact - 1</td>
<td>Your brain itself feels no pain.</td>
<td>The brain feels no pain: Pain from headaches mostly stems from outside the skull, perhaps from muscle pain in your head or neck or when blood vessels swell and press on nerves. Brain surgery can be performed with only a local anaesthetic to the scalp, as the brain has no pain receptors.</td>
<td>If you have a headache, how likely is it that the pain is stemming from your brain? (0-10 Extremely likely – Extremely unlikely)</td>
<td>Your brain itself feels no pain.</td>
</tr>
<tr>
<td>The Brain Fact - 2</td>
<td>The brain is two-thirds fat.</td>
<td>The brain is mostly fat: There are approximately 90 billion neurons in the brain. These neurons are wrapped in myelin, a fat which acts as an electrical insulator. The myelin appears like a string of sausages covering the length of a neuron, allowing the nerve impulses to jump from gap to gap, making the impulses much quicker.</td>
<td>Is the fact that the brain is made up of two thirds fat an accurate way to teach children about biology? (0-10 Definitely not - Definitely)</td>
<td>The brain is two-thirds fat.</td>
</tr>
<tr>
<td>The Brain Fact - 3</td>
<td>Brain activity accounts for 20% of our resting metabolic rate.</td>
<td>The brain needs a lot of energy: In 1953, Louis Sokoloff found that the brain consumed the same energy while resting and while doing arithmetic, so the brain uses a lot of energy regardless of what we are doing. Although the brain is only about 2% of the body’s total weight, brain activity accounts for 20% of our resting metabolic rate.</td>
<td>If our heart accounts for 10% of our resting metabolic rate, how much does our brain account for? (0-20%)</td>
<td>Brain activity accounts for 20% of our resting metabolic rate.</td>
</tr>
<tr>
<td>Item</td>
<td>Claim</td>
<td>Explanation</td>
<td>Inference Question</td>
<td>Fact only claim</td>
</tr>
<tr>
<td>---------------</td>
<td>------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Alcohol</td>
<td>If you cook with alcohol, it will ‘cook off’ and the dish will become non-alcoholic.</td>
<td>Alcohol remains in food: A study in 1992 found that if alcohol was stirred into a dish, 25% of the alcohol remained after a whole hour of cooking. Many people believe that it is easy to cook out the alcohol due to alcohol’s low boiling point. Yet even after setting alcohol alight to flambé a dish, a whopping 75% of alcohol remains.</td>
<td>Would you give Christmas pudding flambéed with brandy to a four year old child? (0-10 Definitely not - Definitely)</td>
<td>Dishes with alcohol will remain alcoholic, even after cooking them.</td>
</tr>
<tr>
<td>Myth - 1</td>
<td>Alcohol promotes sleep.</td>
<td>Alcohol disturbs sleep: Drinking alcohol before bed leads to REM sleep being disrupted. This is followed by abnormally shallow sleep, causing multiple awakenings. The more alcohol consumed prior to sleep, the more pronounced these effects are. So although alcohol may help the onset of sleep, sleep quality is adversely affected.</td>
<td>If your insomniac friend told you they were planning on drinking two glasses of wine before bed to help them sleep, would you advise them otherwise? (0-10 Definitely not - Definitely)</td>
<td>Sleep is adversely affected by alcohol.</td>
</tr>
<tr>
<td>Alcohol</td>
<td>Alcohol kills brain cells.</td>
<td>Brain cells survive alcohol: Neurons are the longest living cells in the body, and they are very difficult to kill. Alcoholics and non-alcoholics have the same number and density of neurons in the brain. Alcohol only temporarily affects brain functioning.</td>
<td>Do you think that ‘alcohol kills brain cells’ is an accurate way to teach high school children about the negative effects of alcohol (0-10 Definitely not - Definitely)</td>
<td>Alcohol negatively impacts brain functioning, but not due to cell death.</td>
</tr>
<tr>
<td>Myth - 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myth - 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item</td>
<td>Claim</td>
<td>Explanation</td>
<td>Inference Question</td>
<td>Fact only claim</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Alcohol Fact 1</td>
<td>Men and women process alcohol differently.</td>
<td>Alcohol gender differences: ‘Alcohol dehydrogenase’ is an enzyme in the stomach lining which helps break down alcohol. Women possess significantly less of this enzyme than men, which means that even if a man and a woman have equivalent body sizes, the women will feel the effects of alcohol sooner and to a greater extent.</td>
<td>Your friends Ben and Emma are having a drinking competition, shot for shot. They both weigh 60 kilograms. How likely is it that Emma will fall of her chair first? (0-10 Extremely unlikely – Extremely likely)</td>
<td>Men and women process alcohol differently.</td>
</tr>
<tr>
<td>Alcohol Fact 2</td>
<td>Mixing alcohol with sugary drinks makes you more drunk.</td>
<td>Diet soda increases the effects of alcohol: A recent study in the journal Alcoholism: Clinical and Experimental Research found that alcohol mixed with diet soda led to participants being more intoxicated than if they consumed the same beverage was mixed with full sugar soda. While participants reported that they did not feel any difference, those drinking diet mixers had significantly slower reaction times and their breathalyser readings were elevated by 18%.</td>
<td>If you wanted to have a few drinks with friends prior to driving, would you choose to mix your drink with diet soda? (0-10 Definitely not - Definitely)</td>
<td>Mixing alcohol with sugary drinks makes you more drunk.</td>
</tr>
<tr>
<td>Alcohol Fact 3</td>
<td>Alcohol improves your sense of smell.</td>
<td>Alcohol can improve sense of smell: The brain automatically dampens our sense of smell, and if these inhibitory signals to the brain are affected, our sense of smell can be enhanced. Consuming alcohol is one example of how these inhibitory signals can be alleviated. A study in the journal of Behavioural Brain Research found that people who had consumed low amounts of alcohol could smell with greater accuracy than those who were sober.</td>
<td>Would a wine expert’s sense of smell improve after drinking the first few tastings? (0-10 Definitely not - Definitely)</td>
<td>Alcohol improves your sense of smell.</td>
</tr>
<tr>
<td>Item</td>
<td>Claim</td>
<td>Explanation</td>
<td>Inference Question</td>
<td>Fact only claim</td>
</tr>
<tr>
<td>------------</td>
<td>----------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Animals Myth -1</td>
<td>The daddy long legs spider is extremely venomous, but fortunately its fangs are too small to pierce human skin</td>
<td>Daddy long-legs are only mildly venomous: While the fangs of the daddy long legs spider are indeed very small (estimated at 0.25mm in length), they are capable of piercing human skin. Tests have revealed that the venom of the daddy long legs spider is no more powerful than that of most other spiders, and not dangerous to human health.</td>
<td>Scientists believe it is quite likely that the daddy long legs spider will develop larger fangs due to a genetic mutation: do you agree that health officials should fund projects to develop antivenom? (0 strongly disagree, 10 strongly agree)</td>
<td>While the fangs of the daddy long legs spider can indeed pierce human skin, the venom is only mildly potent.</td>
</tr>
<tr>
<td>Animals Myth -2</td>
<td>Urine is an effective treatment for a jellyfish sting.</td>
<td>Urine aggravates jellyfish sting: Urinating on the area affected by a jellyfish sting will only aggravate the sting, not to mention making your day even worse than it already is. While this “home remedy” worked for a character on the television show ‘Friends’, in reality there is no scientific evidence to support it.</td>
<td>Should there be more funding for research isolating which exact properties of urine might be responsible for sting relief? (0-10 Definitely not - Definitely)</td>
<td>Urine will only aggravate a jellyfish sting.</td>
</tr>
<tr>
<td>Animals Myth -3</td>
<td>Bulls are mostly colour-blind, but can see the colour red vividly.</td>
<td>Bulls see mostly blue and yellow: Matadors hold red capes purely due to tradition; it is the movement of the bullfighter’s cape that causes the bull to charge. Bulls can see colour, but they only have two types of cones in their eyes, rather than three types like us humans. This means that they can see blues and yellows, but ironically, not reds.</td>
<td>What is the likelihood that a bull would charge at a matador brandishing a blue cape? (0: Extremely unlikely – 10 Extremely likely)</td>
<td>Bulls can see blues and greens, but not reds.</td>
</tr>
<tr>
<td>Item</td>
<td>Claim</td>
<td>Explanation</td>
<td>Inference Question</td>
<td>Fact only claim</td>
</tr>
<tr>
<td>--------------------</td>
<td>------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Animals Fact - 1</td>
<td>The colour of a chicken’s egg is related to the chicken’s ear lobe.</td>
<td>Egg colour is related to chicken’s earlobes: Chickens’ earlobes are quite large, yet are covered with feathers so are rarely visible. The earlobes do not have a direct impact upon the colour of the egg, but reflect trends in genetics, like people with black hair being more likely to have brown eyes. Chicken breeds with white earlobes are likely to lay white eggs, and chickens with red earlobes are likely to lay brown eggs.</td>
<td>If you bought a box of white eggs, what is the likelihood that the chickens that laid them have white earlobes? (0: Extremely unlikely – 10 Extremely likely)</td>
<td>The colour of a chicken’s egg is related to the chicken’s ear lobe.</td>
</tr>
<tr>
<td>Animals Fact - 2</td>
<td>Dogs can smell cancer.</td>
<td>Dogs smell cancer: Cancer patients have traces of chemicals (like alkanes and benzene derivatives) in their breath, and dogs can detect chemicals in concentrations as small as a few parts per trillion. The University of California did the first double blind study, where dogs correctly detected 99% of lung cancer samples, and made a mistake with only 1% of healthy controls.</td>
<td>How much extra funding should go towards programs supporting training for cancer sniffing dogs? (0-100%)</td>
<td>Dogs can smell cancer.</td>
</tr>
<tr>
<td>Animals Fact - 3</td>
<td>A cockroach can live for over a week without its head</td>
<td>Cockroaches survive headless: A cockroach’s body works very differently to that of a human, which allows it to survive after it has been decapitated. A cockroach does not breathe through its mouth, but through ‘spiracles’, which are little holes in its body segment. Additionally, its nervous system is located throughout their body, so they do not need their brain to control all bodily functions. Lastly, they need much less food than a human.</td>
<td>What proportion of cockroaches live after they have been decapitated? (0-100%)</td>
<td>A cockroach can live for over a week without its head</td>
</tr>
<tr>
<td>Item</td>
<td>Claim</td>
<td>Explanation</td>
<td>Inference Question</td>
<td>Fact only claim</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td>Flu vaccine</td>
<td>The side effects of the flu vaccine</td>
<td>Flu is worse than vaccine side effects: The worst side effect you're likely to get with injectable vaccine is a sore arm. The nasal mist flu vaccine might cause nasal congestion, runny nose, sore throat and cough. The risk of a rare allergic reaction is far less than the risk of severe complications from influenza.</td>
<td>Would concerns regarding the side effects of the flu vaccine be sufficient to deter you from getting vaccinated? (0-10 Definitely not - Definitely)</td>
<td>Flu is much worse than any side effects of the flu vaccine.</td>
</tr>
<tr>
<td>Myth - 1</td>
<td>are worse than the flu.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flu vaccine</td>
<td>Only older people need the flu vaccine.</td>
<td>Flu vaccine is good for all ages: Adults and children with conditions like asthma, diabetes, heart disease, and kidney disease need to get flu vaccine. People who are active and healthy can benefit from the protection the flu vaccine offers. The flu vaccine is beneficial to all age groups.</td>
<td>How likely is it that you will get a flu vaccine in the future? (0: Extremely unlikely – 10 Extremely likely)</td>
<td>The flu vaccine is beneficial to all age groups, not only the elderly.</td>
</tr>
<tr>
<td>Myth - 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flu vaccine</td>
<td>You must get a flu vaccine before winter for it to be effective.</td>
<td>Flu vaccine still effective in winter: Flu vaccine can be given before or during the flu season. It only takes about two weeks for antibodies to develop in the body, so as long as the current season’s flu is in circulation, the vaccine will be beneficial. While the best time to get flu vaccine is in autumn, getting immunized in winter can still protect you against the flu.</td>
<td>If it was winter, would you be less inclined to get a flu vaccine than in autumn? (0-10 Definitely not - Definitely)</td>
<td>If you receive the flu vaccination in winter, you will still be protected.</td>
</tr>
<tr>
<td>Myth - 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item</td>
<td>Claim</td>
<td>Explanation</td>
<td>Inference Question</td>
<td>Fact only claim</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------</td>
</tr>
<tr>
<td>Flu vaccine</td>
<td>Not everyone can take the flu vaccine</td>
<td>Some people should avoid flu vaccine: You might not be able to get the flu vaccine if you are allergic to eggs (used in making the vaccine), are very ill with a high fever, or have had a severe reaction to the flu vaccine in the past. Other rare conditions such as Guillain-Barre syndrome, an uncommon paralysing illness, may also prevent you from receiving the flu vaccine.</td>
<td>Would you ask your friend about their allergies before recommending the flu vaccine to them? (0-10 Definitely not - Definitely)</td>
<td>Not everyone can take the flu vaccine</td>
</tr>
<tr>
<td>Flu vaccine</td>
<td>Even if I get the flu vaccine, I can still get a mild case of the flu.</td>
<td>Mild flu and the vaccine: The flu vaccine protects most people from the flu. Sometimes a vaccinated person will still get the flu, but they will be far less sick than without the vaccine. Flu vaccines will not protect you 100% from mild cases of the flu.</td>
<td>How likely is it that you would encourage your own friends or relatives to get a flu vaccine? (0-10 Extremely unlikely – extremely likely)</td>
<td>Even if I get the flu vaccine, I can still get a mild case of the flu.</td>
</tr>
<tr>
<td>Flu vaccine</td>
<td>People can die from the flu.</td>
<td>The flu can kill: Influenza (flu) is a highly infectious disease of the lungs, and it can lead to pneumonia. Each year about 114,000 people in the United States are hospitalized and about 36,000 people die because of the flu. Most who die are 65 years and older. But small children less than 2 years old are as likely as those over 65 to have to go to the hospital because of the flu.</td>
<td>How would you rate the potential benefits of flu vaccination? (0-10)</td>
<td>People can die from the flu.</td>
</tr>
<tr>
<td>Item</td>
<td>Claim</td>
<td>Explanation</td>
<td>Inference Question</td>
<td>Fact only claim</td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
<td>-------------</td>
<td>--------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Hypnosis Myth -1</td>
<td>A talented hypnotist can hypnotise you against your will.</td>
<td>Hypnosis is voluntary: Hollywood movies and dramatized hypnosis performances have distorted people’s view of hypnosis. During hypnosis, individuals remain fully aware of their behaviour. They can easily refuse to comply with the hypnotist’s suggestions if they wish to do so.</td>
<td>How likely is it that a talented hypnotist could make you jump off a cliff? (0: Extremely unlikely – 10 Extremely likely)</td>
<td>It is impossible for any hypnotist, no matter how talented they are, to hypnotise you against your will.</td>
</tr>
<tr>
<td>Hypnosis Myth -2</td>
<td>Hypnosis is useful for retrieving repressed memories.</td>
<td>Memories “recovered” using hypnosis usually false: While it is plausible that an individual could suddenly recall a traumatic event they had repressed, the memory is more likely to be valid if it occurs spontaneously, rather than working towards a revelation in therapy or during hypnosis. Memories recovered during hypnosis are more likely to be false memories resulting from suggestion.</td>
<td>If your friend tells you about a memory that was helped to be uncovered by a hypnotherapist, would you assume that it is a true event? (0: Extremely unlikely – 10 Extremely likely)</td>
<td>Repressed memories uncovered during hypnosis are likely to be false memories.</td>
</tr>
<tr>
<td>Hypnosis Myth -3</td>
<td>Under hypnosis, a person can perform better physically and mentally.</td>
<td>Hypnosis unrelated to performance: Hypnosis is likely to result from a mixture of relaxation, conformity, obedience, suggestion and role-playing. Some studies find that participants perform worse while under hypnosis, as participants make more errors, and show less motivation to do well. Other studies find no performance differences in intellectual or physical performance.</td>
<td>Do you think you could do a crossword puzzle quicker and more accurately if you were hypnotised, than under normal circumstances? (0-10 Definitely not - Definitely)</td>
<td>A person’s abilities will be approximately the same, whether or not they are hypnotised.</td>
</tr>
<tr>
<td>Item</td>
<td>Claim</td>
<td>Explanation</td>
<td>Inference Question</td>
<td>Fact only claim</td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
<td>-------------</td>
<td>--------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Hypnosis Fact -1</td>
<td>Hypnosis can be used in pain management</td>
<td>Hypnosis reduces pain: People tend to have a higher pain threshold whilst hypnotised. It has been performed during childbirth and other medical procedures such as the dressing change of burn patients, and seems to help patients manage their pain.</td>
<td>If you were in a great deal of pain, and hypnosis was offered, would you accept? (0-10 Definitely not - Definitely)</td>
<td>Hypnosis can be used in pain management</td>
</tr>
<tr>
<td>Hypnosis Fact -2</td>
<td>It is possible to ‘hypnotise’ animals.</td>
<td>Animals can be hypnotised: In animals, hypnotic states are commonly referred to as ‘tonic immobility.’ It is thought to be an involuntary reflex that an animal enters in response to a threat. For example, many species of shark become immobile if you turn them upside down, and lobsters can be put into this state if you stroke their head. It is possible to hypnotise an animal as it plays a role in their survival, helping the animal not be noticed, or ‘play dead’.</td>
<td>If we gave you an animal, and instructions on how to hypnotise it, what is the likelihood that it would work? (0: Extremely unlikely – 10 Extremely likely)</td>
<td>It is possible to ‘hypnotise’ animals.</td>
</tr>
<tr>
<td>Hypnosis Fact -3</td>
<td>The susceptibility to be hypnotised varies from person to person.</td>
<td>Individual differences in hypnosis: There are scales such as the ‘Stanford Hypnotic Susceptibility Scale’, which measure an individual’s susceptibility to hypnosis. Your motivations (how much you want to be hypnotised), and expectations (how much you believe you will be hypnotised) influence the likelihood that you will be hypnotised.</td>
<td>If your friend really wanted to be hypnotised, but you did not – would the hypnotist be able to hypnotise both of you? (0-10 Definitely not - Definitely)</td>
<td>The susceptibility to be hypnotised varies from person to person.</td>
</tr>
</tbody>
</table>
Appendix C: Chapter 5 Stimuli
## Table 12
Trump and unattributed items and their corresponding explanations

<table>
<thead>
<tr>
<th>Item number</th>
<th>Trump item</th>
<th>Unattributed item</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Misinformation - 1</td>
<td>Donald Trump said that there are 30-34 million illegal immigrants residing in the US.</td>
<td>There are 30-34 million illegal immigrants residing in the US.</td>
<td>According to the Department of Homeland security’s most recent estimate, the number of illegal immigrants currently residing in the US is 11.4 million people. In 2014, the Pew Research Center placed this number at 11.3 million people, while the Center for Migration Studies estimated approximately 11.0 million individuals.</td>
</tr>
<tr>
<td>Misinformation - 2</td>
<td>Donald Trump said that the real unemployment rate is between 24-48 %.</td>
<td>The real unemployment rate is between 24-48 %.</td>
<td>The US Bureau of Labor Statistics states that the official unemployment rate is 5.5 %. There are more lenient measures of unemployment that includes people who have part time jobs but would prefer full time work, and people who are not looking for work. If these individuals are included, the unemployment rate only rises to 10.8 %.</td>
</tr>
<tr>
<td>Misinformation - 3</td>
<td>Donald Trump said that the gross domestic product (GDP) growth rate has never been below zero.</td>
<td>The gross domestic product (GDP) growth rate has never been below zero.</td>
<td>The GDP growth rate indicates how much a country’s production has increased in comparison to the previous year, and is an indicator of a country’s economic strength. The US Department of Commerce Bureau of Economic Analysis reports that there were 42 occurrences since 1946 where the US growth rate was below zero.</td>
</tr>
<tr>
<td>Misinformation - 4</td>
<td>Donald Trump said that vaccines cause autism</td>
<td>Vaccines cause autism</td>
<td>There is strong consensus in the scientific community that vaccines are not linked to autism. For example, one study by the Danish Epidemiology Science Center tracked all children born in Denmark from 1991 to 1998 and concluded that there was no increase in the rate of autism for vaccinated as opposed to non-vaccinated children.</td>
</tr>
<tr>
<td>Item number</td>
<td>Trump item</td>
<td>Unattributed item</td>
<td>Explanation</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Fact - 1</td>
<td>Donald Trump said that the US debt is $18 trillion.</td>
<td>The US debt is $18 trillion.</td>
<td>Each day the US Department of the Treasury releases the exact amount of US debt. The debt currently sits at $18.15 trillion. The public holds $13 trillion, which is debt held by individuals and corporations. A further $5 trillion is intragovernmental debt, which is the government borrowing from federal trust funds.</td>
</tr>
<tr>
<td>Fact - 2</td>
<td>Donald Trump said that the US spent $2 trillion on the war in Iraq.</td>
<td>The US spent $2 trillion on the war in Iraq.</td>
<td>A report by the Watson Institute found that as of 2013, the US spent $1.7 trillion on the war in Iraq. While the appropriations for the war were under $800 billion, the Watson report also included the cost of disability, Defense Department base spending costs, and homeland security expenditures attributed to Iraq.</td>
</tr>
<tr>
<td>Fact - 3</td>
<td>Donald Trump said that the US is ranked 26th in the world in education.</td>
<td>The US is ranked 26th in the world in education.</td>
<td>The Program for International Student Assessment is a test for children 15 years of age. It is administered every three years, and largely focuses on math, reading and science. The most recent test was administered in 2012, when the US ranked between 24th and 35th on the three measures, scoring below average in each category.</td>
</tr>
<tr>
<td>Fact - 4</td>
<td>Donald Trump said that Nabisco, the company that manufactures Oreo cookies, is moving jobs to Mexico.</td>
<td>Nabisco, the company that manufactures Oreo cookies, is moving jobs to Mexico.</td>
<td>Nabisco is a food snack company that is known for products such as Oreos and Ritz crackers. It announced that it will open a new factory in Mexico, rather than investing the $130 million in their current factory in Chicago. Over the next year, half the workers at the Chicago-based bakery will lose their job, which totals 600 employees.</td>
</tr>
</tbody>
</table>
Appendix D: Post-Explanation Additional Analysis
If the post-explanation items are analysed separately, we see similar trends as to the accuracy score analyses. For the fact items, we likewise see a main effects of source, $F(1,1538) = 19.79; p < .001; MSE = 2.38; \eta^2_p = .01$, retention interval, $F(1,1538) = 190.48; p < .001; MSE = 2.38; \eta^2_p = .11$, and an interaction of source and retention interval, $F(1,1538) = 9.00; p = .003; MSE = 2.38; \eta^2_p = .006$. In addition, there is a Trump support and source interaction, $F(2,1538) = 3.28; p = .038; MSE = 2.38; \eta^2_p = .004$, indicating that Trump support influences the degree to which the Trump attribution influences belief. A planned comparison indicates that Democrats do not update their belief in the factual items to the same extent as the Republican groups if the information is attributed to Trump, $F(1,1538) = 5.12; p = .024; MSE = 2.37$. There is also an interaction of Trump support and retention interval, $F(2,1538) = 3.44; p = .032; MSE = 2.38; \eta^2_p = .004$. While Democrats over both time periods are worse at updating their belief in the facts if information is attributed to Trump, Republicans immediately update their belief equally in the Trump and unattributed conditions, yet after one week belief in the Trump information reduces below the unattributed condition, $F(1,1538) = 5.08; p = .024; MSE = 2.38$.

The post-explanation misinformation items reveal three main effects. A marginal main effect of source, $F(1,1538) = 3.78; p = .052; MSE = 3.20; \eta^2_p = .002$, indicating that the Trump attribution led to less accurate belief, and a main effect of Trump support, $F(2,1538) = 33.35; p < .001; MSE = 3.20; \eta^2_p = .04$, indicating that Republican supporters had higher belief in the misinformation than Democrats and Republican supporters, $F(1,1538) = 53.00; p < .001$. Finally, a main effect of retention interval, $F(2,1538) = 64.50; p < .001; MSE = 3.20; \eta^2_p = .04$ indicating belief increased over time, all groups forgetting that the misinformation was in fact false. There was no interaction of source and retention interval, indicating that unlike the fact scores (where Trump attribution led to less accurate beliefs particularly over time), the information associated with Trump is considered to be less accurate over both time periods.
Appendix E: Feelings Thermometer Analysis
A 2 × 2 × 3 factorial ANOVA on the feelings-change index revealed an interaction of retention interval and Trump support, $F(2,11530) = 21.67; p < .001; MSE = 139.37; \eta^2_p = .03$, indicating that Republican non-supporters and Republican supporters changed their feelings towards Trump more than Democrats. Mimicking voting preferences, over the course of a week Republican supporters indicated feeling “cooler” towards Trump, and Republican non-supporters indicated feeling “warmer”.

*Figure 18.* Feelings-thermometer scores over time across partisanship and sources. Error bars denote 95% confidence intervals.