Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

The operation of confirmation bias: Discourse analysis of witnesses' evidence about the conduct of a sexual abuser

A thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Psychology at Massey University, Palmerston North, New Zealand

S. Bourne

Embargoed to 30 June 2020

Abstract

Confirmation bias is a cognitive bias where beliefs formed at an early stage are unreasonably sustained even in the face of new information that makes those beliefs untenable or at least questionable. In these circumstances, new information that does not fit with the earlier-formed beliefs may be ignored, while information supporting those beliefs is accepted readily as lending credence to them. During 2010 and 2011, Jonathan Lord was employed by the YMCA in New South Wales, Australia, until a child disclosed that he had been inappropriate touched by Mr Lord. This led to Mr Lord being convicted of 13 representative offences including multiple aggravated indecent assault charges and two counts of sexual intercourse with a child under 10, relating to 12 children enrolled in the YMCA Before and After School Care service. Subsequently, several of the children's parents, and some of Mr Lord's YMCA co-workers, gave evidence to the Royal Commission into Institutional Responses to Child Sexual Abuse that was conducted in Australia from 2012 to 2017. This study shows how confirmation bias may lead to tragic or destructive outcomes in some circumstances. The use of discourse analysis in this study has afforded a 'micro context' understanding of how Mr Lord's abuse of children associated with the YMCA service persisted undetected for more than a year, despite Mr Lord breaching YMCA rules and being observed engaging in other questionable conduct regarding children in that period.

Acknowledgements

Firstly, I acknowledge the victims of Jonathan Lord, and the victims' parents and Mr Lord's YMCA co-workers who gave evidence before the Royal Commission. It is clear from their testimony that these parents and workers wanted the best for the children in their care, and genuinely believed that the YMCA Before and After School Care service would afford a safe environment for the children.

Secondly, I acknowledge my family and friends who have been patient and understanding while I worked on this thesis. I especially acknowledge my parents, Ian and Margaret, who have encouraged my studies throughout my life although they suggest I might have done enough already.

Thirdly, I acknowledge my supervisor John Fitzgerald PhD who always gave sage advice despite my sometimes erratic approach to study routines and outputs.

The operation of confirmation bias: Discourse analysis of witnesses' evidence about the conduct of a sexual abuser

Contents

Chapter 1 – Introduction

The purpose of this chapter is to introduce my study, and describe the structure of this thesis.

My study is about confirmation bias, a cognitive bias where beliefs formed at an early stage are unreasonably sustained even in the face of new information that makes those beliefs untenable or at least questionable. In these circumstances, new information that does not fit with the earlier-formed beliefs may be ignored, while information supporting those beliefs is accepted readily as lending credence to them.

The study here builds on the work on many distinguished researchers, in particular German psychologist Gerd Gigerenzer (2001, 2004), American psychologist Raymond Nickerson (1998), Nobel prizewinners Amos Tversky and Daniel Kahneman (1974), and Cornell psychology professor Thomas Gilovich (Gilovich, Griffin & Kahneman, 2002).

How confirmation bias leads to errors in decision-making

One specific area where confirmation bias can have tragic or destructive consequences is when beliefs are formed on limited data regarding the integrity and trustworthiness of another person who is entrusted with the care of a child, such as a co-worker or childcare provider. Those initial beliefs may then constrain the objective assessment of subsequent information about that other person. Individuals who prey on or take advantage of children for their own self-gratification often become very adept at promoting themselves as trustworthy, thereby lulling work colleagues and children's parents into a false sense of security. A difficulty that arises in identifying and assessing confirmation bias, even though it is highly prevalent, is that each situation is unique and it is necessary to keep in mind that:

- before all relevant information has been received and considered, our unconscious is already formulating the basis for a decision using the information to hand so far, and is establishing its own 'ground rules' about what further information may be deemed useful or not useful;

- good decision making requires the logical and rigorous assessment of all relevant information, to ensure that fallacies and unsupported information do not inform or influence decision-making;

- when we are under pressure to make decisions quickly, biases like confirmation bias are more likely to affect the outcome;

- 'keeping an open mind' means giving careful consideration to all information received right up to the time when a decision must be made, and always being alert to the possibility of bias;

- in some cases, confirmation bias is almost impossible for an individual decision-maker to overcome: robust decision-making strategies are required to avoid outcomes being affected by bias.

A common type of decision-making used to study confirmation bias is the decision-making of jurors. For example, a jury member may form an early belief as to the guilt or innocence of the accused person and may hold to that belief despite later information supporting the opposite conclusion. However, in Australia it is not permitted to interview jurors after a trial, and so research into confirmation bias affecting jurors is mainly conducted in other countries or using mock juries. The widely varying rules that affect legal processes in other countries may call into question our ability to apply that research to Australia, and mock juries are inevitably artificial constructs that may not reliably represent real-life situations. Nonetheless, we have a reasonable basis now to believe that some specific circumstances can tend to bias jury decision-making, such as when a

defendant is on trial for more than one charge (Goodman-Delahunty, Cossins & Martschuk, 2016, pp.46-7).

Royal Commission of Inquiry into Institutional Responses to Child Sexual Abuse 2012-17

In this thesis, I am wanting to show situations where confirmation bias and its operationalisation is found in real life, influencing decision making and leading to negative outcomes. The Royal Commission into Institutional Responses to Child Sexual Abuse ("Royal Commission") was established in November 2012, and ran hearings and conducted meetings with victims of child sexual abuse over a period of more than 4 years. The Royal Commission delivered its final report on 15 December 2017. A comprehensive description of the Royal Commission's terms of reference, activities and outputs is contained in Appendix 1 of this thesis.

The first two case studies conducted by the Royal Commission related to the conduct of two men, Steven Larkins and Jonathan Lord, who each held trusted positions where they had access to young children. Larkins and Lord were responsible for the abuse of multiple children, and both were sentenced to time in jail for their offending. The relevance of these case studies to my research arises from the clear evidence of confirmation bias affecting the decision making of co-workers of the offenders and parents of the abused children. Very often co-workers of Mr Larkins and Mr Lord failed to see warning signs, perhaps due to their pre-existing beliefs regarding the offenders' professional commitment to children's welfare. In Mr Lord's case, co-workers noted his apparent commitment and "enthusiasm" for child care, that led to him being trusted by workers and parents alike.

Parents dealing with Mr Lord during his time at the YMCA in Carringbah were also misled by how much he was liked by their child or children, and his willingness to go the extra mile to engage with the children and support their interests – acts that we might now describe as 'overt helpfulness' raising the

spectre of grooming. Even when Mr Lord's behaviour was clearly outside the rules of the institution that employed him – such as being the sole carer present on bus during trips with the children, taking them on trips in his own car, and offering to babysit the children outside of his work duties – that was not enough to raise the alarm. Some parents later realised that Mr Lord had been grooming them as much as he had been grooming their children.

The transcripts of witnesses' evidence to the Royal Commission reflect (in the main) a genuine attempt by witnesses to describe their experiences and knowledge of matters relevant to the commission's inquiry. Much of the evidence heard by the Royal Commission was harrowing, as individuals who were sexually abused recounted incidents of abuse and the impacts of those incidents, parents described their distress and confusion about what happened to their child, and co-workers of perpetrators described how a person they trusted was later found to be engaged in child abuse activities such as grooming, possessing and distributing child abuse material, and incidents of direct sexual abuse against the children and young persons who they were employed to protect. The commission also heard from leaders and managers of institutions, some of whom did not have personal first-hand knowledge of the abuses, but who could speak about policies and how specific incidents were handled within their respective organisations.

The scope of this thesis is limited to the case study regarding Mr Lord, and a list of relevant witnesses can be found in Appendix 2. The transcripts of these witnesses' evidence are revealing in relation to the formation of an initial belief about the offender, and the persistence of that belief. Moreover, they reveal factors that might in other circumstances have called the initial beliefs into question, and provide some insight into why they did not. A limitation of this study is that it is wholly reliant on the evidence given by the identified witnesses to the Royal Commission, and therefore some relevant context is unknown.

It is sobering to realise that during a period of just over two years in his trusted role as a child care worker at the YMCA, Jonathan Lord sexually abused at least

12 children. In 2013, Mr Lord was convicted of 13 offences including multiple aggravated indecent assault charges and two counts of sexual intercourse with a child under 10 (Royal Commission, 2014, p.4). However, these were representative charges and do not reflect the full magnitude of the offending. In the case of one child, referred to by the Royal Commission as "AO", his mother "AN" gave the following evidence of her personal beliefs:

> AN.29 I trusted and believed that if Jonathan Lord was a supervisor at the YMCA, he was more than qualified.
> AN.30 … In around December 2010 Jonathan Lord started babysitting [AO] every Saturday. I would
> AN.31 pay him $100 for the day. I now know that Jonathan indecently assaulted [AO] the very first time that he
> AN.32 babysat him and that the abuse occurred every Saturday thereafter until Jonathan stopped babysitting
> AN.33 [AO] in October 2011.

Research purpose, tasks and methodology

Considering all of the above, and that the main purpose of the study is to describe confirmation bias and illustrate its operation with the help of discourse analysis of witness statements in a real-life situation, the research objectives are:

1. To describe the theoretical background related to confirmation bias: Human Biases and Cognitive Heuristics;
2. To understand the meaning of confirmation bias and how it is operationalised;
3. To identify relevant statements of parents of sexually abused children, and co-workers of Mr Lord, using discourse analysis;
4. To demonstrate how confirmation bias can influence decision making and lead to adverse outcomes.

For present purposes I have regarded the Royal Commission reports of witnesses' evidence as evincing 'real-life' material for discourse analysis, from which beliefs and meanings related to this topic may be distilled.

The research question is how confirmation bias is represented within witnesses' evidence in a real-life context. This phenomenon is researched by considering the beliefs of the perpetrator's co-workers and the victims' parents in logical categories (chapter 5). The discursive methodology helps to reveal the influential role of individual subjectivity – i.e., confirmation biases informing and influencing decision making – in real-life situations that may lead to tragic or destructive consequences.

Structure of this thesis

Any study of a specific bias needs to be viewed in the context of human biases and cognitive heuristics generally. While many people may regard biases negatively, they play a very important role in our unconscious decision-making. Without biases we might not have survived as a species. For example, a natural bias towards the possibility that a long, thin 'snaky' object is a snake is more likely to inform a life-saving decision to run away from that object than if additional time is taken to decide if it is a snake or a stick. In that specific example we are probably not engaging a confirmation bias, but different biases can have some similar characteristics.

To provide context for my focus on confirmation bias, the next chapter (chapter 2) is a short introduction to human biases and cognitive heuristics. This is intended to establish a context for Chapter 3, where I focus on confirmation bias. Chapter 4 addresses the methodology employed in this study, and Chapter 5 discusses the analysis that I have conducted using a discursive approach. Finally, Chapter 6 summarises this study and my conclusions.

Chapter 2 – Briefly About Human Biases and Cognitive Heuristics

The purpose of this chapter is to introduce the subject of human biases and cognitive heuristics broadly, attempting to illustrate and explain the main types of biases and heuristics that affect our decision-making and behaviours. I have also included a section on informal fallacies. Including these fallacies may be controversial, but some modern researchers in this area make a strong case for representing informal fallacies as heuristic 'paraschemes' that affect decision making (Cummings, 2014, p.5).

Gigerenzer and Bright (2009, p.136) propose that a hypothetical "grand planner" might have had three alternative options to consider when designing the optimal human mind:

- Perfect memory, so that all previous experiences become a guide to future actions.

- Highly developed processing ability to make accurate decisions based on available data.

- A mind that can make lightning fast inferences based on a limited set of data.

However, previous experiences could only be a guide to future actions in a predictable world, and even the fastest 'processing unit' requires time to consider all the possible alternatives in a situation and choose the most suitable one. That leaves inference-based decision-making, which is of course exactly what we have available to us. While sometimes our inferences are spectacularly wrong, more often they are right enough to get us through the day unscathed.

Any thinking process that involves inferences or subjective reasoning is likely to fall within the ambit of *biases.* Biases often involve prejudice developed over time due to cultural influences or personal experiences, and a narrow definition of biases might describe them as psychological responses that favour certain outcomes over others based on pre-existing preferences. Sometimes we refer to biases as 'cognitive biases' to reflect how they arise from cognitive rules or shortcuts known as *cognitive heuristics.*

In this thesis, I suggest defining cognitive heuristics as strategic mental processes to facilitate timely decision-making using limited data. These mental processes are derived mainly from our experiences (Ross, 2014, p.33). This means that how a person unconsciously deals with a situation in the future is likely to be informed by how a similar situation unfolded in his or her past.

It is important to keep in mind that cognitive heuristics and cognitive biases are not the same, although heuristics are involved in creating biases and in turn are shaped by biases. Both biases and cognitive heuristics play a significant role in our decision-making, often concurrently.

2.1. Biases

Although most of us will readily acknowledge that we sometimes express bias towards a particular person or a particular outcome, we are all largely oblivious to the full extent of the biases and intuitive leaps that influence almost every aspect of our daily lives. We may believe that our decisions are always rationally considered, even though dispassionate analysis of a decision-making process will typically reveal otherwise. Many factors are involved here, some more within our conscious control than others.

It is widely accepted that we often perceive bias negatively, particularly when we feel "biased against", but in theory it is a neutral process, neither negative nor positive. That said, the outcome of a decision informed by bias could in practice

be very negative or very positive depending on how the decision plays out. Lauwereyns (2010, p.xiv) describes bias as follows:

> "bias is a core brain mechanism that attaches different weights to various information sources, prioritizing some cognitive representations at the expense of others."

Seen in this way, bias is simply a method to evaluate data. However, influences on the formation and operation of biases should not be ignored as they can significantly affect how bias manifests itself in a particular situation. Put simply, biases inform expectations based on our pre-formed beliefs about what is most preferred, but reality can and does intrude in the form of external stimuli affecting our perceptions and the implementation of the bias. In this way, our biases and sensitivity to stimuli act in concert (Lauwereyns, 2010, p.79).

The pursuit of happiness?

Even if biases are not always negative factors, a valid question remains: why have biases evolved to be so integral to our behaviour, often bypassing available logic that would sometimes produce a better result for us? Lauwereyns (2010, pp.83-84) suggests this might be associated with the "pursuit of happiness", and the idea that happiness could be the "systemic goal of informed decision making" even though that satisfaction may be short-lived and "explicable as computational processes gone awry." Other researchers have proposed that biases evolved as a risk or error management strategy (Kahneman & Tversky, 1979, Haselton & Buss, 2000, Haselton & Galperin, 2013, Nettle, 2004, Yamagishi et al., 2007, and Johnson, 2009, all as cited in Johnson, Blumstein, Fowler & Haselton, 2013, p.474). Johnson et al. give the example of how biases might be more likely to lead us to mistake a stick for a snake (a "harmless" error) than to mistake a snake for a stick (a potentially fatal error). Similarly, if we are conscious of the possibility of being run down by a car when crossing the road, we are less likely to die in that way even though the consistent vigilance comes at a cognitive cost. The

rationality of this approach is described in this way, referring here to cognitive bias as 'behavioral bias': (Johnson et al., p.476)

> "To a third party, this *behavioral bias* may appear to be ‘economically irrational’ because we are consistently wasting time worrying about an event that rarely happens, but if it helps us to maximize our chance of survival (and reproduction), it is ‘adaptively rational’. In other words, our behavior is not tuned to fit a model of expected *payoffs,* but a model of expected *fitness.* For this reason, a behavior that seems 'biased' would not seem biased if we had complete information about the consequences of those actions for survival and reproduction."

Therefore, despite their apparent irrationality cognitive biases can lead to outcomes that improve species survival. Johnson et al. (2013, p.478) suggest that bias-informed strategies make sense because they most accurately address "the asymmetric costs of false-positive and false-negative errors made under uncertainty."

Happy outcomes? Not so much

Sometimes we can identify how a bias has formed, such as if we have won Lotto using a specific set of numbers and feel more confident using those same numbers after that. While this example of biased confidence ignores the reality that every number set has the same probability of winning Lotto, the general principle that a successful formula should be repeated is deeply entrenched in the human psyche. This is not necessarily a bad thing, as very often a run of successes in a non-random scenario can indeed be attributed to a pattern of behaviour based on earlier success. We can imagine this working for our forebears on an African savanna, as they honed their method of catching wildebeest for food. Eventually the method would inform a powerful bias, overwhelming any stray thought of trying a different method. However, a change

in the environment might then have catastrophic results, as the old bias failed to produce the same outcomes as were previously enjoyed.

Another great challenge is the manner in which a belief once formed becomes hard to dislodge. This is one manifestation of confirmation bias, and is a lot more common than most people realise. In some cases, an individual will cling tenaciously to a belief even after he or she has been shown conclusive proof that the belief is erroneous. However, it is not necessarily irrational or senseless for us to conduct ourselves subconsciously in this way: (Kahan et al., 2017, p.77)

> "It is perfectly rational, from an individual welfare perspective, for individuals to engage with decision-relevant science in a manner that promotes culturally or politically congenial beliefs… [because] forming a belief at odds with the one that predominates on it within important affinity groups of which such a person is a member could expose him or her to an array of highly unpleasant consequences (Kahan, 2012)."

However, the reality that some errors do become consequential is a reasonable basis to suggest that encouraging people to independently review the available evidence and reach their own conclusions is prudent, even if it may be an uphill battle to get them to do this.

Broad categorisations of biases

Considering biases more generally, Boyd and Richerson (1985, pp.134-5, as cited in Enfield, 2014, p.25) have proposed that biases fall into one or more of three categories:

- direct: directly advantages the biased individual.

- indirect: individual advantage depends upon other group members interpreting the relevant conduct and identifying the individual as part of the group.

- frequency-dependent: behaviours that are common within the group are copied (even if they are not necessarily advantageous).

In practice, it may be difficult to identify the bias or biases responsible for specific individual behaviour at a point in time, but an essential feature of all frequency-dependent biases is that they are connected to the transmission of cultural information within the group. Some direct and indirect biases that influence group behaviour are also definable as transmission biases.

Context biases and content biases

From an evolutionary perspective, it may be more useful to distinguish between 'content biases' and 'context biases'. Context biases a.k.a. 'model-based biases' are also transmission biases because they involve transmission of cultural information (Boudry, Blancke & Pigliucci, 2014, p.5). A good illustration of a context bias is the situation where a teenager emulates the dress style of a pop idol such as Justin Bieber, influenced by Mr Bieber's dress sense because it is popular within the teenager's personal context.

Emulation of a celebrity or someone who is seen as important within the group implicates a context bias more specifically known as 'prestige bias' (Richerson & Boyd, 2005, p.124). Whether dressing like Mr Bieber is a good strategy depends on the individual's objectives in life, but for a teenager it is possible that success might be measured in terms of being recognised by one's peers as committed to the Bieber 'look'. This kind of bias may be direct or indirect, and does not necessarily depend on frequency.

Another form of context bias is 'conformist bias', which leads us to adopt – i.e., conform with – the views that are most strongly held by others around us. Conformist bias therefore leads us to emulate beliefs or behaviours that are common within our personal context. Richerson and Boyd (2005, p.120) note that "conformity is not just simple cultural influence", but instead involves unconsciously choosing a model of behaviour according to the frequency of a specific trait. Conformist bias may also be direct or indirect or both, as the relevant belief or behaviour may be perceived as benefitting the individual and the group in varying degrees, but what is critical here is that the conformist bias is frequency-dependent. However, we cannot assume that all individuals in a group that exhibits conformist bias will necessarily be conformists, as some "maverick" members may not respond to behavioural frequency in the group (Efferson, Lalive, Richerson, McElreath & Lubell, 2008, p.56). This is significant because it suggests that while conformity is important to survival, some degree of nonconformity is tolerable and perhaps even serves a valuable purpose.

Unlike context biases, content biases such as confirmation bias build on our innate mental model of how our world is constructed, including the content of our world and our intuitions (Spelke & Kinzler, 2007, as cited in Boudry, Blancke & Pigliucci, 2014, p.5). These intuitions are typically influenced by 'personality factors' that increase or decrease an individual's propensity to adopt beliefs or behave in a particular way (Enfield, 2014). At times it appears as though a war is raging between science and our intuition, as new discoveries challenge our 'commonsense' understanding of the world. However, that 'commonsense' often prevails for a long period beyond the emergence of a competing scientific paradigm, because science inevitably struggles in its path to consistent certainty, and also because intuition is immediately at hand while scientific truth is far less accessible until it becomes ubiquitous.

How biases affect groups

Significant experiences relevant to a group as a whole are more likely to be transmitted within the group than isolated personal experiences that are only relevant to one individual group member or to a small sub-group. This dynamic was sharply evident from divergent reactions to the outcome of the infamous O.J. Simpson arrest and trial that preoccupied the United States for much of 1994 and 1995. In relation to the verdict in that trial, Brigham and Wasserman (1999, p.1368) comment:

> "Because people view the world through the prism of their own experiences, and because Whites and African Americans have remarkably different experiences when it comes to discrimination and special treatment as a result of race, it is no wonder that the perceptions of the two races can be remarkably different... It is these life experiences which may condition a juror's view of the evidence, and thus influence the verdict."

Brigham and Wasserman (1999, p.1368) suggest that individual life experiences are typically formed through engagement in a specific group or community, so that less malleable individuals in the group or community become role models for more malleable individuals and "attitudes are passed from person to person."

Biases as a force for good

Rather than rejecting bias, Lauwereyns (2010, p.14) suggests that it "deserves to be exonerated, polished, and used properly... as a fundamental property of human thinking, perceiving, and decision making."

Lauwereyns (2010, p.200) says he cannot imagine a world without bias, because our "fears and desires" give meaning to life. However, biases are capable of being managed in many situations, and in particular the ability to bring biases into "our conscious awareness, allows us to apply them more carefully, more

systematically, and more successfully" (Lauwereyns, 2010, p.228). This sentiment is echoed by Johnson, Blumstein, Fowler and Haselton (2013, p.480), who see error management theory – described as "a unifying framework for understanding decisions made under uncertainty, wherever there are asymmetric costs of false-positive and false-negative errors" – as a way to mitigate damaging errors while also allowing "the risks that have driven remarkable feats of human endeavour."

Biases and cognitive heuristics

The phenomena of biases and cognitive heuristics are not mutually exclusive. In fact, our cognitive heuristics are frequently informed by our biases, with a consequential effect on our unconscious decision-making.

Confirmation bias a.k.a. confirmatory bias is a cognitive bias but also works in conjunction with cognitive heuristics. Formation of a belief, for example that a person is honest, may be motivated by the perception of specific stimuli, but how that belief is then implemented heuristically affects the influence of the bias in a particular situation. Plous (1993, p.233) suggests that the label 'confirmation bias' "usually refers to a preference for information that is consistent with a hypothesis rather than information which opposes it." Klayman and Ha (1987, p.220, as cited in Plous) identify how confirmation biases arise from a "positive test strategy" which, "like any all-purpose heuristic ... is not always optimal..."

The effect of confirmation bias is that once a belief has informed the rule-base of a heuristic process, subsequent information may be interpreted or preferred in a manner that is consistent with the belief. For example, if a juror in a criminal trial forms a belief as to the defendant's innocence early in the trial, that juror may pay greater attention to subsequent evidence supporting the defendant's innocence than to evidence supporting guilt. In analysing this situation, we may consider that the juror has become motivated towards favouring evidence of innocence. From a cognitive perspective, the juror may be unconsciously using an availability

heuristic – a type of cognitive heuristic that prefers the most easily recalled information – to determine the value of the evidence, which potentially has the effect of preserving the bias if the most easily recalled information is information supporting innocence. You can see from this example that the interplay between biases and cognitive heuristics can be complex, and sometimes the effective agent cannot be determined.

2.2. Cognitive heuristics

You have probably at some stage played the popular fund-raising game where people are given the opportunity to guess the number of jellybeans in a jar (this can also be done with Smarties or M&Ms or other lollies) usually in return for a small donation that goes towards the prize for guessing the correct number. My experience is that individual ability to estimate the number varies widely in these games, but most people will nonetheless make an attempt. Many of those individuals will then express confidence that they have a chance of winning – up until the jellybeans are counted out. Although a few people will make a determined effort to calculate the number of jellybeans, most will simply write the first number that comes into their head relying on their 'intuition'. Intuition does not require conscious reasoning.

Any cognitive process that we use to unconsciously find an answer or solve a problem may be referred to as a 'cognitive heuristic'. The term 'heuristic' distinguishes the process from a more conventional (typically slower) process of computing the correct or most accurate answer where that can be done – i.e., where all the information required to make an accurate determination is available – although we tend to instinctively use heuristics even if that information is available. For example, if you are attending a conference and someone asks you how many people are in the room, you could stand on a chair and count the number of heads you see, but many people will simply make an estimate based on an intelligent guess. For most purposes that intelligent guess will be adequate, but if you ask 10 attendees the same question you will almost certainly get 10

different responses. This level of inaccuracy would not be acceptable if, for example, the head count was being used to justify a funding application. Moreover, our prior knowledge such as how many people attended the previous year's conference might bias us towards a greater or lesser estimate for the current event. This is a good example of how biases and cognitive heuristics, while different, can work in concert to affect an outcome.

Some writers argue that the word heuristic is ambiguous, as its meaning has changed over time and it has been used to denote diverse concepts (e.g., Evans, 2009, p.36, as cited in Chow, 2014, p.978). Historically, Simon (1990, p.11, as quoted in Shah & Oppenheimer, 2008, p.207), proposed that heuristics are "methods for arriving at satisfactory solutions with modest amounts of computation." Implicit in this definition is the notion that some potentially relevant information will not be taken into account, and that the decision-making will be faster on account of the reduced computation time.

However, I might take issue with the word "satisfactory" in Simon's definition, as clearly some heuristic outcomes are far from that. Furthermore, a "modest amount" of computation will in some cases be virtually none, perhaps resting entirely on one factor (Gigerenzer and Gaissmaier, 2011, pp.460-3). My own working definition of cognitive heuristics is:

> "Strategic mental processes to facilitate timely decision-making using limited data."

While making decisions based on incomplete information may sometimes be dangerous, our heuristic response can be an excellent starting point. It is surprising how often that heuristic response turns out to be the best answer. Allowing our cognitive heuristics to guide us in this way can often reduce the effort that would otherwise be required to reach useful conclusions. For example, if you receive a burn from touching a hot stove element you will probably be more

vigilant next time you are cooking without giving any conscious thought to the previous incident (Ross, 2014, p.33).

However, while keeping ourselves safe from hot stove elements and other dangers remains important for modern homo sapiens, it is sometimes problematic for us that our cognitive heuristics are still operating in much the same way they have for many millennia without an evolved capacity to identify situations when they should switch off. For example, Milliman (1986, p.288) found that restaurant patrons exposed to fast music ate their meals significantly faster than patrons exposed to slow music. In Milliman's study, this had the consequence that the "slow-tempo" patrons spent nearly 50% more on drinks than the "fast-tempo" patrons. Next time you are in a restaurant, ask yourself whether the background music might be having an effect on how long you are taking to eat your meal and whether the restaurant owner might be trying to sell more drinks!

In situations where we must act quickly, cognitive heuristics typically do most of the heavy lifting. A good example is our 'fight-or-flight' instinct, where there is insufficient time for us to make a considered decision. At that moment of serious and immediate threat, our cognitive heuristics – informed by our life experiences up to that time – spring into action and (we hope) take us automatically to a place of safety.

Satisficing versus maximising

Humans and most animals are programmed to judge situations inferentially rather than optimise responses to stimuli, unless specific circumstances and the creature's mental abilities support a more systematic process (Chase, Hertwig & Gigerenzer, 1998; Hertwig & Herzog, 2009, pp.670-1; Stiegler & Gaba, 2015, pp.133-4). Simon (1959, pp.262-3) uses the terms 'satisficing' and 'maximizing' respectively, to distinguish a good enough response – based on aspiration level – from the best response. This is, in effect, an argument against reliance upon classical notions of 'rational choice' as a model for human behaviour.

While rational choice may be our preferred option for decision-making, Simon's theory posits that "maximizing is too hard for us", and that our needs are typically satisfied by ignoring probabilistic comparisons and instead preferring the first option that satisfies our requirements (Byron, 2005, pp.312-3). This has been expressed, by way of example, as us being satisfied when the proverbial haystack yields the first needle that is *sharp enough,* rather than demanding the best needle in the haystack (March & Simon, 1958/1993, as cited in Hertwig & Herzog, 2009, p.671). Similarly, in our quest for a life partner it may be more practicable to accept a few faults rather than wait for the most compatible man or woman. The alternative, that we approach all decision-making from a utilitarian perspective – e.g., finding, from those available, the best needle or the most compatible mate – requires us to attach "definite pay-offs" to every possible outcome of a decision (Simon, 1955, pp.103-4).

By settling for 'good enough' decisions, humans can choose from a "richer set of properties of the real world" (Smith, 1979, p.498). In this context, good enough is not just a consequence of our cognitive limitations, but also relates to the requirements of the task at hand (Gigerenzer & Brighton, 2009, p.108). Such requirements vary dynamically – reflecting the 'aspirational level' at a moment in time – as the individual engaged in the task adapts to the difficulty of the task and how easily he or she can identify suitable responses (Simon, 1955, p.111). This means that rather than continuing the task until every possibility is identified and evaluated, we accept the 'good enough' response in hand.

This human ability to produce and adopt a rough, good enough answer quickly is a vital aspect of our efficient functioning as an adaptive organism. There is simply not enough time in many instances to process even the available information and compute a response that can be empirically justified in some way. In this situation, we are in much the same position as the female peacock trying to choose a mate. Although it is not entirely clear why female peacocks prefer males that have more spectacular tail feathers, that preference is a good example of cognitive heuristics

at work: it would be impossible for the female peacock to evaluate all the individual characteristics that make for a good mate, so instead she instinctively engages in satisficing, basing her mate selection decision on the visual appearance her suitor's tail such as how many 'eye-spots' it has (Petrie & Halliday, 1994, pp.214-6). Where human or animal cognitive capacity to solve the complex problem of finding a good mate is "very small compared with the size of the problem", the optimal answer may be to adopt a practical solution where pragmatism substitutes for strict rationality (Simon, 1957, p.198, as cited in Barros, 2010, p.459). In this way, satisficing can be regarded as a type of 'bounded rationality' (Gigerenzer, 2004, p.65).

As I write this chapter I am engaged in satisficing. While my aspiration is to produce a comprehensive and well-thought-out discussion of the subject, the reality is that many factors conspire to prevent the most optimal outcome. When I search "cognitive heuristics" in Google Scholar (http://scholar.google.com), I find there are more than 9,000 results (9,110 as at 4 June 2017) and the conventional Google search system produces even more: 49,200 results. That's a lot of data.

If I were to read all 9,110 articles identified in Google Scholar at the rate of 10 articles each day, that would take me 2.5 years. The compromise, that I adopt instinctively and to some degree consciously, is to use the sources available to me to identify writers who appear to stand out as leaders in the field, and follow the breadcrumbs they have laid in order to gain a good enough understanding of the subject for present purposes. Like the female peacock I look (metaphorically) for the fanciest tail feathers proximate to me because my mental processes have concluded that this is the most efficient way to produce a chapter like this in a reasonable time-frame. I hasten to add that I try hard to base my assertions on plausible theories, even though I cannot possibly find every morsel of data that might be relevant.

Solving problems quickly and efficiently

When humans were primarily hunter-gathers, our ability to roughly estimate the distance between ourselves and an animal we were hunting may have been the difference between having food to eat or starving to death. The dynamics of trapping a wild beast seldom allow time to carry out careful measurements and calculations, or to refer to the wealth of human knowledge about the best ways to undertake this task. Instead, our forebears relied upon the hunting skills passed down by their elders, and on their innate abilities. Although nowadays we would prefer to be seen as rational purveyors of our learning, always acting consciously rather than instinctively, the reality is that almost all tasks we do in life – even making major decisions, like buying a house or choosing a life partner – are undertaken or aided by unconscious mental processes that we are not wholly aware of.

Without these unconscious mental processes, we would have to identify and weigh all relevant factors for every task, calculating the statistical likelihood of one or other decision leading to a specific outcome, and evaluating the costs and benefits of all possible outcomes, even though we do not have the memory capacity or cognitive ability to process all of this in real time (Johnson et al., 2013, p.476). Achieving survival despite these "biological limitations" has required that we evolve a smarter way to tackle this mammoth exercise.

Russell and Norvig (1995, p.94, as cited in Walton, 2010, p.163) suggest that heuristics were originally regarded as "problem-solving techniques" and later as "rules of thumb that domain experts could use to generate good solutions without exhaustive search." On this basis, we might say that whereas biases typically give effect to simplistic pre-existing preferences or aversions, cognitive heuristics are more sophisticated mechanisms to compute solutions in the absence of comprehensive data. However, both biases and cognitive heuristics are shaped by context and prior experiences to inform a unified scheme that is both finely

tuned and omnipresent. All day every day my biases and cognitive heuristics are quietly working away in the background.

My (limited) ability to catch a ball – a simple, everyday task that most of us can do if the ball is not too fast – relies on cognitive heuristics that automatically estimate the trajectory of the ball and where it is going to land. The unconscious processing capability that enables us to translate visual stimuli into real-time movements – a capability that humans share with most other predatory creatures – is so natural to us that we seldom give it much thought. It is, however, a remarkable capability in terms of its speed and efficiency that cannot be matched by any computer so far designed by humankind.

Our mental search processes usually stop working when we instinctively consider that we have enough data in hand to answer the problem at hand (Barros, 2010, p.465). In the current era of 'big data', this heuristic phenomenon means that our estimations can be wildly wrong because we are relying on a very small sample of all relevant data. Nonetheless, there is some evidence that our adaptability is coming to the rescue. For example, in relation to our ability to determine the veracity of online data, Metzger and Flanagin (2013, p.214, citing Metzger, 2010) suggest that where we have neither the time nor the cognitive resources to undertake a systematic review of data produced by an online search, we may instead "invoke a diversity of heuristics to evaluate credibility, loosely titled reputation, endorsement, consistency, self-confirmation, expectancy violation, and persuasive intent." These heuristics – that collectively we may call 'credibility heuristics' – probably derive from the conjunction of existing capabilities, which exemplifies the way that multiple heuristics can work together to optimise decision making (Metzger & Flanagin; Gigerenzer & Todd, 1999, as cited in Metzger & Flanagin).

However, there will always be situations when our cognitive heuristics will lead us into the swamp if we let them. This requires us to adopt new strategies, based on understanding our cognitive heuristics, compensating for their shortcomings

and developing their strengths (Metzger & Flanagin, 2013, p.218). For humans, the use of such strategies may be regarded as an essential life skill. In particular, the engagement of a cognitive heuristic as an automatic response to stimuli may be mediated by our "awareness of a potential influence" such as where we consciously seek to avoid a particular stereotype (Bargh, 2014, pp.12-13). Bargh (p.13) comments:

> "One cannot be aware of the actual occurrence of accessibility or stereotypic influences because of the fast, effortless, and immediate (i.e., preconscious) way in which those mental structures capture and interpret relevant environmental input. Nonetheless, through education and other consciousness-raising techniques, one can become aware that one might be influenced."

A good example of a situation where "consciousness-raising techniques" may help us is when we are undertaking a task that requires us to focus our attention on a specific activity. Inattentional blindness is a well-known phenomenon where focusing on one activity can result in failure to see a non-salient feature or event even though it is in plain sight. In a study by Simon and Chabris (1999, pp.1066-9), subjects were asked to watch a video recording of a ball game and mentally keep track of the number of ball passes by a nominated team. After about 45 seconds of play, a person dressed as a gorilla walked through the area of play, but subsequently 46% of those present did not recall seeing the gorilla at all (this was the average of various test conditions). In another study, 60% of radiologists given x-rays to review failed to see that the patients' collarbones were missing (Potchen, 2006, as cited in Drew, Vo & Wolfe). In their own study, Drew, Vo and Wolfe inserted a gorilla image into chest x-rays, then asked radiologists to click on lung nodules that appear in the x-rays as small white ovals. Only 4 of the 24 radiologist participants reported seeing the gorilla (p.1850). Drew, Vo and Wolfe (p.1852) conclude that the problem arose because the radiologists were focused on the nodules and were not expecting to see the gorilla image.

A force for good and bad

Our cognitive heuristics are still at work even if we are not under any pressure to make a quick decision. This is partly because we are innately inclined to use cognitive heuristics even when other processes are available, and also because we are generally overconfident about our innate ability to assess data and reach conclusions. In particular, as Simon (1955, 1959, 1979) identifies, we are very poor at estimating probability. However, it is important to appreciate the difference between the operation of cognitive heuristics and the impact of 'motivational factors' on our decision-making (Tversky & Kahneman, 1974, p.1130). Tversky and Kahneman suggest that motivational factors may include 'wishful thinking', rewards, and penalties, where these distort our judgement. The significance of this distinction is apparent when we consider the gambler's fallacy where the desire to win may be a factor in favouring an option that is believed to be a 'good omen', although there is no doubt that cognitive heuristics play a significant and often decisive role in our estimations of measure and probability.

Even when individuals are incentivised to expend extra cognitive effort, the influence of biases and cognitive heuristics persists largely unchanged. This raises the question whether it would be apt to regard "human decision makers as systematically flawed bumblers" (Ortmann & Hertwig, 2000, as cited in Gilovich, Griffin & Kahneman, 2002, p.8). However, this is a dubious critique that ignores the vital role our cognitive heuristics have played in human survival as a species, fine-tuned by evolution to promote our longevity and reproductive success. Although we may now find these cognitive mechanisms inconvenient on occasions, they are properly regarded as the efficient automation of cognitive tasks that would in other circumstances overwhelm us, and there is really no possibility at all that we could function without them.

In fact, studies suggest that less information and less computation may in some circumstances lead to greater accuracy, which has been dubbed the 'less-is-more' effect (Gigerenzer & Gaissmaier, 2011, p.453). However, this is not a linear

effect where accuracy increases as information decreases, but rather represents "the existence of a point at which more information or computation becomes detrimental, independent of costs" (Gigerenzer & Brighton, 2009, pp.110-1). The problem remains that while in some scenarios that "point" falls within the operational ambit of our cognitive heuristics, very often it does not and so in that case we make poor decisions that could have been improved with more information and more conscious thought.

Nonetheless with some effort we can be trained to approach problems more logically, for example employing greater 'statistical sophistication', and can thereby reduce our individual susceptibility to some biases to some degree (Tversky & Kahneman, 2002, p.231). Moreover, in the context of society and community we can and do build comprehensive systems and safeguards into our daily lives that save us from the worst consequences of our biases. For example, while some aspects of our individual decision making associated with buying a new car – e.g., choice of colour – may involve our cognitive heuristics, and specific features that we require the vehicle to have will usually be at the forefront of our consciousness, critical issues such as safety and fuel efficiency are addressed by relevant laws that recognise our inability to adequately assess those factors heuristically or otherwise.

2.3. Informal fallacies

In theory, 'biases and cognitive heuristics' could include any cognitive process that enables us to make rapid decisions when we perceive external stimuli. Arguably, it should not make a difference whether the cognitive process incorporates pre-existing notions of efficient computation, or inefficiently allows logic errors to lead us astray. Either way, our brain is processing external stimuli unconsciously unless we assert conscious control. However, there is some controversy about including informal fallacies in a study of biases and cognitive heuristics, and so it is useful to briefly examine the contribution of logic to our cognitive activity.

By way of definition, informal fallacies involve the misapplication of inductive logic. For example, if we accept an assertion as truth only because the assertion was made by a person in authority, we may have fallen victim to the fallacy *argumentum ad verecundiam,* which broadly translates to 'fallacious appeals to authority'. Similarly, if our strong desire for an outcome encourages an unrealistic belief that undertaking a specific task will achieve that outcome – a kind of 'wishful thinking' – we may be affected by the fallacy *ad consequentiam* (an argument based on consequences). An example of the fallacy *ad consequentiam* is when we throw a coin into a wishing well and make a wish for new car or a pay rise.

Many fallacies also have a negative form. For example, the fallacy *ad consequentiam* can be expressed as the belief that an assertion is false because an outcome is not desirable. In the U.K. in 1974, six men were arrested for planting bombs in Birmingham that killed 21 people. Following their trial, the six were sentenced to life imprisonment, but then applied to have the convictions overturned. In his judgment in the Court of Appeal refusing that application, *McIlkenny v Chief Constable of the West Midlands* [1980] QB 283, Lord Denning said: (original judgment at p.22)

> "Just consider the course of events if this action were to proceed to trial... If the six men win, it will mean that the police were guilty of perjury, that they were guilty of violence and threats, that the confessions were involuntary and were improperly admitted in evidence: and that the convictions were erroneous... This is such an appalling vista that every sensible person in the land would say: It cannot be right that these actions should go any further."

Although Lord Denning (d. 1999) is generally regarded as one of England's greatest judges, his comments in this *McIlkenny* judgment have been widely criticised as fallacious, and the decision was subsequently overturned by the House of Lords in *Hunter v Chief Constable of the West Midlands Police* [1982] AC 529 (The Guardian, 1999; Kennedy, 2011, pp.128-9; Wikipedia, 2018).

Brief introduction to <u>formal</u> logic and fallacies

Cognitive heuristics and biases do not involve formal logic. Formal logic can be expressed as a way of identifying how a set of known information can lead to a logical conclusion. For example, if group A contains all members of group B, and all members of group A have a specific characteristic, we can say that all members of group B have that same characteristic. We also call this 'deductive logic' or 'deductive augmentation'. 'Augmentation' refers to the manner in which the known information is parlayed into an argument. Deductive augmentation uses logic to make a truthful argument. Unfortunately, there are several problems with deductive logic as a universal approach to making decisions:

- To make a deduction we need relevant facts which may not be available.

- Although many deductive scenarios are similar to my group A and group B example above, most are far more complex making it hard to work out the logical answer.

- Often we have extensive information to inform a deduction, but it may not be the deduction we need. This is frequently the case where we can see a correlation between two events but cannot deduce which event (if either) is causative.

McFerran (2015, p.4) refers to Aristotle's "classic example" of a deduction based on formal logic:

> "All men are mortal and Socrates is a man, so Socrates is mortal (or A –> B and C –> A, so C –> B)."

We don't usually have much difficulty with logic at this level. But how about this example: (McFerran, p.3):

> "All Republicans are conservatives, and Senator Short is a conservative, so he must be a Republican."

You can see the illogic here. We call this example a 'fallacy', although it must be regarded as a 'formal fallacy' as distinct from an informal one because it is really a 'pattern mistake' – i.e., if we adopt a symbolic approach we can see that the error is in the pattern itself: it is incorrect to say that B –> A just because A –> C and B –> C (McFerran, 2015, p.4). This pattern can never be deductive. However, if we fix the pattern so it is logically correct, the outcome can then be deductive subject to the information provided. With this McFerran example as it stands, the most we can say is that *it is possible* that Senator Short is a Republican, but it also remains possible that he is not. We don't have enough information. Even if we knew in addition that most conservatives are Republicans, we still could not say for sure (deduce) that Senator Short is a Republican.

Informal or 'inductive' logic

If it seems very probable to us that conservatives are more likely to be Republicans than Democrats, we could perhaps make an 'inductive argument' that Senator Short is a Republican. In some cases, such an argument would be a strong one, but it depends on the content and the context. There are many circumstances where such arguments are not strong at all – e.g., they may rely on 'inductive augmentation' that is unreliable ('unacceptable augmentation') which in some cases could involve 'informal fallacies'.

The strength of induction derives from the content of the information, and sometimes its context, so inductive error usually arises from misuse of the content. Walton (2010, p.175) identifies several fallacies – e.g., "equivocation, amphiboly, accent, begging the question, [and] fallacies of irrelevance, like red herring and wrong conclusion" – that could not be regarded as heuristic, but lists the following as potentially capable of being represented in heuristical terms:

– *ad misericordiam* (argument based on distress)
– *ad populum* (argument based on popular opinion)
– *ad hominem* (e.g., circumstantial arguments)
– *ad baculum* (argument based on threat or fear)
– straw man (argument based on commitment)
– slippery slope schemes
– *ad consequentiam* (argument based on consequences)
– *ad ignorantiam* (argument based on ignorance)
– *ad verecundiam* (argument based on expert opinion)
– *post hoc* (causal argument based on correlation)
– composition and division (argument based on composition or division)
– false analogy (argument based on analogy)

Conventionally, a reasoning error based on flawed logic has been described as a fallacy and therefore distinguishable from errors arising from cognitive heuristics or biases (Ross, 2014, p.4). A simple logic error is not, in Ross's view, enough to qualify as a heuristic outcome. However, Johnson (2005, as cited in Cummings, 2014, p.5) identifies that heuristic processing as a 'parascheme' works alongside more systematic mental activities, taking over when we receive 'cues' that can inform a snap judgment. Systematic mental activities typically involve formal logic to analyse content and reach a logical decision when that is possible. Walton (2010, p.x) uses the term paraschemes to represent the heuristical structure facilitating a "fast and frugal" approach to making decisions that our cognitive resources would struggle to address systematically in any practicable time-frame if at all (Gigerenzer et al., 1999, as cited in Walton, p.161).

The mediating role of cognitive heuristics

Walton (2010, as cited in Cummings, 2014, p.5) is credited by Cummings as the first writer to treat certain informal fallacies as cognitive heuristics. Walton (2010, p.160) describes heuristics as the "psychological dimension" of some fallacies,

and suggests that three specific heuristics – deference to expert opinion, accepting not false as truth, and avoiding fear – may be regarded as "a mediating concept between the notions of fallacy and defeasible argumentation scheme." The three examples given by Walton may be analysed in the following way: (pp.163-4)

- deference to expert opinion: if a person recognised as an expert says that A is true, A must be true. This corresponds to the fallacy *argumentum ad verecundiam,* or 'fallacious appeals to authority'.

- accepting not false as truth: if A is not false, it must be true. This corresponds to *argumentum ad ignorantiam,* or 'arguments from ignorance'.

- avoiding fear: if consequence C induces fear, avoid activity that might lead to consequence C. This corresponds to *ad baculum* arguments, or 'arguments that appeal to threats'.

You can see immediately that these examples do not involve deductive reasoning, although they may be persuasive in some circumstances. Considering deference to expert opinion, if the opinion-giver has the requisite expertise and we have a basis for having confidence in his or her opinions within that field of expertise, then accepting the opinion as fact may be pragmatic. In fact, we do this every day. If our car mechanic says we need new spark plugs we usually accept this, just as we may accept legal advice from a qualified lawyer operating in the relevant area of law. There is of course no guarantee that either advice is correct, and sometimes we will want to obtain a second opinion, but otherwise we may rely on the expert because that makes sense to us. Even where we have very little information to rely on when evaluating our expert, our biases will assist us to arrive at a decision. In some cases, it will be a wrong decision, perhaps disastrous for us, but statistically we may be just as well off employing our cognitive heuristics.

In practice, our cognitive heuristics operate at a level of complexity that belies their apparent simplicity. Although fallacious appeals to authority may in some cases deserve condemnation, behind the scenes our heuristics are considering a sophisticated set of questions: (Walton, Reed & Macagno, 2008, p.310, as cited in Walton, 2010, p.165)

- expertise: is our opinion-giver a known 'expert source'?
- field: is our opinion-giver a subject matter expert?
- opinion: how does the opinion relate to the proposition?
- trustworthiness: is the opinion-giver known to be a reliable source of opinions?
- consistency: is the opinion consistent with other expert assertions?
- evidence: is the opinion supported by evidence?

Thinking about our car mechanic, we might consider all these factors:

- expertise: is our mechanic a known 'expert source' in relation to car engines?
- field: is our mechanic an expert in relation to spark plugs?
- opinion: what does our mechanic say about the spark plugs and what do we infer from that?
- trustworthiness: is the mechanic known to be a reliable source of opinions?

- consistency: is the mechanic's opinion consistent with other expert assertions? (here we might refer for example to data about the usual life of spark plugs, if that data is known to us)

- evidence: is the opinion supported by evidence? (perhaps the mechanic has shown us one of the old spark plugs and pointed out the contact erosion that supports his or her opinion)

Recently, in 2017, I had an experience very much like this, where the mechanic looking at my car suggested that a drive belt needed replacing. I don't recall a conscious process of going through all the above steps, but when I check through them in my mind now I realise that all these factors were present including that last one: he showed me the worn belt. What I do recall is that when he asked if I wanted the belt replaced, without a moment's hesitation I replied "yes."

The confidence-giving role of cognitive heuristics

Although we may get to know some tradespeople and professionals quite well, in many situations we will have very limited information about an authority-figure whose advice we are needing to rely on. For example, if we see a doctor about stomach pain we may have no idea about the doctor's skills in that particular area or whether he or she is trustworthy or consistent. Perhaps we have never seen the particular doctor before. Nonetheless, when the doctor pronounces a diagnosis we seldom question it. To some extent we are relying on broader confidence in doctors as a profession, but we also know that doctors sometimes make horrendous mistakes. Relying on a doctor in these circumstances is fallacious, a clear case of *argumentum ad verecundia,* but we do it nonetheless. Put simply, our cognitive heuristics give us confidence – whether it is justified or not – that the diagnosis is correct, and we wander off to the pharmacy with our prescription happy in the belief that tomorrow we will be feeling much better.

In a similar way, we are often heavily influenced by popular opinion. *Ad populum* arguments are implicitly imprudent, even though the bias we have towards popular opinion is most often harmless and may even lead us to a sensible approach depending on our interests and objectives.

In relation to arguments based on ignorance, *argumentum ad ignorantiam,* Walton (2010, pp.175-6) gives the example of Roman soldiers receiving medals:

> "… there is no evidence that Roman soldiers received posthumous medals in war, only evidence of living soldiers receiving such awards. From this lack of evidence, the conclusion can be drawn by inference that Roman soldiers did not receive posthumous decorations in war... It is commonly called the 'lack of evidence argument' in the social sciences or the *ex silentio* argument in history, where it is regarded as a reasonable but defeasible argument."

A very practical example of this 'lack of evidence argument' can be found in the U.K. Government's reassurance to the general public about eating beef following the bovine spongiform encephalopathy (BSE) crisis, which relied entirely on the lack of evidence that beef was not safe for consumption (Cummings, 2014, p.7). BSE is popularly known as "mad cow disease" and can be transmitted to humans in some circumstances.

In our minds, a bias may be regarded as a preconscious argument that appears to be a better argument than it really is (Hansen, 2002, p.152, as cited in Walton, 2010, p.180). While drawing a clear demarcation between fallacy and bias might be an impossible task, certainly one outside the scope of this study, there seems little doubt that some fallacies are simply "a very natural way of unreflective thinking" (Walton, 2010, p.181).

Chapter 3 – Focus on Confirmation Bias

One day my car had a flat tyre. Attempts to fit the spare wheel were unsuccessful, so I walked down the street to a vehicle repair shop that I had not visited previously. A young lad in the office said I would have to pay a "$60 call-out fee" even though my vehicle was clearly visible less than 100 metres up the road from the repair shop. Fuming about this outrageous fee (that I refused to pay) as I walked back to the car alone, I removed the old wheel and rolled it back to the repair shop to be fixed. I subsequently told my friends about the "idiot" at the repair shop. This soon became the "idiots" at the repair shop, and then a vow that I would never ever take my car there again. All perfectly logical of course.

However, there's another side to this story that I reluctantly share. About a year later I urgently needed a vehicle fitness check (called a pink slip in some states). I'd left it too late, and most workshops were closed or booked up. Grinding my teeth a bit I resigned myself to the misfortune of breaking my vow, and drove into the original repair shop. When they looked up my car on their computer and confirmed my status as a "valued existing customer" I was welcomed like a long-lost friend. Moreover, I was reminded that the first time I went to them I'd actually purchased three tyres from them. I had no recollection of this aspect of that transaction, although clearly the tyres had performed as well as they should have. I only recalled my irritation at how I was quoted $60 for a call-out to a vehicle that was just a block away. This last visit they were nothing short of professional and helpful in giving the vehicle its fitness check.

However, I still couldn't quite get rid of the notion that they were "idiots"... It was a bit irrational, but my initial belief about this repair shop hung around in my head, a black mark that was (probably) as unreasonable as it was persistent.

3.1. Forming a belief

Nickerson (1998) records that the English philosopher Francis Bacon identified as far back as the year 1620 that we tend to celebrate successes and forget failures, which is a typical characteristic of the confirmation bias phenomenon. A more refined definition of confirmation bias is provided by Nickerson himself, who took the term to "represent a generic concept that subsumes several more specific ideas that connote the inappropriate bolstering of hypotheses or beliefs whose truth is in question" (p.175). Noting the lack of intentionality that typically distinguishes unconscious 'case-building' from the more deliberate biases of attorneys and debaters, Nickerson suggests that a fundamental aspect of confirmation bias is: (pp.175-6)

> "unwitting selectivity in the acquisition and use of evidence... without intending to treat evidence in a biased way or even being aware of doing so."

The starting point is the holding of a belief. In some situations, such as that facing a juror in a criminal trial, the belief we are concerned about is the one that is formed before all relevant information is to hand – i.e., failing to wait until all the evidence has been heard, the lawyers' summing up has been completed, the judge has given directions, and the jury is about to retire to reach a verdict. In relation to 'belief', McKay and Bennett (2009, p.493) propose that:

> "A belief is a functional state of an organism that implements or embodies that organism’s endorsement of a particular state of affairs as actual."

However, actuality may never be known in some cases, and it is hard to see why a belief must have any link to actuality at all. Belief is, I suggest, simply a matter of having an opinion about a situation or state of affairs, which may be a strongly-held opinion or perhaps may be quite tenuous. Moreover, a belief can be formed in the context of knowing that it may possibly be erroneous.

3.2. Persistence of belief

I am writing the first draft of this paragraph on the day after the inauguration of Donald Trump as America's 45th president, a man who has been accused of many untruths (e.g., Stanley, 2016; Kessler & Kelly, 2018). However, as Stanley asserts, it is quite likely that Mr Trump reaches and proclaims beliefs in a way that is knowingly "insensitive to reality" and demonstrates a kind of laziness about the truth that is both convenient and cynical. Convenient because it enables him to define a reality that aligns with the ambitions of his supporters. Cynical because it takes improper advantage of electors' susceptibility to "authoritarian rhetoric." What Mr Trump's approach to truth speaks to is the malleability and attributional quality of belief. A reasonable inference we can take from Mr Trump's statements is that he persistently believes he is hugely successful.

If we consider the situation of the juror who has heard the first few days of a long trial and has already formed a belief about the guilt or innocence of the accused, we might be tempted to regard that simplistically – e.g., "the accused is probably guilty of the crime" – but a nuanced belief is more likely, such as "having heard the evidence so far, it seems probable that he committed the crime, but having seen the accused in the dock I think the crime would be out of character." This construction (or deconstruction) opens the door to an infinity of possibilities, in terms of how confirmation bias may affect the juror's acquisition and interpretation of information from that point onwards in the trial.

This can be contrasted with '*commitment* bias' a.k.a. 'escalation of commitment bias', where the holder of a belief is strongly committed to it even where maintaining the belief leads to conduct that is irrational, at least in social contexts (Painter, 2004, pp.405-6; Back, 2010, p.27). Escalation of commitment bias has been defined as "persistence... with a failing course of action" (Brockner, 1992, pp.40 & 52-3; Garland, 1990, p.731; Markovitch, 2014, p.302).

3.3. Operationalisation of confirmation bias

Manifestations of confirmation bias can give the impression that entirely different factors are in play, but a critical factor is that the process where it arises is always a form of 'decision making', whether it leads to a better (or worse) test outcome, or a life-changing decision to follow a particular course of action (Panda, 2014; Rassin, 2008, p.88). Rassin notes that confirmation bias both affects our perceptions of stimuli, and how we evaluate perceived stimuli in the context of a specific theory or belief. However, the use of 'confirmatory decision strategies' is strongly dependent on the circumstances of the decision making, even though some individuals may be more prone to using them than others (p.92).

There is also an 'accumulation effect', where information received and evaluated over time eventually reaches a point when "a critical event occurs that halts the decision process" – e.g., due to tiredness, the acquisition of significant new information, or a deluge of information that exceeds the individual's cognitive capacity to process it (Devine, 2012, p.25). As discussed in the next section, I don't agree wholly with Devine's theorisation on this specific point, although there's no doubt that the accumulation of information and the formation of biases can be related.

Later in his book, Devine (2012) discusses joinder of charges in criminal proceedings – i.e., cases where a person has multiple charges heard by a jury (or judge alone) at the same time – and the phenomenon first demonstrated by Kerr and Sawyers (1979, as cited in Devine) and Horowitz, Bordens, and Feldman (1980, as cited in Devine) where the conviction rate for a 'focal charge' increases significantly if a jury is asked to consider a second charge at the same time. However, having more than one alleged offender tried at the same time (in the same trial) has not been found to significantly affect the outcome compared to having separate trials for each defendant (Leipold and Abbasi, 2006).

3.4. Proneness to confirmation bias

Although Devine (2012, p.25) suggests that confirmation bias arising from an 'accumulation point' -- i.e., where information favouring a particular belief has become persuasive -- effectively "halts the decision process", this is not supported by Devine's own work or by other studies identified so far. In a lengthy trial, it is virtually inconceivable that a typical juror reaches a point of complete resistance to (for example) highly compelling contrary evidence, due to the decision-making process having *halted* for that juror. That situation could only arise in circumstances of 'escalation of commitment bias' or some other entrenched bias or other factor that is not susceptible to rational thinking. However, any such factor would almost certainly involve the juror having a mental investment in the trial outcome, representing a perverse partiality that is outside the scope of confirmation bias. It is more likely, and consistent with other studies, that the accumulation point identified by Devine is the threshold when a mental model is formed, and that the subject is then *susceptible* to confirmation bias favouring that model. Under the heading 'accumulation prejudice', Goodman-Delahunty, Cossins and Martschuk (2016, pp.46-7) comment as follows regarding joinder of charges in criminal proceedings:

> "As Leipold and Abbasi submitted, intuitively at least, the more counts the defendant is charged with, the more likely it is that a jury will conclude that the defendant is guilty of *something.* The hypothesis is that juries are prone to conclude that because there are multiple charges or multiple witnesses supporting the prosecution’s case, the defendant is guilty."

I think the words "prone to conclude" represent a good understanding of the point that Devine is apparently trying to make, which is that with cognitive bias in play – e.g., when significant information is received by jurors at an early stage of the case, such as the number and type of charges faced by the defendant – the die is loaded against the defendant, increasing the likelihood that some of the jurors

will reach a belief about the defendant's guilt well before all of the evidence in the case has been heard.

3.5. Persistency over time

My objective in this chapter has been to convey some of the complexity associated with confirmation bias, even though it is often presented as conceptually simple. It is of particular relevance to my thesis that once a theory or belief has formed in an individual's mind, it can be highly persistent *over time.* This means, for example, that child protection caseworkers responsible for evaluating the risk to a child may be biased in their assessment of specific risk factors due to their previous experiences and beliefs they have formed about child safety over a long period. Similarly, children who experience abuse are very likely to form persistent beliefs about the impact of that abuse and what kinds of responses to it are appropriate. These beliefs will then unconsciously inform the acquisition and use of information over the individual's lifetime, perhaps leading to a tragic misunderstanding of the patterns of that lifetime and of the origins of adversity in it. However, this outcome and the confirmation bias phenomenon need to be considered against the much broader tableaux against which our human existence plays out. While personal tragedy is life at its worst for the affected individual and those close to him or her, such tragedy can only be evaluated by an empirical measure of life in all its fullness including its best moments. If we focus only on the negative we may fail to see that our cognitive processes, imperfect as they are, have taken us to heights that no other species known to us has every achieved.

3.6. Types of confirmation bias

Much has already been written and published about the different forms of confirmation bias, and it is not necessary for me to cover them here in great detail. A modest summary is sufficient for present purposes, using Nickerson's (1998) categorisation more-or-less – i.e.:

– Seeking and interpreting information according to our theories or beliefs
– Flawed reasoning
– Primacy effect and how a belief persists
– Evaluation of our own judgment

To that categorisation I have added two further categories for biases that appear to meet the definition of confirmation bias, but which are not easily allocated to one of Nickerson's category areas:

– Low signal-to-noise ratio and change blindness
– Delusional thinking

In establishing the last category, I acknowledge that not all delusional thinking involves confirmation bias, and that the boundary between confirmation bias and sustained delusional thinking is uncertain.

Seeking and interpreting information according to our beliefs

Nickerson (1998) notes that confirmation bias often involves actively seeking evidence that supports a particular belief or giving undue weight to an interpretation that is consistent with the belief. For example, if a person believes in the existence of a deity, he or she is likely to seek out evidence supporting the belief, and may interpret ambiguous or contrary evidence in a way that also supports the belief. Everyday life abounds with opportunities to attribute good fortune to a kind and benevolent deity, while regarding misfortune as punishment for failing to live up to the same deity's expectations. The notion that all or most such events are better attributed to chance is hard for most people to grasp, as our daily lives are built on the premise that most events are able to be (and therefore are) controlled by ourselves or other intelligence.

When we observe an event that is consistent only with our belief, not with any competing or inconsistent belief, that observation is said to be "diagnostic" – i.e., it diagnoses the truth of our belief. Accordingly, an observation that supports an alternative belief is not diagnostic. This is a critical factor in assessing belief, because our unwavering adherence to a particular belief means that we are not open to any possibility that there is another, inconsistent belief that is more likely to be true. We therefore fail to test that other belief, and in fact tend to ignore that possibility entirely. We are rather adept at this. Kuhn (1989, p.677) conducted a study where subjects were provided with information that contradicted their belief about a matter. According to Kuhn:

> "Subjects either failed to acknowledge discrepant evidence or attended to it in a selective, distorting manner. Identical evidence was interpreted one way in relation to a favored theory and another way in relation to a theory that was not favored…"

Significant, the subjects remembered the information as more in line with their belief than it was, distorting the information or simply not recalling it correctly so as to facilitate the maintenance of their belief.

We also tend to see an 'illusory correlation' between unrelated objects or events, such as where we see specific objects or events in the same context, or where we perceive that a correlation "ought to exist" (Nisbett & Ross, 1980, p.97, as cited in Nickerson, 1998, p.183). The illusory correlation may then inform our understanding of particular relationships, so that our learned behaviour is changed to reflect that new understanding (Goldberg, 1968, 493, as cited in Nickerson, p.183). This is closely related to stereotyping, where our belief about a relationship between particular factors – e.g., poverty and crime – leads to our being "more likely to seek and find evidence to support the belief than evidence to oppose it, somewhat independently of the facts" (Nickerson, p.183). This 'illusion of consistency' reflects our being more attentive – Nickerson uses the

word "attuned" – to the information that supports our belief, and less attentive to other information.

Flawed reasoning

Nickerson (1998) describes the classic experiment conducted by Wason (1966; 1968, pp.274-5) involving two-sided cards. It is a simple but – in its outcome – disarming experiment that uses just four cards to establish our almost universal failure to attend to a selection task. Nickerson describes the experiment as follows:

> "[P]articipants see an array of cards and are told that each card has a letter on one side and a number on the other. Each of the cards they see shows either a vowel, a consonant, an even number, or an odd number, and participants are asked to indicate which cards one would have to turn over in order to determine the truth or falsity of the following statement: *If a card has a vowel on one side then it has an even number on the other side."*

Consider this example: The visible cards are [E] [X] [6] [9]. Studies consistently show that most people will select only card [E], or else cards [E] and [6]. They ignore cards with a consonant or an odd number because those are apparently not related to the problem at hand. Most people struggle to see that proving that [E] is backed by an even number does not prove the statement, as one of the other cards may refute it. Those who select [E] and [6] remain ignorant of whether the statement is disproved by card [9], as it would be if the back of card [9] was a vowel. The two cards that can prove the statement are [E] and [9] together, because it does not matter what is behind [X] or [6].

This 'card trick' is often seen by subjects as more interesting than significant, but it goes to the heart of our reasoning process when presented with a situation that requires us to decide whether a statement or assertion is true or false. We tend to focus on the factors that are most apparently relevant to our belief about the

situation, and accordingly fail to attend to factors or options that may give us a more accurate outcome or understanding.

Although the broad range of findings in studies using variants of Wason's card experiment are not all explained by confirmation bias, the experiment demonstrates the existence of the phenomenon and its "substantive role" in many cases (Nickerson, p.187).

Primacy effect and how a belief persists

The 'primacy effect' refers to how the first observations or information received by an individual about an issue tend to inform the interpretation of subsequent information on the issue. For example, Bruner and Potter (1964, as cited in Nickerson, 1998) found that when they showed subjects a series of images, the first of which was unfocused and unrecognisable, that first image nonetheless promoted a belief that persisted even as the images progressively revealed a particular object – i.e., those shown the initial image struggled to recognise the object when it became clearer, while those who were not shown the initial image were able to identify the object easily from the later images.

In another study, observers watching a group of problem solvers work through a variety of problems rated the problem solvers more highly if they solved more earlier problems and less later problems, than those who did the reverse. The initial impression persisted despite the problem solvers' actual overall results (Jones, Rock, Shaver, Goethals & Ward, 1968, p.336). In a study by Ross et al. (1975, p.882) the subjects were asked to assess the genuineness of suicide notes. Some were then told that their assessments were better than the average, others were led to believe the opposite, and a third group were told that they were average performers. Subsequently all of the subjects were told that their outcomes were "determined before they entered the experiment, that they had received feedback unrelated to their actual performance" (Ross et al., p.882). Nonetheless, when asked to state their competence at making such

assessments, the subjects who had been told they performed above average rated themselves more highly than those who received the negative feedback.

The persistence of a belief can become more entrenched as subjects – or individuals acting similarly in real life situations – look for and find further information that tends to confirm the belief, so that the belief is not merely based on "the fraudulent feedback" received at the outset but is also supported by the subsequently acquired information (Nisbett & Ross, 1980, as cited in Nickerson, 1998, p.188). This suggests that even in research situations, where it is intended to debrief subjects who have been given false information in the course of a study, there is a real risk that belief in the false information will persist beyond the scope of the study and the debriefing.

The persistence of belief is also assisted by what Rossmo (2016, p.218) describes as 'groupthink', which involves individual constraint in challenging the prevailing theories or beliefs within the group. Citing Janis (1982), Rossmo suggests that individuals in the group may become close-minded to the possibility of different theories or beliefs being true, and give undeserved weight to the apparent consistency of opinion within the group.

Evaluation of our own judgement

Earlier I discussed the study by Ross et al. (1975), where the subjects were asked to assess the genuineness of suicide notes. In that study, predetermined feedback was given to the subjects. However, in real life we are often placed in the position of having to evaluate our own competence at performing a task, which can be referred to as a 'confidence judgment'. Many studies have found that we are more likely to over-rate our competence than to under-rate it (e.g., Einhorn & Hogarth, 1978; Fischhoff, 1982; Lichtenstein & Fischhoff, 1977; all as cited in Nickerson, 1998) – the 'above-average effect' (Kruger & Dunning, 1999, p.1122) – and demonstrate reluctance to recant an unrealistic rating even when forced to undertake a more robust self-evaluation (Nickerson).

This 'above-average effect' even affects subject matter specialists – for example doctors and psychology students who in different studies were found to exhibit overconfidence in their ability to achieve specific outcomes – particularly where a prediction or diagnosis is not founded on statistical data (Keren, 1987, p.100, citing Christensen-Szalanski & Bushyhead, 1981, Lusted, 1977, and Lichtenstein and Fischhoff, 1977). It is therefore ironic that such professionals may tend to have more confidence in their personal assessment than an assessment derived from statistical data, even when the statistical data challenges the personal assessment. Nickerson (1998, p.189) theorises that an individual's overconfidence in his or her own knowledge is sustained by the way in which memory retrieval follows the line of a "tentative answer" to produce consistent information, while other available information that may be inconsistent is not found.

While an individual's tendency to regard him or herself as above average is undoubtedly a cognitive issue, it is not necessarily connected to confirmation bias. However, failure to recognise the relationship between one's own shortcomings and a pattern of errors may be indicative of 'self-erasing errors' (Redelmeier & Scales, 2015, p.280). Redelmeier and Scales (citing Sherman, 1980) define this kind of confirmation bias as arising "when people do not know enough to realize they do not know enough."

Low signal-to-noise ratio and change blindness

Although the term 'signal-to-noise ratio' is widely used to denote potential for confusion between desired information and background 'noise', the concept can also be applied to our human perceptive processes. This may, for example, be related to persistence of belief, where later information is not heard so well as to force a re-evaluation of an earlier-formed belief, so the problem is one of attention rather than availability. However, the 'noise' category also stands on its own because persistence of belief may in some cases defy the logical evaluation of

information received later, while persistence of belief arising from confusion between different sources of information or even misunderstood information from the same source is not necessarily illogical. Rossmo (2016, p.224) proposes that deficient attention to the bigger picture often arises from the combination of confirmation bias and low signal-to-noise ratio.

It is easy to think of illustrations of how signal-to-noise ratio might conspire with confirmation bias to cause a wrong outcome in the end. For example, police investigations often involve evaluating hundreds or even thousands of leads, some of which will be dead-ends that should be weeded-out before they consume too much time. However, just one error – e.g., a potential suspect being overlooked – could be disastrous. In this illustration, the sheer volume of leads anticipated as dead-ends inherently introduces 'noise' that may obscure a potential suspect within the leads.

Noise can also come from external factors, such as public pressure. For example, where a serial killer is on the loose, public pressure may demand that the killer be brought to justice. The consequent pressure that policing management bring to bear on investigating officers amounts to a form of noise that impacts the methodical evaluation of evidence and encourages officers to form conclusions that may be wide of the mark.

In our private lives too, noise encourages confirmation bias. For example, a purchase decision may be influenced by advertising; if we constantly hear on television and radio that the ABC brand gismo that our kids want for Christmas is the most popular toy ever, then on Christmas eve as we do our last-minute shopping this noise crowds out the notion that another toy or brand may be just as good at half the price.

Low signal-to-noise ratio can also involve 'noise' that is linguistic or malformed, such as where the information given to a person is couched in terms that obscure its true meaning. An example might be where a risk warning is entirely true, but

the words used to express it mislead the individual hearing the warning. Gigerenzer (2001, pp.93-94) tells the story of a psychiatrist who warned patients prescribed the anti-depressant Prozac that they had a 30-50% risk of developing a sexual problem while on the drug. The information was correct, but was nonetheless misunderstood by most patients who perceived that they would have sexual problems such as sexual dysfunction or low libido more than a third of the time. Reference instead to "three out of ten patients" developing a sexual problem meant the warning was more easily understood by patients.

Delusional thinking

Building on their definitions of belief and misbelief, McKay and Dennett (2009) explore the proposition that humans are engineered by evolution to formulate beliefs approximating current beliefs informed by available evidence, and why this proposition is subject to "routine exceptions" (p.494). McKay and Dennett distinguish between two types of misbelief, being malfunctioning of the systems that formulate beliefs, and misbelief that arises from normal functioning. The first of these types may also be referred to as arising from 'doxastic dysfunction' – i.e., as being "the faulty output of a disordered, defective, abnormal cognitive system" (p.496). They suggest that *delusions* fall into this type

Mishara and Corlett (2009) suggest that delusions can be attributed to:

> "... aberrations of perception which occur when neuronal noise induces mismatches between expectancy (Bayesian priors) and experience (sensory inputs/evidence), but in terms of the single factor, prediction error."

In simple terms, this suggests that delusional thinking can routinely arise when 'neuronal noise' – i.e., random neuronal misfiring, a biological activity that is always present to some degree – exceeds a certain threshold such that the individual becomes inclined to jump to conclusions more quickly. This could be because evaluating available inputs in those circumstances is confounded by the

apparently greater "value of choices inferred from probabilistic reasoning" (Jardri & Denève, 2013, p.296). This mechanism may also be an explanation for disorders such as schizophrenia, as studies suggest that patients typically exhibit greater levels of delusional paranoia when neuronal “noise” is present at higher levels (Moutoussis et al., 2011, p.425).

3.7. Confirmation bias in daily life

In the previous discussion of types of confirmation bias, I have given a few real-life illustrations. However, it is easy to dismiss such illustrations as pertaining to other people, not oneself. In this section I set out a few examples of erroneous beliefs that were widely held, and some that still are. Fortunately, the long-enduring notion that the Earth is flat appears to have largely succumbed in the face of photographic evidence to the contrary, but superstition and errant rationalisation remain hard at work.

Coincidence informing popular belief

Nickerson (1998) gives the wonderful example of the mathematical relationships that Taylor (1859) and Smyth (1864, 1890) found when reviewing measurements of the Great Pyramid of Egypt:

- The ratio of the pyramid's base to the width of a casing stone was 365:1.
- Multiplying the pyramid's height by 109 produced a figure roughly equivalent to the moon's distance from the earth.
- Many other measurements of the pyramid could be related to natural distances.

Von Daniken (1969, as cited in Nickerson) pronounced these mathematical relationships as evidence that the earth was previously visited by aliens; his

books – reflecting "cosmic explanations for biblical stories" earlier attributed to Immanuel Velikovsky (Boudry, Blancke & Pigliucci, 2014, p.7) – were widely popular, and in one study were found to influence graduate students just as much as they influenced first-year students (Bainbridge, 1978, as cited in Feder, 1984, p.526).

Even now, these Velikovskian ideas about extra-terrestrial interference in human culture inform a light-hearted belief, sustained by spin-off fantasies such as the imaginary world depicted in the popular sci-fi movie and television series *Stargate,* that perhaps aliens really did design the pyramids, even though we now know that the mathematical relationships that started it all are merely coincidences (Gardner, 1957, as cited in Nickerson), or at best – in relation to some measures – attributable to the known Egyptian ability to "derive the general case from a specific numerical example" (Shutler, 2009, p.350).

Sources of harm and adversity

Another historic example of confirmation bias described by Nickerson (1998) is in relation to witchcraft, which apparently flourished in 17th century England to the extent that more than 40,000 "witches" were put to death by burning or hanging in that century alone. By the start of that century, the notion that a man or woman was presumed innocent until proven guilty was well-entrenched in English law, and yet the reverse was true when it came to witchcraft. This arose from the highly persistent belief that pervaded England and Europe at the time, that many adverse events were the work of witches and that extraordinary measures were required to protect the populace from this scourge. There is no doubt that many adverse events did occur, and that many so-called witches confessed their profession when tortured sufficiently, and so the popular belief was confirmed perpetually over several centuries despite an almost complete lack of any evidence of wrong-doing by the majority of these unfortunate souls who happened to be in the wrong place at the wrong time.

Lest we think that modern-day men and women are more discerning about the sources of adversity, consider the persistent of beliefs in the western world about the culpability of Muslim people in relation to terrorist acts. Vogel and Kebbell (2011, p.615) observe that confirmation bias operates in "forensic settings" such as in the evaluation of witnesses and the interrogation of suspects. Police and military officers investigating crimes are undoubtedly prone to confirmation bias when a suspect fits their existing beliefs, and the case of Dr Mohamed Hanreef discussed by Vogel and Kebbell is an interesting example of this. In short, Dr Hanreef – later found to be innocent of the allegations made against him – was arrested leaving Australia when erroneous information connected him to a bombing in London. Investigating officers reviewing ambiguous information about Dr Hanreef's movements and items in his possession, leapt to a conclusion that was primed by their belief that Dr Hanreef was guilty of assisting those responsible for the terror attack. The possibility of another explanation was not considered in the rush to charge Dr Hanreef and deprive him of his Australian citizenship.

Injustice in justice

Did O.J. Simpson murder Nicole Brown-Simpson and Ronald Goldman? After a trial lasting more than 8 months, Simpson was found 'not guilty' by a jury of his peers. That of itself is unremarkable, as criminal charges must be proved 'beyond reasonable doubt' and that level of certainty is often elusive when (as in Simpson's case) there was no eye-witness to the murders and the evidence was largely circumstantial. What was remarkable was the partisan reaction to the outcome. Echoing other polls around the United States, a Los Angeles telephone poll after the verdict was delivered found that "blacks were more than four times more likely than whites to think Simpson was not guilty" (Decker, 1995). Nearly 50% of whites polled were "angry" at the acquittals, compared to about 4% of blacks. Although this was clearly a racial division, that does not fully explain the extraordinary divide in circumstances where every aspect of the trial was widely publicised and communities on both sides of the racial divide furiously debated

the significance of each new piece of evidence or revelation from the witness stand. In the end, the absence of certainty allowed bias to persist unimpeded, and what we saw was bias self-sustained in each group with limited regard for the jury's determination.

In relation to court proceedings generally, Nickerson (1998) notes that the fact-finding phase is typically decoupled from the decision-making phase to avoid the formation of opinions before all of the evidence is in. However, Nickerson suggests that the capacity of jurors to follow the court's admonition to keep "an open mind" is doubtful, and some individual jurors – and even some judges – may "develop their personal mental models of 'what really happened' as a case develops and continuously refine and elaborate those models as evidence continues to be presented" (p.193, citing Holstein, 1985). Such a model can "strongly influence" the interpretation of new information, and an early view as to the accused's guilt may affect the attention given to subsequent evidence and inform a biased interpretation of ambiguous evidence (Nickerson, p.193; Hendry & Shaffer, 1989, p.545). Kalven and Zeisel (1966, as cited in Nickerson, p.194) found that the "final verdicts that juries return are usually the same as the tentative ones they initially form."

I'm dying over here, Doctor

A well-known New Zealand case (well-known to anaesthetists anyway) is that of Dr Margaret Hugel, who was charged with manslaughter after her patient, a thirteen year old boy, suffered serious brain damage when an air filter became blocked. While the case partially turned on evidence about the relevant medical protocol, a critical factor in the boy's death was the occurrence of a "patient-related problem... manifest as difficulty in breathing" well before the filter was or could have been blocked (Merry and McCall Smith, 2001, p.62). Faced with this presentation, Dr Hugel believed the problem was patient-related and persisted with ventilation efforts that ignored the possibility of equipment failure, even

though a blocked filter was the "smoking gun" that ultimately caused the tragic outcome (Merry and McCall Smith, p.21).

Mattox (2012, pp.55-6) notes that the risk of medical error arising from confirmation bias is not limited to medical professionals acting alone, and gives the example of a team engaged in treating a patient where the members individually and collectively failed to diagnose the patient's condition.

In another case, an anaesthetist was unable to insert a 'laryngeal mask airway' in order to ventilate the patient's lungs. None of the doctors (two anaesthetists and a surgeon) initiated the emergency tracheostomy procedure that could have avoided the patient's significant brain damage and subsequent death. A report on the case by Harmer (2005, as cited in White, 2012, p.45) identified that the doctors "lost track of time" as they attempted to insert the tracheal tube.

Mattox (2012, p.56) suggests that knowledge-based cognitive tasks require "extensive mental energy", while Merry and McCall (2001, p.58, citing Reason, 1990) refer to them as "effortful." This makes them especially at risk from confirmation bias (Mattox, p.56), compared to "the automatic, effortless, rapidly responsive processing by which an action is performed or a decision made virtually instantaneously" (Merry & McCall, p.58).

In their empirical study of emergency department (ED) presentations by patients who had engaged in self-harm, Milner, Kolves and Kolves et al. (2013) found that pre-held beliefs about suicide risk and propensities towards self-harm played a significant role in how these patients were treated. For example, female patients with suicide ideation and communication (SIC) received less priority for treatment than male patients who presented after non-fatal suicide attempts. Milner et al. suggest that "female instances of suicidal behaviours in EDs could be seen as 'attention seeking', 'manipulative behaviour' or a 'cry for help', and as such less worthy of treatment than presentations made by males" even though the mental pain experienced by these SIC patients may be more distressing that the physical

pain following a non-fatal suicide attempt (p.37). From this study and others cited by Milner et al., it seems highly likely that doctors and nurses in a hospital ED will often have performed beliefs about patients' deservedness to receive treatment based on patient gender and the type of presentation, and that these beliefs can persist regardless of a patient's specific presentation.

3.8. Putting a positive spin on confirmation bias

Merry and McCall Smith (2001, p.54) note that humans are highly prone to being distracted from the task at hand, and that some tasks do not "play to [our] strengths." However, we have an extraordinary ability to respond to unforeseen circumstances, and to recognise patterns based on our past experiences. Boudry, Blancke and Pigliucci (2014), despite their furious attack on the "fakery" of pseudoscience, nonetheless accept that our human intuitions and modern-day science may "be modeled as two inversely correlated and compensatory sources of cultural stabilization" (pp.2 & 17). Similarly, alluding to the multiplicity of influences to which we are subjected as organic beings, Leonyev (Leonyev & Cole, 2009, p.40) submits that:

> "maintenance of life... is itself also a directed, 'biassed' [sic], process, i.e. one that is inseparably linked with the living body's very existence and that constitutes its most essential and most necessary condition."

It is implicit in Leonyev's argument that our survival as a species is beholden to evolutionary forces that have shaped us to instinctively choose paths involving the least risk. For example, while we may deride the attribution of agency to external entities that are in reality inert or unthinking, 'error management theory' would posit that it is far safer to treat all entities as lethal agents than to erroneously fail to identify a lethal agent due to taking a more critical, scientific stance (Barrett, 2000, as cited in Boudry et al, p.9). Similarly, referring to our pattern matching capabilities, "error management predicts that it is less costly to mistake a spurious correlation for a real causal phenomenon than to mistake a

real pattern for a random fluke", even if this makes us somewhat more vulnerable to superstitious beliefs founded in coincidence (Boudry et al., p.9).

Therefore, while it is clear that in some circumstances our confirmation biases are inconvenient and may even yield outcomes that are harmful to us or others around us – a characteristic they share with most other biases and cognitive heuristics – the confirmatory phenomenon remains a critical contributor to our collective survival.

Probability is not certainty

In practice, statistics are seldom helpful to us in relation to day-to-day decision-making. Consider the question of which bank to go with when arranging a mortgage for your purchase of a new home. Multiple probabilities are involved here, some of which may be measured and evaluated using statistical tools. For example, the probability that bank A will increase its interest rates in a year compared to bank B, or the probability that bank A will be more lenient if we get behind with payments. However, what if – based on past rate rises – bank A is 50% more likely to raise its rates in a year, but bank B has a record of quickly foreclosing on delinquent mortgages. How do we weigh these disparate factors? Moreover, when dealing with larger institutions in a single free market, we may easily conclude that they are "all as bad as each other" (Roberts, 2012). None of this takes us any closer to deciding where to obtain our mortgage. However, we may have already established certain beliefs about the financial institutions around us – e.g., the people at bank C were very nice to deal with when we opened a savings account last year. Using this experiential data, we satisfy our "quest for certainty" in a way that avoids immersing ourselves in superfluous details that may be beyond our capacity or enthusiasm to understand anyway (Dewey, 1929, as cited in Miller, 1984, p.404).

The receptor capabilities of a fly are rather extraordinary. Flies' eyes are believed to have quickest visual responsiveness of any creature on earth, enabling them

to track movements as much as five times more quickly than human eyes (University of Cambridge, 2012; Wardill, List, Li et al., 2012). However, the remarkableness of the "fly eye" is the last thing we are likely to contemplate as we spray insecticide and swat the little pests when they invade our homes. In Australia there is a chain of bread store which – to avoid being sued – I'll call 'Breadx', which makes rather tasty bread products. I recently went to the Breadx store near my office, to buy a bread roll to eat on the train on the way home. To my disgust, there was a fly in the very cabinet where my chosen roll was located. As a child I was told that food is poisoned by flies, who lay their eggs on it and transfer germs to it from whatever dirty places they have visited before. This was therefore a huge dilemma for me. However, not wanting to be a slave to my cognitive heuristics (of course!) I reasoned that the fly had not been on the specific roll that I was eyeing, it might have only just entered the cabinet, and perhaps had not visited any rubbish bins that day. Later, while eating said roll, I felt rather proud of the victory of mind over emotion. Should I have been?

It could reasonably be argued that my purchase of the roll represented the failure of confirmation bias to protect me from myself, or more particularly the non-operation of confirmation bias at a time when it might have been helpful to me. *Now* I know that flies walking on food is a really bad thing – apparently the common house fly is better at transmitting viruses and other pathogens suspended in faeces "than any other substrate or medium" (Graczyk, Knight, Gilman & Cranfield, 2001, p.232) – but only because I went looking for that information while writing this chapter. If, instead, my knowledge of flies came from past experience of watching sick people covered in flies, or seeing flies living in the open toilets of the developing world, my belief about their role in adversity would be significantly stronger regardless of my understanding of the mechanisms at work.

In simple terms, probabilistic science cannot give us certainty in most day-to-day decision-making, and questions like where to obtain a mortgage or whether to buy a bread roll that might have fly poo on it ultimately rest on our own individual

past experiences and the biases formed by those past experiences. In this way, confirmation bias enables us to make decisions that in some cases we could not otherwise make, and often we will never know whether we made the right decision – it's that uncertain.

Hypotheses are not always helpful either

Previously I described the 'Wason Selection Test' (Wason, 1968, as cited in Nickerson, 1998), where subjects must decide which two out of four cards will prove the truth of a proposition that a card with a vowel on one side will always have an even number on the other side. Although this test is routinely used to demonstrate flawed reasoning, in fact the specific test formulated by Wason is a subset of a broader probability equation, the likelihood that B will be true if A is true, and sometimes – e.g., in circumstances of sparsity, such as where A and B occur rarely – confirmation bias can lead to a rational determination using a positive test strategy (Oaksford & Chater, 1994, as cited in Perfors & Navarro, 2009, p.2741).

In the real-world, solving problems that are similar to the test formulated by Wason (1968 – see above) may be assisted by two alternative approaches, one based on 'falsification', the other based on 'expected information gain' (Austerweil and Griffiths, 2011, pp.501-2). Falsification involves asking the question that will most probably falsify the hypothesis (Popper, 1935/1970, as cited in Austerweil & Griffiths, p.502). Austerweil and Griffiths suggest that a positive test strategy (PTS) is more likely to result in falsification than a negative test strategy. Klayman and Ha (1987, as cited in Austerweil & Griffiths, p.522) showed how a PTS can be optimal if the "true rule" would only explain a small number of perceived outcomes. In other words, confirmation bias is beneficial in many circumstances because it favours a PTS that is more likely to result in falsification, and is consequently more likely to help us predict the next event, particularly where the relevant rule arises rarely.

Where a hypothesis cannot be determined because the rules of a test or situation are unknown, it is rational to maximise expected information gain by seeking information that adds to existing knowledge rather than necessarily reaching for the answer to the test or situation; where the world is perceived as deterministic, which equates to non-deterministic outcomes being rare, a PTS is rational (Austerweil & Griffiths, p.507). While this still leaves us with the fact that most people fail the Wason test – failure that is widely perceived in scientific terms as a shortcoming of human cognition – we can nonetheless surmise that confirmation bias and specifically a PTS as one manifestation of confirmation bias might instead be nature's way (refined through evolution) of equipping us to reach *good enough* decisions most of the time. As Navarro and Perfors say, "we might forgive human participants for following an optimal strategy that happens to fail in some cases" (2011, pp.130-1).

Believing in God: a 'lazy way out', or a legitimate personal conviction?

Richard Dawkins, who has written several popular books on scientific topics, is also a well-known atheist. Dawkins (1986, pp.140-141) discusses the remarkable complexity of DNA replication, and – in his view – the sheer impossibility of there being a creator who controls all aspects of the "machinery of replication" and evolution on planet earth. Twenty-five years later, Dawkins (2012, p.261) takes a similar stance in discussing the parable of Jesus turning water into wine:

> "[This feat] would violate some of the deepest scientific principles we know... Molecules of pure water would have to have been transformed into a complex mixture of molecules, including alcohol, tannins, sugars of various kinds and lots of others."

Despite Dawkins' commitment to the logic of his scientific analysis, the persistence of religious belief in every part of the modern world continues undeterred. How do we reconcile this? Perhaps we have to conclude that science is not particularly useful when it comes to matters of faith. For a psychologist this

is a significant issue. On the one hand, we are scientists too, and the idea that some erroneous beliefs may be sustained by coincidences or selective acquisition of information should be abhorrent to us. However, we are also concerned with the mental health of individuals and communities, and we know that in hard times people cling to beliefs as a way of coping.

Ultimately, we have to make a relative assessment of harm. While a belief in God or some other deity may offend Richard Dawkins, the reality is that such a belief is of itself seldom harmful. In the absence of knowledge, we will typically have uncertainty, and there is compelling evidence that in some circumstances uncertainty has adverse consequences for our mental health regardless of the dubiousness of the fictions we might adopt in place of it. Correspondingly, a spirit-filled life sustained by confirmation bias may be a recipe for psychological well-being, even if the association between the bias and mental health is not fully understood and implicitly endorses beliefs that cannot be proved by science (Ellison & Fan, 2008, p.269; Jackson & Bergeman, 2011, p.13; Kashdan & Nezlek, 2012, p.1532).

Believing in ourselves and in what we can do

Earlier on I mentioned the long-running television series *Stargate,* featuring the redoubtable "Macgyver" (Richard Dean Anderson) playing another role where human ingenuity wins the day. The popularity of the series is no doubt connected to the survival every week of the plucky team of humans that battles too-clever-by-half aliens in far-flung corners of the galaxy. The arch-nemesis – at least for a few years – was the *Goa'uld* Anubis – the name plucked straight out of Egyptian mythology – who is smarter than any human being and has some very cool devices to zap people with. However, for a God, Anubis has incredibly bad luck. Every week that little band of humans, who he now knows individually by name, get the better of him. All that alien brain power and technology just doesn't cut the mustard.

While we can easily pigeon-hole these sci-fi series as "harmless entertainment", they are nonetheless value-laden and influential in shaping our personal mythology and beliefs about the world (Langley, 2012, pp.93-94; Buckland, 2000, pp.5-6). Black and Barnes (2015), building on earlier studies that suggested a correlation between reading fiction and 'theory of mind' (Kidd & Castano, 2013, as cited in Black & Barnes, p.423), found the same effect in subjects who watched fictional television dramas.

It follows that in everyday life, confirmation bias can strongly influence the extent to which we strive to achieve laudable goals. For example, the 'Pygmalion effect' has been postulated as the mechanism whereby positive reinforcement leads to improved academic performance (Panda, 2014, p.2). Put simply, the simple belief that one is capable of higher achievement leads to better outcomes. Although conceptually the Pygmalion effect is not necessarily a product of belief being reinforced by selection or interpretation of information available subsequently, in reality it is hard to distinguish the two. The effect came to the attention of the scientific community through the work of Rosenthal (1963), embracing the notion of 'self-fulfilling prophecy' but also relying on social cues reinforcing the expectation of higher achievement (Panda, pp.3 & 5).

Fighting wars and not being killed

One place where conscious attention to the truth of the situation can be fatal is at the 'front' of a war zone. Soldiers often face weeks of tedium, punctuated by short bursts of gun fire. In that situation, war becomes characterised by strange rituals and beliefs, including a fervent belief that the enemy – i.e., every man and women of them – is pure evil. Perhaps it is a necessary belief in wartime. Wright (2004, p.4) gives us the example of U.S. marines in Iraq, who he observed as a reporter, and the following illuminating paragraph:

> "At least one Marine in Colbert’s Humvee seems ecstatic about being in a life-or-death gunfight. Nineteen-year-old Corporal Harold James Trombley,

> who sits next to me in the left rear passenger seat, has been waiting all day for permission to fire his machine gun. But no chance. The villagers Colbert's team had encountered had all been friendly until we hit this town. Now Trombley is curled over his weapon, firing away. Every time he gets a possible kill, he yells, 'I got one, Sergeant!' Sometimes he adds details: 'Hajji in the alley. Zipped him low. I seen [sic] his knee explode!'"

Confirmation bias persists in wartime just as readily as it does in peacetime. Young Corporal Trombley must believe the enemy is not like him, that when you kill "one" it's a just like shooting a wild pig in a hunt. It's not hard, as bullets come flying back, to assess those nameless others as evil. After all, they act like we expect them too, and there's no basis to see them as fellow humans.

And then there's the belief about one's fellow soldiers (p.19):

> "Despite the frictions [between officers and enlisted men], Fick believes in the men he commands. "I have the best platoon," he says repeatedly. Away from his men, Fick cannot talk about them without smiling."

An unbiased appraisal would inevitably find a multitude of faults and weaknesses, but in these situations, belief becomes like a glue that ensures the marines watch each other's backs and demonstrate total loyalty.

3.9. Managing the influence of confirmation bias

Confirmation bias cannot exist in splendid isolation. Necessarily and demonstrably, confirmation bias is a moderator of factors that can influence human behaviour – i.e., if an individual has formed a belief, confirmation bias operates to moderate the influence of subsequent information that this individual is exposed to. However, confirmation bias itself is also susceptible to moderation by other factors including other unconscious biases and conscious thinking

activities. This section explores some of those other factors, and how they may operate in practice.

Information utility

The utility or practical value of specific information has been proposed as having the potential to overcome confirmation bias. Knobloch-Westerwick (2008, p.2273, as quoted in Knobloch-Westerwick & Kleinman, 2012, p.172) suggests that 'information utility' may be a representation of the extent to which we pay attention to information that could potentially address "challenges or gratifications" that are significant, highly likely to arise, need to be addressed soon or immediately, and may be influenced by the said information. Knobloch-Westerwick & Kleinman (2010, citing Purcell, Rainie, Mitchell, Rosenstiel and Olmstead) support the proposition that information utility may be more effective at overcoming confirmation bias affecting regular users of the internet, but Purcell et al. do not mention either information utility or confirmation bias specifically.

The findings advanced by Purcell et al. (2010) include that the internet is now a more popular source of news than newspapers and radio, that typical internet users find the volume of news online "overwhelming" and are loyal to a small number of online news sources, and that internet users are interested in news because they believe that knowledge about the news can enrich their social interactions (pp.3-6). While those surveyed by Purcell et al. were undoubtedly attributing utility to information, it is a long stretch to say that utility in these circumstances was demonstrated to overcome a bias towards preferred news sources. Knobloch-Westerwick and Kleinman themselves acknowledge that the internet is read in a different way to more conventional news sources, and we might infer from their analysis that internet users' greater "effective[ness] in avoiding attitude-discrepant messages" is offset by more extensive reading behaviour (pp.172-3).

Assuming a two-layer approach to information control/attention, in some circumstances we may avoid being exposed to attitude-discrepant messages – i.e., the first layer operating to exclude information that we are not inclined to hear – while in other circumstances reification may operate as an effective second layer protection (against information that gets through the first layer) by categorising unwanted information in a way that makes it easier for us to ignore that information.

Perceptual disfluency

It has long been known that cognitively demanding information is likely to be encoded into long-term memory more robustly than uncontroversial information (Craik & Lockhart, 1972; Craik & Tulving, 1975, pp.289-90). Oppenheimer (2008, p.237) identifies that the subjective challenge of processing such information into memory – i.e., the metacognitive cue known as 'disfluency' – is distinguishable from the objective level of difficulty associated with the information, a point repeated by Diemand-Yauman, Oppenheimer and Vaughan (2011, p.111). Oppenheimer describes that 'fluency' in terms of the "feeling of ease associated with a cognitive operation", as distinct from the actual cognitive task that is required (p.237).

Although the studies conducted by Oppenheimer, and later Diemand-Yauman et al., were primarily concerned with memory, Oppenheimer makes the point that perceptual fluency operates more broadly such as when we are evaluating truth, likability and compatibility – i.e., fluency encourages an affirmative conclusion relying on heuristic cues, while disfluency encourages more thorough evaluation. Reber and Schwarz (1999, p.342) suggest that the effect of perceptual fluency is akin to the effect of familiarity. The premise here is that fluency elicits a comfort level that posits favourability. This takes us to Hernandez and Preston's conclusion that disfluency, while wholly metacognitive (and therefore subject to moderation itself by extracerebral forces) can interfere with, and potentially 'disconfirm', a confirmation bias (p.181).

Multiple biases

Biases frequently – perhaps always – operate in tandem with other biases, even biases of the same type. For example, I might be a juror in the criminal trial of a man charged with murder. The trial has been running for a week, and already I have formed two beliefs from the evidence that I have heard:

(a) The defendant is capable of committing a murder; and
(b) The defendant didn't commit the murder in question.

This duality of belief may mean that I am concurrently open to information that is exculpatory ("he didn't do it") and damning ("he's capable of doing it"). The first belief effectively moderates the second belief or vice versa. This is like the situation where a second similar charge is being considered by a jury at the same time as the primary charge, as opposed to the charges being heard separately, and the possibility that the jury considering the two charges together is more likely to convict on the primary change (Leipold & Abbasi, 2006, as cited in Goodman-Delahunty, Cossins & Martschuk, 2016, pp.46-7).

It is virtually impossible to identify all the biases and cognitive heuristics that might be at play in a given situation, so we need to be cautious that in addressing one recognised bias we do not elevate the influence of another bias that has not been identified. For example, a person may have been bitten by a dog and now fears dogs. When that person sees a dog barking at some young children, he or she may misinterpret the situation as representing danger to the children even though the dog is well-secured and cannot get near them. In this case, there are almost certainly multiple biases affecting the person's behaviour. Bias-busting measures that address just one bias or belief may be ineffective, or may even worsen the situation, if they fail to consider other possible influences. Care must be taken in evaluating scenarios where confirmation bias appears to have played a role.

Chapter 4 – Discourse Analysis Theories and Methodologies

This study is grounded in the 'interpretative paradigm' that interprets individual reality as a construct of human interactions, experiences, understandings and beliefs (Scotland, 2012, pp.11-13). This approach aids what is effectively an exploratory study, as there is no additional evidence or information available that can provide experiential context for the witnesses' statements to the Royal Commission. The interpretive paradigm facilitates the exposure of meaningful patterns and insights. In the present study, the interpretive paradigm is applied to witnesses' narratives that have been placed into the public domain. The logic of the study is inductive, following the theoretical background of the confirmation bias phenomenon described in chapters 2 and 3, using discourse analysis to consider and elucidate the operationalisation of the phenomenon.

The meaning of 'discourse analysis' varies significantly across the different disciplines where the term is used. A good starting point may be to define 'discourse' sans context, which for present purposes must be more than mere words strung together purposelessly. In this thesis the more informal term 'narrative' is intended to have the same meantime as discourse. Some other definitions of discourse are:

- "A continuous stretch of (especially spoken) language larger than a sentence, often constituting a coherent unit, such as a sermon argument joke or narrative" (Crystal, 1992, p.25, as cited in Täuschel, 2004, p.7).
- "stretches of language perceived to be meaningful, unified and purposive" (Cook, 1989, p.156, as cited in Täuschel, 2004, p.7).
- "a system of statements which constructs an object" (Parker, 1989, p.61, as cited in Wooffitt, 2005, p.146).

Although some commentators suggest that a discourse must comprise more than one sentence (e.g., Nunan, 1993, p.6, as cited in Täuschel, 2004, p.10), this does

not appear to be universally accepted, and is not wholly logical. In some contexts, a single sentence can be full of meaning and intention. The issue is not the number of words or how they are arranged structurally, but rather how words/statements through their juxtaposition articulate an intention or purpose.

Cook (1989, p.6, as cited in Täuschel, p.12) proposes that the focus of discourse analysis is the narrative of a person, and the "search for what gives discourse coherence." Coherence exists where the discourse is consistent with "normal" and "familiar" experiences that can be understood by the common-sense application of local knowledge (Täuschel, 2004, p.16, citing Yule, 1996, pp.84-85). Täuschel notes that in "its full richness discourse analysis involves all the levels and methods of analysis of language, cognition, interaction, society and culture" (p.11). This unfettered scope is consistent with discourse being "analysed as a reflection of wider structural and social inequalities" (Wooffitt, 2005, p.144). Wooffitt (p.94) also identifies the need for a critical approach to discourse analysis, even though the discipline is perhaps inherently critical. Wooffitt's point appears to be that to be taken seriously discourse analysis must demonstrate the intellectual rigour conventionally associated with "related critical approaches in social psychology" such as those advocated by Michel Foucault (Wooffitt, pp.94, 137 & 146).

We are dealing here with pre-prepared statements, that in some cases are the subject of little or no cross-examination. In these cases, it would be right to observe that the person doing the communicating has the opportunity to craft narrative in a way that might diverge considerably from the more natural language arising in ex tempore narrative. However, that situation does not invalidate the narrative, it merely requires us to recognise the character of the narrative arising from its manner of creation: (Brown & Yule, 1983, p.5)

> "There are, of course, advantages for the [ex tempore] speaker. He can observe his interlocutor and, if he wishes to, modify what he is saying to make it more accessible or acceptable to his hearer. The writer has no

access to immediate feedback and simply has to imagine the reader's reaction."

We might say that in a complex matter the preparation of a statement in advance enables the writer to cover all the ground that is desired to be covered, without the stressors of thinking on the spot. Even in a formal context such as courtroom, spoken narratives (unless rehearsed) are inevitably drawn from 'top of mind' without opportunity for reflection and revision. Writing on the other hand is "worked over" to produce an output carefully tailored for its intended purpose (Chafe, 1994, p.43). However, if all we have is writing then we must accept the limitation of that and move on with it. In relation to identifying the significance of relevant contexts, it is highly relevant to analysis of witness statements as discourse that the witness has a "legal status" in relation to the court or tribunal (in the present case, in relation to the Royal Commission) and the extant legal process (Wooffitt, 2005, p.22).

Further important areas of divergence between informal writing and legal statements include the level of attention to spelling and grammar. In most cases, the spelling of words in a letter or personal journal will reflect the language abilities of the individual, whereas spelling in a legal statement will at least reflect the capabilities of a computer-based spell-checker. Grammar likewise. In addition, the careful use of structure to produce "ordered context" can improve the accessibility of the narrative, and rewording "naïve representation[s]" may also promote the desired objective (Di Donato, 2011, pp.123-4). For example, a lawyer might insert "clinical" in front of "psychologist" to slightly elevate the apparent status of a relevant expert, replace formal references to a person ("Mr..." or "Ms...") with the person's first name to indicate greater closeness, or portray a dodgy activity as if it were commonplace. As is the case with improving the grammar of a narrative beyond the native ability of the original author, this kind of "legitimation" may produce a slightly skewed impression in the mind of the reader (Leeuwen, 2008, pp.20-21).

A feature of cross-examination is how the lawyer has an "argumentative advantage" -- i.e., it is inherent in efficient cross-examination that the questioner controls the witness and elicits only anticipated responses, rather like the radio host who can "attack a caller's position without having to advance an alternative argument" (Wooffitt, 2005, p.198, citing Hutchby, 1996). However, the wily or determined witness is not wholly powerless, as cross-examination is often conducted imperfectly leaving opportunities for a witness to veer off into narrative that undermines the lawyer's advantage. In my experience, police officers can be particularly adept at eschewing the yes/no responses that cross-examination typically tries to elicit (e.g., "yes, but...").

Implications are often drawn from narratives. For example: "He is an Englishman, he is, therefore, brave" (Brown & Yule, p.31). The implication here is that all Englishmen are brave, which may be a fallacy but is nonetheless an expression of belief that may be significant in the broader understanding of the specific narrative. It is probably nonsensical to regard all Englishmen as brave, and so a pragmatic analysis may (for example, depending on other context) treat the discourse as prioritising "rhetorical appeal to the reader" and inviting the reader to share the presupposition that Englishmen are at least typically brave (Potter & Edwards, 1990, p.412, as cited in Wooffitt, 2005, p.81). The significance of this is that in undertaking discourse analysis we are often faced with narratives that seem rather certain in what we might infer from them. However, there is considerable danger in seizing on such inferences without a strong basis for believing them to be true.

In relation to most narratives, some knowledge of context is essential. Täuschel (2004, p.14) identifies two difference main types of context:

- Linguistic context, being "the language that surrounds or accompanies the piece of discourse under analysis", and

- Non-linguistic context aka experimental context, which is the external context where the discourse occurs.

Non-linguistic context includes historical and circumstantial context, which might also be called "background knowledge" and particularly refers to "pre-existing knowledge structures" such as 'schemata' or 'scripts' that operate as a guide to understanding current situations (Täuschel, 2004, p.17). In some cases, context may be included selectively to lend strength to the credibility of the narrative (Di Donato, 2011, p.123). In other cases, discourse may have been constructed to conform with the expectations of other people (Pillemer, 1992, p.242, as cited in Wooffitt, 2005, p.109). While this thesis is not the place to debate the reliability of human memory, it is critical in undertaking discourse analysis to recognise the limits of perception and memory. There is considerable evidence that our ability to comprehensively assess and accurately record our experiences is rather limited (Chafe, 1994, p.22). Moreover, the 'reconstructive' nature of how we interpret fresh information based on existing schemas means that our ability to distinguish fact and fiction is diminished from the outset, and further diminished by the limitations of how we store memories -- the recall of "once-immediate experience inevitably entail[ing] a greater or lesser amount of invention" (Chafe, 1994, p.33, citing for example Loftus, 1979). Discourse may therefore be unreliable as a "window" looking into our minds (Wooffitt, 2005, p,115). However, the utility of discourse analysis is unaffected by this, because ultimately the discourse once created takes on a life of its own. Therefore, the study of discourse is not an examination of the person or their memories but looks at what he or she does with language and how that can be related to the contexts and identified purposes of the discourse (Wooffitt, pp.115-6). Being able to describe certain matters in great detail sometimes has the effect of lending credibility to other matters touched on by the witness, even though that treatment is probably unwarranted (Bell & Loftus, 1989, as cited in Wooffitt, p.119). If we are looking for credibility, it needs to come from the context not the discourse.

A challenge confronting any study involving discourse is that the context and

objectives of a specific narrative are a 'paradigm' that "necessarily focuses attention on certain phenomenon while obscuring others" (Burns & Carson, in Wodak & Chilton (Eds), 2005, p.288). Moreover, the availability of institutional strategies and practices to address specific problems may inherently influence the formulation of problems -- i.e., so they can be located within existing strategies and practices -- and narrow the scope of possible solutions (Nylander, 2000, and Sutton, 1998, both as cited in Burns & Carson, 2005, p.291). In some situations, the problem-solving capability of existing strategies and practices may be so poor or ineffective that a crisis ensues (Burns & Carson, p.292).

It is necessary to be alert for influences that may exacerbate bias in testimony. One obvious influence is the manner in which evidence is adduced -- i.e., by a written statement being read into the record, and in some cases cross-examination of the witness -- because that scheme tends to favour some forms of expression and constrain others. Less obvious influences on testimony include the capacity of the witness to confront and articulate traumatic events, the sensitivity of non-witness actors to the wellbeing of witnesses, how non-witness actors perceive evidence in terms of relevance or probity, and decisions about cross-examination. For example, a lawyer's decision whether or not to cross-examine a witness, and the extent of any cross-examination, inherently defines the scope of what the witness is allowed to say -- i.e., the discourse of the witness is expanded or confined by the decision of another actor. Where there is no cross-examination, we are left with several possibilities:

- the evidence of the witness was considered to be complete and accurate;

- the evidence of the witness was considered to be unreliable to the extent that cross-examination would be pointless; and/or

- the witness was considered to be vulnerable, and that cross-examination may be harmful to the witness.

It is important to keep in mind that these witnesses were not on trial. The focus of the Royal Commission was the institutional (typically corporate or government agency) responses to child sexual abuse, and in the main the objective of hearing from individual witnesses (victims, parents of victims) was only to gain an understanding of what the institutions should have responded to.

It is commonplace to find human bias represented in narratives, spoken and written, and discourse analysis is often a useful lens through which to view bias and evaluate it in its broader context. Van Dijk (1993, p.262, as cited in Wooffitt, 2005, p.138) suggests that biases reflecting an individual's beliefs may be evident from an examination of the individual's discourses.

In this study, where confirmation bias is a central focus, it is appropriate to go a step beyond exposing individual beliefs, to consider the nuances of bias and specifically how confirmation bias may be detected and its operation described in a specific situation. This research uses transcripts of data gathered from a public source. Having these transcripts, it is possible to respond to a critical question: what are the main tendencies in the narratives according to our understanding of confirmation bias as described in chapter 3. In the analysis of these hearing transcripts I have considered (for each narrative) the context of any belief, the language used, the use of metaphors and idioms, exaggeration, and other rhetorical factors.

Chapter 5 – Discourse Analysis of Witness Statements

The Royal Commission heard from victims of child sexual abuse, parents of victims, co-workers of perpetrators, and managers of relevant institutions. The evidence was typically documented in advance by way of witness statements. When it came time for the evidence of a witness to be heard by the Royal Commission at hearing, the statement would be read into the record by the witness or another person such as the counsel (lawyer) assisting the Royal Commission. A witness present in person could then be cross-examined regarding his or her evidence.

This chapter is broken into two parts, the first being the analysis of parental discourse – i.e., the narratives of parents, mainly mothers, who had a child abused by Jonathan Lord, and second the analysis of co-worker discourse. In most cases, the abuse was sexual. In some cases, it persisted for many months and only came to light after Mr Lord was arrested. Each case was an inestimable tragedy for the affected child and his family.

It is important to note that the witness narratives analysed here are only the tip of iceberg – a small number, less than a dozen, representing the interests of the very large pool of Mr Lord's victims. Keep in mind always that these witnesses were not on trial, and there is nothing here to suggest that any of them were neglectful of children's safety, welfare or wellbeing. On the contrary, these narratives show parents and workers who wanted the best for the children in their care, and who believed that the YMCA Before and After School Care service afforded a safe and enriching experience for these children. Although in one sense the YMCA was on trial before the Royal Commission, the focus of the inquiry was towards the future and how to improve institutional responses to child sexual abuse.

As signalled in the previous chapter, the following analysis of discourses is particularly concerned with how individual beliefs may have informed biases,

particularly confirmatory biases. To that end, I have considered here the contexts of beliefs, the language used by witnesses in describing their experiences, and what the witnesses' use of metaphors and idioms, exaggeration, and other rhetorical factors reveals in relation to their biases.

5.1 Mothers and their sons

Seven parents gave evidence in the Royal Commission's case study relating to Mr Lord. In each case except one, the witness was the mother of a boy enrolled in the YMCA before and after school care program. Girls were also enrolled in the program, but it appears that Mr Lord's abuse was directed towards boys only. The testimony of one mother ("AW") concerns her fears about her daughter having been abused, although no abuse of that child was able to be confirmed.

The parent witnesses and their children are referred to by the unique pseudocodes (identifiers) allocated to them by the Royal Commission, for example "AN", "AO", "AX" etc. The original transcripts produced by the Royal Commission were not line-numbered, so line numbers have been added to my transcript data to aid the analysis. Accordingly, "AN.24" denotes line 24 of the transcript data comprising AN's evidence. See appendix 2 for the full list of witnesses and identifiers.

Other people trusted him, so I did

AN, the mother of victim AO, knew Jonathan Lord's mother Jill Yankos through her work. Mr Lord's mother had recommended Mr Lord as a babysitter for AO, and AN had engaged him in that capacity (lines AN.24-30):

> AN.24 Jill recommended Jon to me as a babysitter. She told me that Jonathan worked at a local YMCA and was
> AN.25 a supervisor there, was great with kids and needed the money. Jill was aware that I needed a babysitter,

AN.26 as I am a single mother and was working every Saturday in a real estate office. Jill Yankos was very

AN.27 proactive in recommending her son to me as a babysitter, telling me how all the children loved Jonathan

AN.28 and that he was great with all children.

AN.29 I trusted and believed that if Jonathan Lord was a supervisor at the YMCA, he was more than qualified to

AN.30 babysit [AO]…

It is not suggested that Ms Yankos was in any way aware of Mr Lord's paedophilia or his abuse of young children. However, from Ms Yankos description of him Mr Lord appeared to AN to tick all the boxes as a person who would be suitable as a babysitter, and AN was therefore inclined to see Mr Lord as an answer to her difficulty securing care for AO on Saturdays when she had to work. AN describes Ms Yankos as "very proactive" in making her recommendation, inferring that Ms Yankos was pressing AN to use Mr Lord for babysitting. AN then goes on to emphasise that regardless of Ms Yankos' recommendation, AN had good reason to trust Mr Lord because he "was a supervisor at the YMCA." In other words, AN infers that as YMCA trusted Mr Lord as a supervisor it was okay for AN to trust Mr Lord as a babysitter. Nonetheless, AN's stated recall of Ms Yankos comments about how "all the children loved Jonathan" and "he was great with all the children" (and earlier "great with kids") suggest that these aspects of Ms Yankos recommendation resonated with AN and were influential in her decision making.

Another mother, AX, also worked with Ms Yankos and similarly recounts a suggestion by Ms Yankos that Mr Lord could babysit her child:

AX.16 Jill Yankos spoke about her son with me. She said that he loved kids and that he had always worked with

AX.17 children. On numerous occasions she suggested that Jon could babysit my children. I never took Jill up

AX.18 on that offer.

This testimony is given in the context of the witness's son and [AN's] son both having subsequently been sexually abused by Mr Lord during YMCA activities. It is significant that the witness emphasises Ms Yankos making suggestions about Mr Lord babysitting the witness's children – i.e., on "numerous occasions..." It is also significant that the witness is similarly emphatic about her response – "I never" – viz: "I never took Jill up on that offer." In this way the witness is moving some responsibility to Ms Yankos, as the persistent offeror, herself as the determined refuser.

AX.21 ... I knew that Jonathan Lord regularly babysat [AO] for [AN]. If [AP] went to
AX.22 play at [AO]'s house, I knew that Jon may have been there, the person looking after the children, not
AX.23 [AN].
AX.24 It was established between [AN] and I that it really didn't matter if she or Jon were looking after the
AX.25 boys, because we thought that they were in good care with Jon Lord

Here again the witness shifts responsibility. It wasn't her decision alone to allow Mr Lord to look after her son, "it was established between [AN] and I... we thought that they were in good care with Jon Lord." However, this arrangement represents a significant deviance from the earlier "I never took Jill up on that offer" and infers that AX was in her mind placing some reliance upon the decision making of AN (mother of AO) when the boys were together at AN's home.

In the case of witness AZ, it was her own son – AE, who later disclosed being sexually abused by Mr Lord – who had suggested to his mother that Mr Lord might babysit AE and his two siblings. AE and his siblings were attending another school that had its own vacation care program. During her testimony, AZ recounted the history of how her child came to be in the YMCA vacation care program, as their own school program had insufficient places:

AZ.17 In April 2011, I enrolled two of my children into the YMCA vacation care program in Caringbah, run out of

AZ.18 Laguna Street Primary School. They also attended YMCA in July 2011 and September 2011 school

AZ.19 holidays. I first enrolled them due to lack of available places at the vacation care program in our school.

AZ.20 I recall my son in particular enjoying it, and as he was getting older, he was also able to go on excursions

AZ.21 for children that were over eight.

In the context of a narrative about the child's abuse, the reference to how the child came to be enrolled in the YMCA vacation care program is significant. The manner of referring to the decision suggests that the school program would have been preferred, but the YMCA was the only alternative. AZ does not give any detail about why enrolling the children in a vacation care program was necessary, or steps taken to check the YMCA program out, or investigation of other alternative programs that might have been available at the time. We might reasonably infer here that no specific enquiries were carried out, and the reference to the lack of available places in the school program is put forward as a way of suggesting that parental choice at the time was limited (by the school having insufficient places) and that if the school program had been available the abuse would not have occurred. The decision to enrol AE in the YMCA program is then further supported by the reference to AE enjoying the YMCA program, and that it involved activities that met his needs as a child "getting older." The witness specifically refers to AE being able to "go on excursions for children that were over eight" (my emphasis). At the time, AE was eight years old. Taken together, these two sentences put forward the proposition that AE's enrolment in the YMCA program was to some extent taken out of the witness's hands by the lack of places in the school program, but that fortuitously the YMCA program appeared to meet the child's needs at that time so the decision to enrol him in the YMCA program was justified. The use of the words "my son in particular enjoying

it" is notable because this indicates that AE was very happy attending the YMCA program, whereas his sister (the other sibling enrolled) might not have been as satisfied. It is known now that Mr Lord paid special attention to little boys. The witness does not comment on why her son enjoyed the YMCA program more than her daughter did.

AZ.22 He told me how much he liked the carers, and one afternoon he said to me, "I would like Jon to babysit

AZ.23 for us." At the time, I didn't think this was particularly out of the ordinary, even though all of the other

AZ.24 childcare centres and preschools that my children had previously attended had policies which prevented

AZ.25 babysitting of children by staff.

In the case of the child AJ, his mother AT describes how she came to place AJ at the after school care operated by YMCA at Carringbah:

AT.23 A. We don't have any family living in Sydney, and I didn't know anyone well enough to trust to look

AT.24 after my children at home, and I thought that the service would be safe and YMCA is a large organisation

AT.25 and I thought it had a good reputation. The school advertised the service at our kindergarten

AT.26 information night and held the service out to be a good service. There were signs on the school gate.

In this situation, the YMCA service responded to AT's vulnerable position as the primary carer of her son who also needed to work to meet the financial needs of her family.

As with the other mothers, AT placed great store in the standing of YMCA as a "large organisation" with a "good reputation." In particular, the school "held the

[YMCA] service out to be a good service" and promoted it. AT expresses all of this in the context of not knowing who to trust. The inference is that in the absence of other options it was reasonable to trust YMCA in view of that organisation's size and repute and also the effective endorsement of the YMCA service by the school.

Although having trust in the integrity of a person or organisation may be regarded in a particular case as a prerequisite for giving that person or organisation a specific responsibility, trust does not of itself drive people to give out responsibility. In the ideal world, all parents making decisions about childcare would be, and feel, completely in control of their decision making, but the examples in this section allude to how little choice some parents may really have. In these examples, there is a theme of parents having to balance multiple demands on their lives, including work obligations and parenting responsibilities, and having limited alternatives to using a specific available service or person for childcare.

He was "willing and available"

From about December 2010, AN used Mr Lord as a babysitter for her son AO. As noted previously, the engagement of Mr Lord for this activity was recommended by his mother and AN trusted Mr Lord based on his YMCA role. Once Mr Lord was engaged in the babysitting role, he proceeded to secure AN's confidence:

> AN.42 In hindsight, I now see that Jonathan Lord groomed me as well as [AO]. Jonathan was always willing and
> AN.43 available to help and became like a member of the family. He would occasionally drop by the house and
> AN.44 come to dinner, where he would talk with me and watch television with [AO]…

AN connects "willing and available" with becoming "like a member of the family. This theme of being helpful and generous with his time is emphasised, such as where AN describes Mr Lord "talking with" her and watching television with her son, activities that build familiarity and trust.

AN.44 … He would sometimes lend
AN.45 [AO] things, such as video games. He would sometimes bring things for [AO] for him to do while he was
AN.46 babysitting, such as making models.

"Lending" in this case denotes another form of perceived generosity, involving the provision of items by Mr Lord that AN infers were desirable to her son, and were also desirable to AN in the sense that she appreciated seeing her son playing happily. Similarly, with the model-making, Mr Lord is portrayed as demonstrating generosity, and apparent thoughtfulness.

Other mothers record similar involvement with Mr Lord. For example, AX's son went to a movie night that Mr Lord arranged:

AX.32 ... Jon Lord told me he was organising the event and I thought it was fine that the children were
AX.33 going to a YMCA movie night with the organiser, even though neither [AP] nor [AO] attended YMCA
AX.34 childcare.

The terminology here is curious. Mr Lord "was organising the event" and it was "fine that the children were going to a YMCA movie night with the organiser." The witness avoids a narrative where she has entrusted her son to Mr Lord; instead it is depersonalised as a formal, planned event supervised by "the organiser." AX continues:

AX.37 ... At 10pm, I texted Jon to see if everything was all right, as it was a bit

AX.38 late. Jon replied, saying that they were nearly finished and that [AP] would be home shortly. Jon

AX.39 dropped [AP] home later that evening.

Here AX portrays that she was a responsible parent, ringing up at 10pm to check that her son was okay. AX then records how she was reassured by Mr Lord – i.e., "nearly finished... home shortly" – and the fact of her son subsequently being dropped home by Mr Lord. It was therefore a promise kept by Mr Lord, and obviated any need for AX to be concerned. There is a theme here of a parent representing herself as having done what a responsible parent might be expected to do, in the context (at the time of providing her evidence to the Royal Commission) of her later-acquired knowledge regarding Mr Lord's crimes.

Another mother, AU who had three children enrolled at the YMCA before and after school care service at Caringbah Public School in 2011, describes Mr Lord's behaviour towards the children and how the children felt about Mr Lord:

AU.23 I understood from speaking with my children that they had grown very fond of Jonathan Lord because

AU.24 he made a big effort to entertain them and pay them attention.

AU.25 I recall that he would often cuddle them or be playful with them by carrying them on his shoulders.

AU.26 When we arrived in the mornings my children would often be the first ones at the centre, and Jon would

AU.27 allow them into the kitchen to get whatever they wanted for breakfast. There were occasions when I

AU.28 dropped my kids off at before school care and Jonathan Lord had opened the centre alone. As a result,

AU.29 my children were left alone with Jonathan Lord until the next staff member arrived.

A significant theme of AU's discourse is Mr Lord's availability and willingness to engage, expressed with rhetorical emphasis – e.g, "big effort", "often cuddle them", and "whatever they wanted for breakfast." There is an obvious question here (that we cannot answer) as to AU's knowledge of the YMCA rules against physical contact between workers and children, and whether she would have realised that Mr Lord's conduct was contrary to those rules. However, AU's specific recollection of this conduct in the context of her children's described fondness for Mr Lord suggests that AU saw the conduct in a positive light, perhaps because it indicated special treatment for her own children over others.

The rhetorical emphasis given to aspects of Mr Lord's conduct continues when AU is discussing Mr Lord's attentiveness to one of her children and herself on other occasions (my emphasis here):

> AU.53 ... I knew that
> AU.54 Jonathan Lord had paid <u>a lot of</u> attention to [AM]. There was one weekend in or about June 2011 where
> AU.55 Jonathan Lord had made an arrangement with [AM] to come and watch him play football on a Saturday.
> AU.56 At the time I did not think it was suspicious. I just saw him as a young guy who liked kids. Jonathan
> AU.57 spent <u>a lot of</u> time trying to get to know me and would inquire as to how my university study was going.

In this way, Mr Lord is portrayed as generous and selfless, and therefore by implication someone who could be trusted by AU. The words "I just saw him as a young guy who liked kids" denote a kind of normalcy – i.e., Mr Lord did not stand out as unusual or deviant in AU's perception of him. Similarly, Mr Lord befriending AU and asking AU about her university study is implied as something a caring person would do, and therefore not consistent with a person having nefarious motives. This is put forward as supporting AU's lack of suspicion of Mr Lord.

The kids didn't say anything

Mr Lord babysat AO. AO's mother AN describes how the children were keen to spend time with Mr Lord:

> 47 [AO] and other kids appeared to love Jonathan. He came to [AO]'s laser tag birthday party voluntarily
> 48 and unpaid. [AO] and his friends fought about who was going to be on Jonathan's team and who would
> 49 get to ride in Jonathan's car. Jonathan had a way with children.

Overt generosity again, this time "voluntarily" coming to a child's birthday party, giving up his time and not expecting any reward for it. From an adult perspective, the children's conduct – AO and his friends fighting about who would be on Mr Lord's team and who would go in his car – is portrayed as compatible with a belief that nothing untoward occurring.

Keeping in mind that AN's narrative is generated after the abuse has been disclosed, it appears likely that AN is embarrassed about being duped by Mr Lord and the harm he has done to her son. Although the words "had a way with children" are put forward as a summation of the behaviours described, they may also be viewed as excusatory – i.e., suggesting that Mr Lord had a unique or rare talent in dealing with children, and that AN should be excused for not seeing past that talent. In this way, Mr Lord is portrayed as the conniving manipulator who duped innocent parents and children.

Later on, AN recounts AO's response when she asked him why he hadn't spoken up about the abuse:

> AN.92 [AO] trusted Jonathan Lord and he saw that I also trusted and relied on him. [AO] was so embarrassed

AN.93 to tell me about what had happened. On the night that he first told me about it, I asked him why he

AN.94 hadn't told me, and he said, "Because I didn't think you'd believe me." I think that he saw that I liked

AN.95 Jonathan and thought that he might get in trouble or make me upset if he said anything.

Here, AN is being protective of her son, effectively saying that he shouldn't be criticised for not speaking up because she had created an environment where it was hard for him to speak up.

Another parent, AS, describes how her son AH attended the YMCA vacation care centre in Carringbah at least twice in each school holiday period from July 2009 to October 2011. It was later learned that AH had been sexually abused by Mr Lord, a staff member at the centre, during the 2010-11 Christmas vacation, and also in April 2011 and July 2011. On at least one occasion, the abuse occurred when Mr Lord was seated between AH and his sister on a bus taking the children on an excursion.

AS noticed that in early 2011 AH started to behave differently:

AS.79 From around the beginning of 2011, I observed that [AH] started rejecting physical contact with people,

AS.80 including his dad and his grandparents. He wouldn't kiss, cuddle or handshake. This affected his

AS.81 relationships with his father and his grandparents.

AS.82 [AH] became increasingly anxious and clingy with me.

AS.83 I was the only person who could get close to him. He would come to our bedroom six times in a night,

AS.84 and in the end he would just climb into bed and would just want to hold my hand.

However, AS records that AH still wanted to attend the YMCA centre, and "would speak a lot about Jonathon Lord." Mr Lord was also known as "Jon." The witness told the Commission that AH "would always speak about Jon in positive terms."

It was similar in the case of AE, the child of AS. As noted above, AE had himself proposed Mr Lord as a person who could babysit AE and his siblings. Although this suggestion was not taken up, and AS was aware that other childcare centres and preschools previously attended by her children "had policies which prevented babysitting of children by staff" (lines AS.24-25), AS "didn't think this [babysitting suggestion] was particularly out of the ordinary." AS records that AE continued to enjoy attending the YMCA events:

> AS.29 I remember the day that my son went on his excursion very clearly. He was eight and he was extremely
> AS.30 excited, and he was extremely excited to be going to the rock climbing centre. When he came home, he
> AS.31 told me that he and his friend had gotten special treatment from Jonathan Lord, who had apparently
> AS.32 bought them both hot chips and cookies from McDonald's.
> AS.33 Again, I didn't think this was particularly suspicious at the time, although I had packed lunch and I do not
> AS.34 recall that the excursion was supposed to stop on the way home.

Here, similarly to the other cases above, the child is portrayed as having been keen about engaging with the planned activity, and happy to be the beneficiary of "special treatment" by Mr Lord. Writing her narrative after the abuse of AE has been disclosed, AS effectively pre-empts issues that might have raised suspicion – i.e., the packed lunch and whether the excursion was supposed to stop on the way home – by recording these factors at the same time as saying that she did not see the situation is being "suspicious at the time."

The persistence of belief

In about September 2011, Mr Lord visited AN (mother of AO) accompanied by his own mother, Ms Yankos. It was disclosed by Ms Yankos that Mr Lord had been suspended from working at the YMCA "because a child made an allegation... that Jon touched him inappropriately on the bus." However, AN's belief in Mr Lord was not shaken. AN recalls saying:

> 72 How could they make that allegation against you? We need to get you a lawyer. There's a guy in
> 73 Cronulla I can recommend.

AN is candid about her trust in Mr Lord, and her bias against information contrary to her pre-existing belief:

> 80 At the time of the above conversation, I still had total trust in Jonathan. It never crossed my mind that
> 81 the alleged inappropriate touching could have actually happened.
> 82 Jill said that they had gone to see a solicitor and she asked me to provide a character reference for
> 83 Jonathan, which I agreed to do.

Later, after reflecting on the conversation, AN spoke to her son, who tragically disclosed that he had been abused by Mr Lord "every Saturday since the first time I met him."

5.2 Co-workers

Twelve employees and former employees of the YMCA gave evidence in the Royal Commission's case study relating to Mr Lord. However, several of these were involved in management capacities, such as to give evidence about the systems employed by the YMCA, and had little or no direct involvement with Mr

Lord. The primary focus of this analysis is co-workers who worked closely with Mr Lord or at least had relevant contact with him in the YMCA workplace where he was employed. I have identified four such workers.

The usual names of the co-workers were associated with their evidence, and I have therefore used their names in this analysis. The Royal Commission did not make any findings against these co-workers, and it is clear from their evidence that they were universally devastated by Mr Lord's abuse of the children in the YMCA's care.

He was "a great worker… very enthusiastic"

YMCA employee Danielle Ockwell ("DO"), talking about how she felt betrayed by Mr Lord's abuse of the children in their care when that abuse came to light, spoke of Mr Lord's inspirational approach and how he was seen by others:

> DO.581 … he made me believe that we had this amazing centre
> DO.582 where the children were safe and happy, and he came across to parents as a lovely guy…

Another YMCA employee, Chloe Starr ("CS") worked with Mr Lord shortly before he moved to the Caringbah centre. Although "in hindsight" Ms Starr agreed that Mr Lord had tended to focus on "some children more than others" (lines CS.320-321), one factor that really stood out was his enthusiasm:

> CS.323 Q. In hindsight, is there any other observation that you had made of him that you now consider in a
> CS.324 different light?
> CS.325 A. At the time when I was being trained by him, his enthusiasm I just thought was - I just thought he
> CS.326 was a great worker and he was very enthusiastic about his job and I just thought he had a great passion

CS.327 for his job. But now, when I look back and I have learnt about grooming, I see that that may have been

CS.328 grooming.

The possibility that this perception of enthusiasm was misconceived was also identified in relation to Mr Lord undertaking babysitting:

CS.331 Q. You say in paragraph 22 that you are now suspicious of - however, you weren't earlier - how much he

CS.332 babysat?

CS.333 A. Yes.

CS.334 Q. That didn't occur to you at the time - a 24-year-old young man who was spending his weekends

CS.335 babysitting young boys?

CS.336 A. At the time, no. Again, he was so enthusiastic about children; I just assumed that that's what he

CS.337 enjoyed doing.

CS.338 Q. You refer in your statement to him having babysitting jobs arranged for Friday, Saturday and Sunday

CS.339 nights?

CS.340 A. Yes. So as we would prepare the afternoon tea before the children had arrived, often you would

CS.341 speak about the staff, what you did on the weekend, and things like that, and he would always say he

CS.342 was basically babysitting.

However, Mr Lord's babysitting did not trigger any alarm bells for Ms Starr, because it was possible to draw a logical connection between his enthusiasm regarding children and the extent of his babysitting activities – "what he enjoyed doing."

Alicia Dellaca ("AD") started working for YMCA as a casual childcare assistant in about 2006, although that was initially just for about a week before she went overseas for 6 weeks. When she returned from her trip, a childcare assistant position was available, and the witness says she commenced working in that capacity and was then working with Mr Lord. In her written statement to the Royal Commission, Ms Dellaca refers to a conversation she had with Mr Lord, where Mr Lord said he had a "soft spot for kids and boys" (lines AD.800-801). Ms Dellaca says this conversation was when she first met Mr Lord and she "thought he has very clear goals and aspirations for his career within the childcare sector" (lines AD.805-806). Under cross-examination Ms Dellaca then asserts that this initial conversation influenced her thinking when she observed Mr Lord with children:

> AD.809 A. Any time I would - as I've listed in my statement, any time I would see him paying attention to
> AD.810 children or going above and beyond in his job role or wanting to stay behind to finish off games, that first
> AD.811 statement that he made to me made me believe that his intentions were good and that, in fact, he just
> AD.812 wanted to have a positive impact on children's lives, and I thought that was a - at the time, I thought that
> AD.813 was a realistic thing and he gave me no reason to believe otherwise.

In terms of forming a bias, it seems very likely that it was this initial conversation with Mr Lord that conditioned the witness to accept his subsequent behaviour in a positive light rather than as the grooming that it was. We cannot draw any conclusion here about what other people might think, hearing Mr Lord speaking of his "soft spot for kids and boys", but the witness is candid that this initial conversation convinced her of Mr Lord's bona fides. This is a good example of the 'primacy effect', discussed in chapter 3, where initial observations or information received by an individual about an issue tend to inform the interpretation of subsequent information on the issue.

It appears that on some occasions Mr Lord would also break or stretch the rules. For example, another co-worker Carine Beer ("CB") recounts how Mr Lord used his mobile phone at work and allowed a child to play with the phone even though both of these were against the rules:

> CB.336 Q. Were there any other occasions when you worked with him where you saw him using his mobile
> CB.337 phone?
> CB.338 A. Yes.
> CB.339 Q. Were there occasions when you saw him having children playing with his mobile phone?
> CB.340 A. Only a special needs child. That's the only one I saw.

What is the significance of the child being “only a special needs child”? It is not clear at this stage of the narrative whether the witness is merely describing the “only” child that she saw, or is attributing some significance to the fact of the child being “special needs.” In the latter case, this could be an example of a fallacious justification for another person breaching known rules – i.e., evidence of Mr Lord's misconduct being framed as a sign of his exceptional dedication to children. This may fall within the scope of *argumentum ad verecundiam* (fallacious appeal to authority) in the sense that Mr Lord could have been regarded by Ms Beer as the best authority on what conduct was appropriate in the circumstances, or alternatively a negative fallacy *ad consequentiam* (argument based on consequences). The latter construction could derive from Ms Beer perceiving the possibility of unfair consequences for Mr Lord and the child if the rules were applied strictly. The cross-examination of Ms Beer continued:

> CB.341 Q. You said earlier that as part of your employment process as coordinator; you were told that there
> CB.342 were to be no mobile phones at work?
> CB.343 A. Yes.

CB.344 Q. So what did you do when you saw Jonathan Lord with his mobile phone?

CB.345 A. Nothing.

CB.346 Q. You didn't report it to anyone?

CB.347 A. I may have. I'm not sure.

CB.348 Q. If you did, would it have been in writing?

CB.349 A. No.

CB.350 Q. You would have done it orally?

CB.351 A. Yes.

CB.352 Q. Do you have a recollection of having done that?

CB.353 A. I can't remember, no.

The witness appears to have a poor memory of the occasion. Initially she says she did nothing when she saw Mr Lord with his phone. Then she is asked a leading question at line CB.346, that may have prompted her to consider the possibility that she made a report. However, she is clear about not having made a report in writing, because that is not what she would have done. She's relying on her memory of what she would usually have done, not any memory of the specific occasion. When asked expressly whether she recalled making an oral report, she gives an ambiguous response – "I can't remember, no."

CB.354 Q. Well, did you think, when you saw him with his mobile phone, that what he was doing was against

CB.355 the rules

CB.356 A. It was for a special needs child and the child really loved The Wiggles, so it was just The Wiggles on it

CB.357 for him.

CB.358 Q. So did you think it was against the rules?

CB.359 A. Yes.

CB.360 Q. So did you then tell a supervisor that you had seen him –

CB.361 A. No.

This is the end of the cross-examination about the phone at that stage. Although the witness is vague, when she is able to be, such as about whether she told anyone, when it comes to the crunch she knows full well that Mr Lord having the phone was against the rules and that she failed to tell a supervisor. No more “I can’t remember.”

The kids didn't say anything

Mr Lord's co-worker, Alicia Dellaca observed that some of the children – specifically the boys – "gravitated to" Mr Lord:

> AD.823 A. I saw that he would often have certain children that he was more likely to do activities with, whether
> AD.824 it be playing sport games or whatnot. It just seemed that the same children seemed to gravitate to him.
> AD.825 Q. What did you think of that?
> AD.826 A. At the time I thought that was typical of boys.
> AD.827 I thought the boys that gravitated to him would have done so because he was one of two male staff
> AD.828 members, so I thought that explained the behaviour.
> AD.829 Q. And he wasn't particularly interested in playing with the little girls?
> AD.830 A. No.
> AD.831 Q. Did that give rise to any thought in you?
> AD.832 A. No, it didn't

Although Ms Dellaca had expressed some concerns about "other people thinking [Mr Lord] had favourites" she was not otherwise concerned about the behaviour at that time:

> AD.838 Q. You were not aware that children sitting on laps provided an access that would not otherwise be

AD.839 available?

AD.840 You didn't think about it in those terms, given your training?

AD.841 A. No, I believe I thought, considering the fact it was done in such a public forum in clear view of other

AD.842 staff, I guess I believed that grooming and other sexually inappropriate behaviours would be done out of

AD.843 sight.

It was put to Ms Dellaca that before the "Lord incident" she "had had a deal of training concerning child safety issues and risks" (lines AD.935-936). The witness effectively denied having (back then) the necessary understanding to recognise Mr Lord's visible behaviours as grooming, but spoke of gaining an improved understanding subsequently:

AD.954 Q. What have you gained from your experience as opposed to the training?

AD.955 A. My experience has given me a first-hand example of how discreet grooming can be and how

AD.956 manipulative paedophiles can be. I believe those two factors contribute to my now understanding.

The persistence of belief

It is recorded above that Mr Lord's co-worker, Alicia Dellaca formed a positive view of Mr Lord's intentions from his assertions that he had a "soft spot for kids and boys" (lines AD.800-801) and his apparently "very clear goals and aspirations for his career within the childcare sector" (lines AD.805-806). In another exchange, Ms Dellaca says Mr Lord questioned the 'no touching' rules:

AD.780 I recall working with Jon Lord one Thursday afternoon at the St Pat's centre, and he said to me, "I think

AD.781 that it's just so bizarre that they say it's important to build strong relationships with families and kids and

AD.782 to be a good role model for them yet the rules say 'don't let kids come close to you' and 'don't do this or

AD.783 that'.

AD.784 I said to him, "You can have a strong relationship as a role model with them without physically touching

AD.785 them, it's about protecting yourself and the

AD.786 children." I remember thinking to myself, "this isn't rocket science Jon, why is this even a dilemma".

It seems that the witness came close to seeing Mr Lord's comments as concerning, but ultimately dismissed them as him being a bit dramatic rather than an indication of an issue that she should have been alarmed by. Even when she thought it was "odd" and "strange" that Mr Lord was coming to the St Pat's centre when he was not rostered on, the witness was not sufficiently alarmed by that conduct to challenge it (lines AD.787-794).

Similarly, despite several incidents that might potentially have given cause for concern, another co-worker Danielle Ockwell did not regard them as sufficiently concerning to report them. For example:

Hugging-

DO.180 Q. Did you see Jonathan Lord hugging children?

DO.181 A. Maybe on the odd occasion, yes.

DO.182 Q. What did you do on the odd occasion you saw him hugging children?

DO.183 A. I would do nothing.

DO.184 Q. Even though you thought that was unnecessary touching?

DO.185 A. Yes.

Here we have conduct by Mr Lord that Ms Ockwell regards as "unnecessary", but not raising concerns that needed to be reported. In the context of a belief that Mr Lord was up to no good (if that belief had existed) we can imagine that Ms Ockwell would have been more inclined to report the touching. At the risk of falling into fallacious error, it may be reasonable to infer that Ms Ockwell's thinking at the time was informed by the context of her belief as to Mr Lord's good intentions.

Focusing on specific children-

> DO.198 … I would often spread myself around the whole centre and
> DO.199 interact with every child, so it made it quite difficult on my job when he would only focus on seven or so
> DO.200 children and I had to focus on the rest.
> DO.201 Q. There is a difference, isn't there, Ms Ockwell, between that making it difficult for you to do your job
> DO.202 and making you a bit uncomfortable?
> DO.203 A. Yes.
> DO.204 Q. What was it that made you a bit uncomfortable?
> DO.205 A. I just thought that him not going to other children was a bit - you know, it was strange, because when
> DO.206 we work in childcare; we have to look after all the kids. But yes.
> DO.207 Q. Did you share your views with anyone that you thought it was strange?
> DO.208 A. No, I did not.

The words "made you a bit uncomfortable" were put to Ms Ockwell under cross-examination and so were not her own words, but she effectively adopts them in describing Mr Lord's conduct as "strange." However, Ms Ockwell's concerns appear to be driver by operational considerations – the need "to look after all the kids" – rather than child safety considerations. Put simply, Ms Ockwell's interpretation of the situation was in the context of expecting Mr Lord to pull his weight in the centre, and there was no heuristic motivation to interpret the situation from a trust perspective.

Child's picture on Mr Lord's mobile phone-

DO.234 ... I saw that the screensaver on Jon's phone was a picture of [AF], who was a child in vacation

DO.235 care. I said to Jonathan Lord, "Oh that's [AF] from vacation care" and Jon said "Yes, that's the boy. I've

DO.236 forgot his name." I didn't say anything else but I thought that it was a bit odd for him to have a child from

DO.237 the centre as his screensaver.

Ms Ockwell doesn't identify this as improper or aberrant behaviour by Mr Lord, just that it was "a bit odd." This particular terminology – i.e., "a bit" - tends to downplay the oddness and excuse the failure to inquire more about the picture on the phone. The use of words like "strange" and "odd" appears likely to represent a subconscious minimisation of seriousness, that may seek to avoid accountability for non-reporting. In relation to confirmation bias, this situation may suggest that Ms Ockwell's established perception of Mr Lord as trustworthy led to a failure to test the "discrepant evidence" – i.e., there was a failure to consider that Mr Lord having the child's picture on his phone might have indicated improper motives regarding that child (Kuhn, 1989, p.677).

Babysitting 'incident'-

DO.241 On another occasion, Jon had told me that he was going to babysit [BA]. The following day he called me

DO.242 on my mobile and told me about it. I asked him how it had gone and he said, "Yeah, it was good, [BA] is

DO.243 a really good kid.” I said, "Yeah he is good" and Jon replied, "Yeah, I love him."

DO.244 I thought that was strange and that it was something that you shouldn't say. But at the time, I just

DO.245 thought that Jon was a loving and friendly guy and that was just his personality.

DO.246 Q. Ms Ockwell, you didn't tell anyone else at work at the YMCA about the discussion with Jonathan Lord

DO.247 in relation to [BA], did you?

DO.248 A. No, I did not.

DO.249 Q. You didn't tell anyone about having seen [AL] on his lap for most of the day?

DO.250 A. No, I did not.

Once again, a non-reported situation is somewhat minimised by use of the word "strange" rather than a term that would regard the situation more seriously. Although a connection can be made between Mr Lord's comments about the first boy and Mr Lord subsequently having another boy on his lap "for most of the day" it is also possible to separate these matters – i.e., different children, different activity – which makes it relatively easy to disregard the connection if there is no existing belief that would promote concerns about the overall situation.

Being fired from working at a children's summer camp-

DO.255 Q. In paragraph 24 you say that Jonathan Lord told you, not long before he was arrested, that he had

DO.256 been fired from a summer camp?

DO.257 A. Yes, that's correct.

DO.258 Q. Can you tell us the context in which you had that conversation with him?

DO.259 A. It was during before school care. I was actually going to apply to summer camp because I was

DO.260 interested in doing it. I had mentioned that to him, and he then told me that he was involved in a camp.

DO.261 I obviously began to ask questions about the camp, like what was it like and he then told me that he had

DO.262 been sent home from camp because he was caught with a boy one on one, but he said that the camp

DO.263 had misunderstood him.

DO.264 Q. What did you understand him to mean by being caught "one on one" with a boy?

DO.265 A. Well, I just know, working in childcare, you are not supposed to be one on one, so I just thought that

DO.266 he happened to get caught just with a child and they didn't approve.

DO.267 Q. You didn't tell anyone in management or a supervisory role about that conversation?

DO.268 A. No, I did not.

DO.269 Q. Why not?

DO.270 A. I just didn't really think much of it at the time.

DO.271 Q. Well, you knew to be one on one was the wrong thing to do?

DO.272 A. Yes.

DO.273 Q. And he was telling you that he had done the wrong thing so as for him to have been fired working

DO.274 with children before he came to the YMCA?

DO.275 A. Yes.

DO.276 Q. You didn't think that was an important thing that your managers needed to know?

DO.277 A. I guess, looking at it now, yes. At that time, when he told me, I just - he told me that, you know, they

DO.278 misunderstood, and I just trusted him.

Much of the context here is of limited import, but there is an interesting aspect to this which is Mr Lord's self-reporting of the dismissal from the summer camp position. Although disclosure of this history by Mr Lord might be seen as contrary to his own interests, by making the disclosure to Ms Ockwell voluntarily and with no apparent pressure to do so he is both maintaining control of the narrative and promoting himself as an honest, forthright person who has nothing to hide. The manner of the disclosure may have had the perverse effect of building rather than

diminishing Ms Ockwell's trust in Mr Lord. Ms Ockwell acknowledges that she trusted Mr Lord and did not think to refer the disclosure to her employers.

Babysitting for free-

> DO.435 [in relation to] babysitting [BA], … [BA]'s mother being told by Jonathan Lord that if she can't afford it,
> DO.436 he will do it for free; do you see that?
> DO.437 A. Yes.
> DO.438 Q. Did that strike you as unusual, that a childcare assistant, who we all know doesn't earn a lot of
> DO.439 money, was prepared to babysit for free?
> DO.440 A. No.
> DO.441 Q. It didn't?
> DO.442 A. I didn't think it was unusual.
> DO.443 Q. Had you ever babysat any of the children at the YMCA for free?
> DO.444 A. No, I didn't.
> DO.445 Q. You didn't think it odd that a 24-year-old man would be offering to babysit an eight-year-old boy for
> DO.446 free?
> DO.447 A. I just thought that he was just a nice guy.
> DO.448 Q. What would you think now if a 24- or 25-year-old man wanted to sit one of the children that you
> DO.449 worked with for free?
> DO.450 A. I would be concerned.

Although the cross-examination here does not test whether Ms Ockwell believed that babysitting is a 'job' that should typically be remunerated, we might reasonably infer that she did not regard babysitting as a normal part of her own work, and therefore did not regard it as a normal part of Mr Lord's work as a childcare assistant. In that context, the fact of Mr Lord being willing to go 'above and beyond' in this manner might in some circumstances have rung alarm bells

for Ms Ockwell, but an adverse interpretation of Mr Lord's conduct was not tested due to the primacy effect of Ms Ockwell's trust in Mr Lord and her perception of him as "just a nice guy."

Ms Ockwell then acknowledges to the Royal Commission that she regarded Mr Lord as a close friend. This friendship and its implications were explored through cross-examination, leading to the following exchange:

DO.618 Q. It is fair to say, isn't it, that until the time that you learnt of the kinds of things that Mr Lord had been

DO.619 charged with, it simply did not cross your mind for a single moment that any of the conduct that was

DO.620 playing out right before your eyes was anything other than him being an affectionate, enthusiastic

DO.621 childcare worker?

DO.622 A. Yes, I trusted him.

DO.623 Q. And you have said that a few times now, but just to complete that, you trusted him to keep the

DO.624 children safe?

DO.625 A. Yes, I did.

DO.626 Q. It was that trust, I take it, that cancelled out in your mind even the merest possibility that what you

DO.627 were seeing played out before you could in any way be interpreted as suspicious or worse conduct on his

DO.628 part?

DO.629 A. Yes.

This narrative largely speaks for itself. In hindsight, Ms Ockwell realises how she has been misled by her confidence in Mr Lord's trustworthiness. However, caution should be exercised in analysing this particular narrative, because much of the cross-examination is *leading* – e.g., through the use of closed questions

that seek a 'yes' or 'no' answer to a proposition. Ms Ockwell agrees with what is put to her, but it is not her own words. The most that we can say is that Ms Ockwell confirmed under cross-examination that she trusted Mr Lord, and was not inclined to deny any of what was put to her as to the consequences of "that trust."

Chapter 6 – Summary and Conclusions

My primary objective in this thesis has been to describe the phenomenon of confirmation bias, and to illustrate the implications of it using a real-life situation.

Chapter 2 introduced human biases and cognitive heuristics generally, as the context for my focus on confirmation bias. Chapter 3 describes confirmation bias itself, and how we form beliefs that influence our subsequent thinking and decision-making. The net effect of this is that the formed belief gains a privilege in our minds that may overwhelm a rational assessment of relevant information. For example, if we form a belief that a specific person is trustworthy, we may dismiss or de-emphasise information that is inconsistent with that belief.

The real-life situation studied here is the sexual abuse of at least 12 boys aged 6 to 10 over a period of about 2 years, in circumstances where their abuser – Jonathan Lord – was a trusted employee of the YMCA, Australia's second largest provider of 'outside school hours care' (Royal Commission, 2014, p.22).

Several parents of the abused children, and several of Mr Lord's co-workers, gave evidence to the Royal Commission. Chapter 4 introduces discourse analysis as the methodology used in this study. The transcripts of the witnesses' evidence are reviewed in chapter 5 using a form of discourse analysis.

Misplaced trust

Several parent witnesses referred to Mr Lord's role as an employee of the YMCA, and inferred from this that he was himself trustworthy, for example: (AN, mother of victim AO, lines AN.29-30)

> "I trusted and believed that if Jonathan Lord was a supervisor at YMCA, he was more than qualified to babysit…"

The critical word here is "trust", which conveys a broad notion of caring and keeping children safe. In at least a couple of cases, parents' confidence in Mr Lord as a prospective carer was also bolstered by a recommendation from Mr Lord's own mother. Such a recommendation is likely to be a powerful influence, because it inherently involves one mother – someone who might be presumed to know the qualities of her adult son – speaking directly to another mother. The first mother is effectively putting herself forward as an authority figure, and the person hearing the recommendation may therefore fall victim to the fallacy *argumentum ad verecundiam* ('fallacious appeals to authority' – see chapter 2). In the case of the child AO, his mother (AN) records that Mr Lord's mother Ms Yankos was "very proactive in recommending her son to me as a babysitter" (lines AN.26-27).

The victims' mothers were also encouraged by Mr Lord's apparently selfless conduct, and the positive reactions of their own children. In one case, Mr Lord became "like a member of the family" and would "drop by" and help out (line AN.43). Consistently, the children – even some of those who were being abused by Mr Lord – extolled his virtues, for example:

- "I understood from speaking with my children that they had grown very fond of Jonathan Lord because he made a big effort to entertain them and pay them attention" (AU, mother of victim AM, lines AU.23-24).

- "[AO] and other kids appeared to love Jonathan. … [AO] and his friends fought about who was going to be on Jonathan's team and who would get to ride in Jonathan's car. Jonathan had a way with children" (AN, mother of victim AO, lines AN.47-49).

These discourses suggest that parents were lulled into a false sense of security through the belief that their children's reactions to Mr Lord would be different if he was behaving inappropriately towards them. This connection between Mr Lord and the children was also perceived in the context of the parents' "trust" in Mr Lord arising from knowing of his involvement with YMCA. In the absence of that

basic building block of trust, the parents may perhaps have been more suspicious of the attention that Mr Lord paid to the children and his efforts to entertain them. This is a fairly classic manifestation of confirmation bias, where subsequent information – in this situation, Mr Lord's conduct towards the children and their reactions to it – is assessed in terms of beliefs already formed – e.g.., beliefs as to a person's trustworthiness.

This was no accident. Just as Mr Lord had a "way with children", he also had a way with their parents. As one mother put it: "In hindsight, I now see that Jonathan Lord groomed me as well as [my son]" (AN, mother of victim AO, line AN.42).

Mr Lord's co-workers were similarly beguiled by Mr Lord's enthusiasm and apparent sincerity:

- "he made me believe that we had this amazing centre" (line DO.581).
- "he came across to parents as a lovely guy" (line DO.582).
- "I just thought he was a great worker and he was very enthusiastic about his job and I just thought he had a great passion" (lines CS.325-6).
- "[the] statement that he made to me made me believe that his intentions were good and that, in fact, he just wanted to have a positive impact on children's lives" (lines AD.811-2).

In this context, words spoken by Mr Lord that might have raised suspicions in the minds of someone removed from the situation – e.g., Mr Lord speaking to Ms Dellaca of his "soft spot for kids and boys" (lines AD.800-1), and saying "Yeah, I love him" to Ms Ockwell when discussing his feelings towards a child he had babysat (line DO.243) – did not trigger any concern for co-workers. Even clear breaches of YMCA rules were interpreted in a positive light, for example the situation observed by Ms Beer where Mr Lord allowed a child to play with his

mobile phone: "It was for a special needs child and the child really loved The Wiggles, so it was just The Wiggles on it for him" (lines CB.356-7).

As previously discussed in chapter 5, Ms Beer's interpretation of the breach of rules relating to the mobile phone may fall within the scope of *argumentum ad verecundiam* (fallacious appeal to authority) in the sense that Mr Lord could have been regarded by Ms Beer as the best authority on what conduct was appropriate in the circumstances, or alternatively a negative fallacy *ad consequentiam* (argument based on consequences) – i.e., Ms Beer perceiving the possibility of unfair consequences for Mr Lord and the child if the rules were applied strictly. However, we are undoubtedly also seeing the primacy effect of confirmation bias, where existing beliefs as to the commitment and trustworthiness of Mr Lord affected the decision making of these co-workers. This in no way makes them accountable for Mr Lord's criminal conduct, but shows how even the most well-intentioned individuals can be deceived by a manipulative offender.

A significant aspect of confirmation bias is how it affects decision making. In relation to parents, the trust that parents invested in Mr Lord led to decisions to allow him to babysit their children in some cases, and to allow children to go on excursions with him. These decisions were not unreasonable in their context, but part of that context – as was the case with the co-workers – was the parents' established bias towards regarding Mr Lord as reliable and trustworthy. In some cases that bias was encouraged by the children's own endorsement of Mr Lord as a person they liked and could relate to.

The extent of the parental and co-worker bias is indicated by how their trust in Mr Lord persisted despite incidents and disclosures that in other circumstances may have been seen as pointers to the possibility of Mr Lord being a child abuser. In the case of co-workers, Ms Ockwell saw Mr Lord hugging children and focusing on specific children to the exclusion of others, as well as incidents that she regarded as "odd" (Mr Lord having a vacation care child's picture as his phone screen-saver – lines DO.234-7) and "strange" (Mr Lord saying he "loved" a child

who he babysat – line DO.243). In addition, Ms Ockwell learned that Mr Lord had been dismissed by a children's summer camp for being "one on one" with a child, but accepted Mr Lord's explanation that (in her words) "he happened to get caught just with a child and they didn't approve" (line DO.266).

A similar persistence of belief in Mr Lord's trustworthiness is evident from the evidence of AN, mother of victim AO. It appears certain that Mr Lord groomed both AO and AN, including taking time to build a connection with AN and showing interest in her and her family. By any measure, Mr Lord's engagement with that family, including dropping by the home, sharing meals at the home, and spending time watching television at the home, went significantly beyond the level of contact that would usually occur between a child care worker and the family of a child enrolled with the child care service employing the worker. We now (from AN's evidence to the Royal Commission) have cause to believe that AO might have been sexually abused by Mr Lord on as many as 100 occasions, but AN and her family were oblivious to this at the time. Even when Jill Yankos, Mr Lord's mother brought him around to AN's home and disclosed that Mr Lord was suspended by the YMCA due to an allegation of inappropriate touching, AN retained her trust in Mr Lord:

> "[I said] How could they make that allegation against you? We need to get you a lawyer. There's a guy in Cronulla I can recommend… I still had total trust in Jonathan. It never crossed my mind that the alleged inappropriate touching could have actually happened... [Jill] asked me to provide a character reference for Jonathan, which I agreed to do" (lines DO.72-83).

Closing

This thesis considers a very small aspect of the evidence heard by the Royal Commission, and is not intended to be an exhaustive or even comprehensive review of the relevant evidence about Mr Lord. My purpose has been to describe confirmation bias and illustrate the operation of it in a real-life scenario. My study

shows how confirmation bias may lead to tragic or destructive outcomes in some circumstances.

Illustration of the primacy effect and other implications of confirmation bias is helpful in demonstrating how confirmation bias can operate insidiously to facilitate child sexual abuse. The use of discourse analysis in this study has afforded a 'micro context' understanding of how Mr Lord's abuse of children associated with the YMCA service persisted for over a year undetected. While discourse analysis does not provide 'black and white' answers to relevant questions, it can nonetheless inform theoretical models of behaviour that improve our understanding of how such abuse might be detected sooner.

The main focus of the Royal Commission was *institutional responses* to child sexual abuse. Arguably, on the basis of what I have found here, that focus is wholly appropriate. There is no doubt that the co-workers and parents who gave evidence to the Royal Commission were concerned for the welfare of the children involved, and honestly believed that the children were safe, but were individually powerless against the deliberate and callous acts of a highly manipulative sexual predator. It seems very likely that the safety of children in the context of engagement with institutional employees and contractors can only be assured by having robust systems to check the bona fides of prospective employees and manage their compliance with relevant policies. As can be seen here, even just the fact of a person being employed by a reputable institution was enough to overcome barriers than might otherwise have improved the safety of children in situations outside the scope of the service offered by the institution. This puts an onus on such institutions to not only have rules regarding workplace conduct, but to also adopt a holistic approach to child safety that recognises the way in which the trust of other people is gained and potentially abused.

Accordingly, one path to improving children's safety almost certainly lies in the direction of ensuring that institutions involved with the welfare of children implement appropriate systemic approaches to protecting children. That said,

such approaches will never be enough on their own. Optimal outcomes will only be achieved through a combination of systemic approaches and the training of individual staff and workers to ensure they understand their individual susceptibility to confirmation bias and can recognise abusive situations and respond to them appropriately despite that susceptibility.

S. Bourne

References

Aue, T. & Okon-Singer, H. (2015). Expectancy biases in fear and anxiety and their link to biases in attention. *Clinical Psychology Review, 42*, 83-95. DOI: 10.1016/j.cpr.2015.08.005.

Austerweil, J.L. & Griffiths, T.L. (2011). Seeking confirmation is rational for deterministic hypotheses. *Cognitive Science, 35*(3), 499-526. DOI: 10.1111/j.1551-6709.2010.01161.x.

Back, I.H. (2010). Commitment bias: Mistaken partner selection or ancient wisdom? *Evolution and Human Behavior, 31*(1), 22-28. DOI: 10.1016/j.evolhumbehav.2009.06.006.

Bargh, J.A. (2014). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition (Chapter 1). (In) R.S. Wyer Jr. & T.K. Srull, *Handbook of social cognition: Volume 2*, 1-40. US: Psychology Press.

Barros, G. (2010) Herbery A. Simon and the concept of rationality: Boundaries and procedures. *Brazilian Journal of Political Economy*, 30(3), 455-472.

Biddington, M. (2017). *Funding for the Royal Commission into institutional responses to child sexual abuse: Budget review 2013-14 index.* Commonwealth of Australia, http://www.aph.gov.au, Retrieved 17 January 2017.

Black, J. & Barnes, J.L. (2015). Fiction and social cognition: The effect of viewing award-winning television dramas on theory of mind. *Psychology of Aesthetics, Creativity, and the Arts, 9*(4), 423-429. DOI: 10.1037/aca0000031.

Bleich, E., Nisar, H. & Abdelhamid, R. (2016). The effect of terrorist events on media portrayals of Islam and Muslims: Evidence from New York Times headlines, 1985–2013. *Ethnic and Racial Studies, 39*(7), 1109-1127. DOI: 10.1080/01419870.2015.1103886.

Boudry, M., Blancke, S. & Pigliucci, M. (2014). What makes weird beliefs thrive? The epidemiology of pseudoscience. *Philosophical Psychology, 1-22*. DOI: 10.1080/09515089.2014.971946.

Brigham, J.C. & Wasserman, A.W. (1999). The impact of race, racial attitude, and gender on reactions to the criminal trial of O.J. Simpson. *Journal of Applied Social Psychology, 29*(7), 1333-1370.

Brockner, J. (1992). The escalation of commitment to a failing course of action: Toward theoretical progress. *Academy of Management Review*, 17(1), 39-61.

Byron, M. (2005). Simon's revenge: or, incommensurability and satisficing. *Analysis, 65*(4), 311-15.

Cacioppo, J., Petty, R. & Kao, C.F. (1984). The efficient assessment of need for cognition. Journal of *Personality Assessment, 48*(3), 306-307.

Carey, B. (2014). Remembering, as an extreme sport. *The New York Times*, http://www.nytimes.com, 19 May 2014.

Chase, V.M., Hertwig, R. & Gigerenzer, G. (1998). Visions of rationality. *Trends in Cognitive Sciences, 2*(6), 206-214.

Chow, S.J. (2014). Many meanings of ‘heuristic'. *The British Journal for the Philosophy of Science, 66*(4), 977-1016.

Clark, A. (1998). Magic words: How language augments human computation. (In) P. Carruthers & J. Boucher (Eds), *Language and thought: Interdisciplinary themes*, 162-183. Cambridge, UK: Cambridge University Press.

Cosgrove, P. (2017). *Terms of reference: Letters Patent* (Royal Commission into Institutional Responses to Child Sexual Abuse). Commonwealth of Australia, 13 November 2014.

Cullen, S. (2012). Gillard launches royal commission into child abuse. *ABC News,* Australian Broadcasting Corporation, 12 November 2012.

Cummings, L. (2014). Informal fallacies as cognitive heuristics in public health reasoning. *Informal Logic, 34*(1), 1-37.

D'Antonio, M. (2016). *The truth about Trump*. New York, NY: Thomas Dunne Books.

Davis, D.B. (2011). *The mystic on the concrete: Exploring how reported experiences of divine guidance are interpreted for meaning and decision-making*. Master's Thesis, Humboldt State University, May 2011.

Dawkins, R. (1986). *The blind watchmaker: Why the evidence of evolution reveals a universe without design*. New York & London: WW Norton.

Dawkins, R. (2012). *The magic of reality: How we know what's really true*. London, UK: Transworld Publishers.

de Waal, F. (2016). *Are we smart enough to know how smart animals are?* London, UK: Granta Publications.

Decker, C. (1995). The Times poll: Most in county disagree with Simpson verdicts. *Los Angeles Times*, 8 October 1995.

Devine, D.J. (2012). *Jury decision making: The state of the science*. New York, NY: New York University Press.

Dewey, J. (1933). Philosopher's search for the immutable (Chapter II). (In) *The question of certainty. New York*, NY: Capricorn Books.

Diemand-Yauman, C., Oppenheimer, D.M. & Vaughan, E.B. (2011). Fortune favors the Bold (and the Italicized): Effects of disfluency on educational outcomes. *Cognition, 118*(1), 111-115. DOI: 10.1016/j.cognition.2010.09.012.

Drew, T., Võ, M.L.H. & Wolfe, J.M. (2013). The invisible gorilla strikes again: Sustained inattentional blindness in expert observers. *Psychological Science*, 24(9), 1848-53. DOI: 10.1177/0956797613479386.

Efferson, C., Lalive, R., Richerson, P.J., McElreath, R. & Lubell, M. (2008). Conformists and mavericks: The empirics of frequency-dependent cultural transmission. *Evolution and Human Behavior*, 29(1), 56-64. DOI: 10.1016/j.evolhumbehav.2007.08.003.

Ellison, C.G. & Fan, D. (2008). Daily spiritual experiences and psychological well-being among US adults. *Social Indicators Research, 88*(2), 247-271. DOI: 10.1007/s11205-007-9187-2.

Enfield, N.J. (2014). *Natural causes of language: Frames, biases, and cultural transmission.* Berlin, Germany: Language Science Press, langsci.github.io.

Feder, K.L. (1984). Irrationality and popular archaeology. *American Antiquity, 49*(3), 525-541.

Fest, G. (2005). Offshoring: Feds take fresh look at India BPOs; Major theft has raised more than a few eyebrows. *Bank Technology News, 18*(9), 1 September 2005.

Furness, G.B. (2013). *The response of YMCA NSW to the conduct of Jonathan Lord: Submissions of Counsel assisting the Royal Commission.* Royal Commission into Institutional Responses to Child Sexual Abuse, 21 October 2013.

Garland, H. (1990). Throwing good money after bad: The effect of sunk costs on the decision to escalate commitment to an ongoing project. *Journal of Applied Psychology*, 75(6), 728-731.

Gawande, A. (2010). *Complications: A surgeon's notes on an imperfect science*. London, UK: Profile Books.

Gawande, A. (2010). *The checklist manifesto: How to get things right*. London, UK: Profile Books.

Gegenfurtner, K.R. (2016). The interaction between vision and eye movements. *Perception, 45*(12), 1333-1357. DOI: 10.1177/0301006616657097.

Gigerenzer, G. & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science, 1*(1), 107-143. DOI: 10.1111/j.1756-8765.2008.01006.x.

Gigerenzer, G. & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, 62, 451-482. DOI: 10.1146/annurev-psych-120709-145346.

Gigerenzer, G. (2004). Fast and frugal heuristics: The tools of bounded rationality (Chapter 4). (In) D. Koehler & N. Harvey (Eds.)., *Blackwell handbook of judgment and decision making*, 62-88. Oxford, UK: Blackwell.

Gigerenzer, G. (2001). Content-blind norms, no norms, or good norms? A reply to Vranas. *Cognition, 81*(1), 93-103.

Gigerenzer, G., Todd, P.M. & ABC Research Group (1999). How good are simple heuristics? (In) *Simple heuristics that make us smart*, 97-118. New York, NY: Oxford University Press.

Gilovich, T., Griffin, D. & Kahneman, D. (2002). Heuristics and biases: *The psychology of intuitive judgment*. New York, NY: Cambridge University Press.

Goodman-Delahunty, J., Cossins, A. & Martschuk, N. (2016). *Jury reasoning in joint and separate trials of institutional child sexual abuse: An empirical study*. Sydney, NSW: Royal Commission into Institution Responses to Child Sexual Abuse.

Graber, M.L., Franklin, N. & Gordon, R. (2005). Diagnostic error in internal medicine. *Archives of Internal Medicine, 165*(13), 1493-99.

Graczyk, T.K., Knight, R., Gilman, R.H. & Cranfield, M.R. (2001). The role of non-biting flies in the epidemiology of human infectious diseases. *Microbes and Infection, 3*(3), 231-235.

Hernandez, I. & Preston, J.L. (2013). Disfluency disrupts the confirmation bias. *Journal of Experimental Social Psychology, 49*(1), 178-182. DOI: 10.1016/j.jesp.2012.08.010.

Hertwig, R. & Herzog, S. M. (2009). Fast and frugal heuristics: Tools of social rationality. *Social Cognition, 27*(5), 661-698.

Hilmo, J., Naesheim, T. & Gilbert, M. (2014). 'Nobody is dead until warm and dead': Prolonged resuscitation is warranted in arrested hypothermic victims also in remote areas -- A retrospective study from northern Norway. *Resuscitation, 85*(9), 1204-1211. DOI: 10.1016/j.resuscitation.2014.04.029.

Jackson, B.R. & Bergeman, C.S. (2011). How does religiosity enhance well-being? The role of perceived control. *Psychology of Religion and Spirituality, 3*(2), 149-161. DOI: 10.1037/a0021597.

Jardri, R. & Denève, S. (2013). Computational models of hallucinations. (In) *The neuroscience of hallucinations*, 289-313. New York, NY: Springer.

Johnson, D.D., Blumstein, D.T., Fowler, J.H. & Haselton, M.G. (2013). The evolution of error: Error management, cognitive constraints, and adaptive decision-making biases. *Trends in Ecology & Evolution, 28*(8), 474-481. DOI: 10.1016/j.tree.2013.05.014.

Jones, E.E., Rock, L., Shaver, K.G., Goethals, G.R. & Ward, L.M. (1968). Pattern of performance and ability attribution: An unexpected primacy effect. *Journal of Personality and Social Psychology*, 10(4), 317-340.

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

Kashdan, T.B. & Nezlek, J.B. (2012). Whether, when, and how is spirituality related to well-being? Moving beyond single occasion questionnaires to understanding daily process. *Personality and Social Psychology Bulletin, 38*(11), 1523-1535. DOI: 10.1177/0146167212454549.

Kassin, S.M., Reddy, M.E. & Tulloch, W.F. (1990). Juror interpretations of ambiguous evidence: The need for cognition, presentation order, and persuasion. *Law and Human Behavior, 14*(1), 43-55.

Kelley, N. (2012). Why do flies exist? *Quora*, http://www.quora.com, 3 April 2012.

Keren, G. (1987). Facing uncertainty in the game of bridge: A calibration study. *Organizational Behavior and Human Decision Processes*, 39, 98-114.

Kessler, G. & Kelly, M. (2018). *President Trump has made more than 2,000 false or misleading claims over 355 days*. The Washington Post, 10 January 2018.

Kirsh, D. (1991). Today the earwig, tomorrow man? *Artificial Intelligence, 47*(1-3), 161-184.

Knobloch-Westerwick, S. & Kleinman, S.B. (2012). Preelection selective exposure: Confirmation bias versus informational utility. *Communication Research, 39*(2), 170-193. DOI: 10.1177/0093650211400597.

Konrad, B.N. (2014). *Characteristics and neuronal correlates of superior memory performance*. PhD dissertation, Ludwig-Maximilian University, Munich, October 2013.

Kruger, J. & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*(6), 1121-1134.

Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review*, 96(4), 674-689.

Langley, R.M. (2012). *Ameritocracy: Hollywood blockbusters and the universalisation of American values*. PhD thesis, College of Arts and Law, University of Birmingham, September 2011.

Lauwereyns, J. (2010). *The anatomy of bias: How neural circuits weigh the options*. Cambridge, MA: The MIT Press.

Leipold, A.D. & Abbasi, H.A. (2006). The impact of joinder and severance on federal criminal cases: An empirical study. Vanderbilt Law Review, 59(2).

Leontyev, A.N. & Cole, M. (2009). The development of mind: Selected works of Aleksei Nikolaevich Leontyev. Pacifica, CA: Marxists Internet Archive.

Levy, N. (2014). Addiction as a disorder of belief. Biology & Philosophy, 29(3), 337-355. DOI: 10.1007/s10539-014-9434-2.

Lock, A. & Shotter, J. (2011). *The Social Nature of Human Development.* Draft Chapter, 15 July 2011.

Lopez, L. (2016). One explanation for why Donald Trump forgets things the second he says them. *Business Insider,* http://www.businessinsider.com.au, 6 December 2016.

Lovallo, D. & Kahneman, D. (2003). Delusions of success. *Harvard Businss Review,* 81(7), 56-63.

Lustig, R. (2017). The Trump Inauguration: Darkness Descends. *The Huffington Post,* http://www.huffingtonpost.co.uk, 13 January 2017.

Macbrayne, R. (2000). Inquest focus on teen's breathing. *The New Zealand Herald,* http://www.nzherald.co.nz, 30 June 2000.

Malpas, J. (1997). Space and sociality. *International Journal of Philosophical Studies, 5*(1), 53-79. DOI: 10.1080/09672559708570845.

Markovitch, D.G., Huang, D., Peters, L., Phani, B.V., Philip, D. & Tracy, W. (2014). Escalation of commitment in entrepreneurship-minded groups. *International Journal of Entrepreneurial Behavior & Research*, 20(4), 302-323. DOI: 10.1108/IJEBR-08-2013-0127.

Marsh, D.M. & Hanlon, T.J. (2007). Seeing what we want to see: Confirmation bias in animal behavior research. *Ethology, 113*(11), 1089-1098. DOI: 10.1111/j.1439-0310.2007.01406.x.

Mattox, E.A. (2012). Strategies for improving patient safety: Linking task type to error type. *Critical Care Nurse, 32*(1), 52-78. DOI: 10.4037/ccn2012303.

McFerran, D. (2015). *Logic in practice: A first guide to both deductive (formal) and inductive (informal) reasoning*. Los Angeles, CA: Douglass McFerran.

McKay, R.T. & Dennett, D.C. (2009). The evolution of misbelief. *Behavioral and Brain Sciences, 32*(06), 493-510. DOI: 10.1017/S0140525X09990975.

Metzger, M.J. & Flanagin, A.J. (2013). Credibility and trust of information in online environments: The use of cognitive heuristics. *Journal of Pragmatics, 59*, 210-220. DOI: 10.1016/j.pragma.2013.07.012.

Miller, R.M. (1984). Science teaching for the citizen of the future. *Science Education, 68*(4), 403-410.

Miller, T. S. (2014). Forms of perspective and Chaucer's dream spaces: Memory and the catalogue in The House of Fame. *Style, 48*(4), 479-495.

Milliman, R.E. (1986). The influence of background music on the behavior of restaurant patrons. *Journal of Consumer Research, 13*(2), 286-289.

Milner, A., Kõlves, K., Kõlves, K., Gladman, B. & De Leo, D. (2013). Treatment priority for suicide ideation and behaviours at an Australian emergency department. *World Journal of Psychiatry, 3*(2), 34-40. DOI: 10.5498/wjp.v3.i2.34.

Mishara, A.L. & Corlett, P. (2009). Are delusions biologically adaptive? Salvaging the doxastic shear pin. *Behavioral and Brain Sciences, 32*(06), 530-531. DOI: 10.1017/S0140525X09991464.

Mlodinow, L. (2008). *The drunkard's walk: How randomness rules our lives*. London, UK: Allen Lane.

Moutoussis, M., Bentall, R.P., El-Deredy, W. & Dayan, P. (2011). Bayesian modelling of jumping-to-conclusions bias in delusional patients. *Cognitive Neuropsychiatry*, 16(5), 422-447. DOI: 10.1080/13546805.2010.548678.

Navarro, D.J. & Perfors, A.F. (2011). Hypothesis generation, sparse categories, and the positive test strategy. *Psychological Review, 118*(1), 120-134. DOI: 10.1037/a0021110.

Nelson, C. (2015). The domestic is political, and the political is gendered: An analysis of veiled subjects, gendered epistemologies, and Muslim bodies. *Islamophobia Studies Journal, 3*(1), 106-114.

Nickerson, R.S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*(2), 175-220.

Oppenheimer, D.M. (2008). The secret life of fluency. *Trends in Cognitive Sciences, 12*(6), 237-241. DOI: 10.1016/j.tics.2008.02.014.

Painter, R.W. (2004). Convergence and competition in rules governing lawyers and auditors. *The Journal of Corporation Law*, Winter 2004, 397-426.

Panda, S. (2014). Great expectations: The Pygmalion Effect. *Summer Ventures in Science and Mathematics*, http://cstem.uncc.edu, 2014.

Perfors, A. & Navarro, D. (2009). Confirmation bias is rational when hypotheses are sparse. *Cognitive Science Society, 31st annual conference, COGSCI 2009,* 2741-46.

Petrie, M. & Halliday, T. (1994). Experimental and natural changes in the peacock's (Pavo cristatus) train can affect mating success. *Behavioral Ecology and Sociobiology, 35*(3), 213-217.

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

Prime Minister (2012). *Transcript of press conference: Royal Commission into child sexual abuse.* Department of the Prime Minister and Cabinet, Australia, 12 November 2012.

Purcell, K., Rainie, L., Mitchell, A., Rosenstiel, T. & Olmstead, K. (2010). Understanding the participatory news consumer: How internet and cell phone users have turned news into a social experience. *Pew Internet and American Life Project,* 1 March 2010.

Qiu, L. (2015). Is Donald Trump's 'Art of the Deal' the best-selling business book of all time? *PolitiFact,* http://www.politifact.com, 6 July 2015.

Rassin, E. (2008). Individual differences in the susceptibility to confirmation bias. *Netherlands Journal of Psychology,* 64(2), 87-93.

Rawden, J. (2016). Why let's make a deal is so unpredictable, according to host Monty Hall. *Cinemablend,* http://www.cinemablend.com.

Reber, R. & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition, 8,* 338-342.

Redelmeier, D.A. & Scales, D.C. (2015). Missing the diagnosis of brain death as a self-erasing error. *American Journal of Respiratory and Critical Care Medicine, 192*(3), 280-282. DOI: 10.1164/rccm.201503-0499OE.

Roberts, J. (2012). Safeguarding the customer's interest. *Marketing Week,* http://www.marketingweek.com, 31 October 2012.

Rosenthal, R. (1963). On the social psychology of the psychological experiment: The experimenter's hypothesis as unintended determinant of experimental results. *American Scientist, 51*(2), 268-283.

Ross, H.J. (2014). *Everyday bias: Identifying and navigating unconscious judgments in our daily lives*. Lanham, Maryland: Rowman & Littlefield.

Ross, L., Lepper, M.R. & Hubbard, M. (1975). Perseverance in self-perception and social perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology*, 32(5), 880-892.

Rossmo, D.K. (2016). Case rethinking: A protocol for reviewing criminal investigations. *Police Practice and Research, 17*(3), 212-228. DOI: 10.1080/15614263.2014.978320.

Royal Commission (2014). *Report of case study no. 2: YMCA NSW's response to the conduct of Jonathan Lord.* Barton, ACT: Royal Commission into Institutional Responses to Child Sexual Abuse, June 2014.

Royal Commission (2014a). *Interim Report: Volume 1.* Barton, ACT: Royal Commission into Institutional Responses to Child Sexual Abuse.

Royal Commission (2017). *A brief guide to the final report.* Sydney, NSW: Royal Commission into Institutional Responses to Child Sexual Abuse.

Rudski, J. (2003). What does a 'superstitious' person believe? Impressions of participants. *The Journal of General Psychology, 130*(4), 431-445.

Sanborn, A.N., Griffiths, T.L., & Navarro, D.J. (2010). Rational approximations to rational models: Alternative algorithms for category learning. *Psychological Review, 117*(4), 1144-1167. DOI: 10.1037/a0020511.

Scheuplein, R.J. (1993). Uncertainty and the 'flavors' of risk. *EPA Journal, 19*(1).

Schmidt, U., Fritz, K.W., Kasperczyk, W. & Tscherne, H. (1995). Successful resuscitation of a child with severe hypothermia after cardiac arrest of 88 minutes. *Prehospital and Disaster Medicine, 10*(01), 60-62. DOI: 10.1017/S1049023X00041686.

Scotland, J. (2012). Exploring the philosophical underpinnings of research: Relating ontology and epistemology to the methodology and methods of the scientific, interpretive, and critical research paradigms. *English Language Teaching*, *5*(9), 9. DOI: 10.5539/elt.v5n9p9.

Shah, A.K. & Oppenheimer, D.M. (2008). Heuristics made easy: An effort-reduction framework. *Psychological Bulletin, 134*(2), 207-222. DOI: 10.1037/0033-2909.134.2.207.

Shutler, P.M.E. (2009). The problem of the pyramid or Egyptian mathematics from a postmodern perspective. *International Journal of Mathematical Education in Science and Technology, 40*(3), 341-352. DOI: 10.1080/00207390802641692.

Simon, H.A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics, 69*(1), 99-118.

Simon, H.A. (1959). Theories of decision-making in economics and behavioral science. *The American Economic Review, 49*(3), 253-283.

Sommers, S.R. & Kassin, S.M. (2001). On the many impacts of inadmissible testimony: Selective compliance, need for cognition, and the overcorrection bias. *Personality and Social Psychology Bulletin, 27*(10), 1368-1377.

Stanley, J. (2016). Beyond lying: Donald Trump's authoritarian reality. *The New York Times,* http:// www.nytimes.com, 4 November 2016.

Stiegler, M.P. & Gaba, D.M. (2015). Decision-making and cognitive strategies. *Simulation in Healthcare, 10*(3). DOI: 10.1097/SIH.0000000000000093.

Tallentire, V.R., Smith, S.E., Skinner, J. & Cameron, H.S. (2011). Understanding the behaviour of newly qualified doctors in acute care contexts. *Medical Education, 45*(10), 995-1005. DOI: 10.1111/j.1365-2923.2011.04024.x.

Tanford, S., Penrod, S. & Collins, R. (1985). Decision making in joined criminal trials: The influence of charge similarity, evidence similarity, and limiting instructions. *Law and Human Behavior, 9*(4), 319-337.

Tversky, A. & Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. *Science, 185*, 1124-1130.

Tversky, A. & Kahneman, D. (2002), in Eldar Shafir (Ed.)., *Preference, belief, and similarity: Selected writings: Amos Tversky.* Massachusetts, MA: Massachusetts Institute of Technology.

University of Cambridge (2012). Surprising solution to fly eye mystery. *Research,* http://www.cam.ac.uk, 11 October 2012.

Usborne, D. (2016). Donald Trump reveals his fear of humiliation and addiction to media attention in newly released tapes. *The Independent,* http://www.independent.co.uk, 27 October 2016.

Vogel, L.K. & Kebbell, M.R. (2011). Judgement and decision-making in the controversial Dr Haneef counter-terrorism operation: A simulation approach. Psychiatry, *Psychology and Law, 18*(4), 612-625. DOI: 10.1080/13218719.2010.543401.

Walton, D. (2010). Why fallacies appear to be better arguments than they are. Informal *Logic, 30*(2), 159-184.

White, N. (2012). Understanding the role of non-technical skills in patient safety. *Nursing Standard, 26*(26), 43-48.

Wikipedia (2017). Star Trek. *Wikipedia,* http://en.wikipedia.org, Retrieved 1 January 2017.

Wikipedia (2017). Stargate. *Wikipedia,* http://en.wikipedia.org, Retrieved 1 January 2017.

Wikipedia (2018). Appeal to consequences. *Wikipedia,* http://en.wikipedia.org, Retrieved 2 January 2018.

Wilson, E.O. (2012). Altruism (Chapter 7). (In) *On human nature,* 155-175. Toronto, CA: Bantam Books.

Wright, E. (2004). *Generation kill: Devil Dogs, Iceman, Captain America and the new face of American war.* New York, NY: G.P. Putnam's Sons.

Zhu, M., Billeter, D. & Inman, J.J. (2010). Can pictures in advertisements curb consumption? *Advances in Consumer Research, 37,* 50-51.

Appendix 1 – Royal Commission Terms of Reference, Activities and Outputs

The Royal Commission into Institutional Responses to Child Sexual Abuse was established by the appointment of a chair and five other commissioners on 11 January 2013. The commissioners comprised two judges, a former police commissioner, a productivity commissioner, a consultant psychiatrist, and a former state senator. The decision to establish the commission was made by the Australian federal cabinet on 12 November 2012, and then Prime Minister Julia Gillard described the "heartbreaking" character of child sexual abuse (Prime Minister, 2012; Cullen, 2012):

> "Any instance of child abuse is a vile and evil thing. Australians know that, and Australians know from the revelations that they've read in recent weeks that too many children have suffered child abuse. They have also seen other adults let them down. They've not only had their trust betrayed by the abuser, but other adults that could have acted to assist them have failed to do so. There have been revelations of child abusers being moved from place to place rather than the nature of their abuse and their crimes being dealt with. There have been too many revelations of adults who have averted their eyes from this evil. I believe in these circumstances that it's appropriate for there to be a national response through a Royal Commission. This I hope will help the healing, but I specifically hope that its recommendations will help us ensure that this never ever happens again."

The decision to establish a commission was widely supported, and the Prime Minister's announcement was followed by a period of consultation to determine the appropriate terms of reference for the commission. Initially funding approved for the commission was A$434.1 million (Biddington, 2017). In 2014, the commission sought a two year extension, which was estimated to cost A$104 million (Royal Commission, 2014a, pp.11 & 203).

Terms of reference and program of inquiry

The letters patent formally establishing the commission were 'witnessed' by the Governor General of Australia on 11 January 2013 (Letters Patent, per Bryce). The initial expectation was that the commission would produce its final report by 31 December 2015, but this was subsequently extended to 15 December 2017 (Letters Patent, per Cosgrove). The commission was given a broad brief to investigate child sexual abuse in the context of institutions, which does not include the institution of the family. The focus of the commission was particularly on 'systemic issues' arising in the 'institutional context', which extended to child sexual abuse perpetrated by an individual official of an institution where that

abuse had been "created, facilitated, increased, or in any way contributed to" by the institution or its activities.

In volume 1 of its interim report (Royal Commission, 2014a, p.1) the Royal Commission notes the broader brief given to it compared to earlier enquiries about child abuse, and describes its three-pronged approach involving private sessions with survivors of abuse, public hearings where evidence is heard, and research and policy activities – "an extensive research program... [focusing on] prevention, identification, response and justice for victims." 'Survivors' and 'victims' are distinguished as follows (p.31):

> "... we use the term ‘victim’ when referring to the person at the time when the abuse occurred. We generally use the term ‘survivor’ when discussing a person’s experiences after the abuse. This includes when they share their story, access support services or seek redress. Where the context is unclear, or the discussion is general, we will usually use the term ‘victim’."

The scale of the commission's undertaking is indicated by the number of staff and contracted employed, 250 full-time-equivalents mainly based in Sydney (Royal Commission, 2014a, p.2). As at 31 May 2014, the commission had conducted 1,677 private sessions, and by 30 June 2014 there had been 13 public hearings in capital cities across Australia (pp.3-4). In the same time frame, 21 research projects were completed and over 30 more were in progress or under consideration (p.5).

In addition to the "three pillars" of its work – private sessions, hearings and research – the commission had the power to refer suspected perpetrators to appropriate authorities, and could investigate compensation and other remedies for survivors (Royal Commission, 2014a, pp.29-30). Based on the definition in Price-Robertson (2012, as cited in Royal Commission), the commission adopted the following definition of child sexual abuse (p.31):

> "Any act which exposes a child to, or involves a child in, sexual processes beyond his or her understanding or contrary to accepted community standards. Sexually abusive behaviours can include the fondling of genitals, masturbation, oral sex, vaginal or anal penetration by a penis, finger or any other object, fondling of breasts, voyeurism, exhibitionism, and exposing the child to or involving the child in pornography. It includes child grooming, which refers to actions deliberately undertaken with the aim of befriending and establishing an emotional connection with a child, to lower the child’s inhibitions in preparation for sexual activity with the child."

Many of the witness transcripts considered in this thesis describe the ambiguity that may attend adult behaviours towards children, particularly grooming behaviours, and the common difficulty of distinguishing between innocent acts of

kindness and predatory acts. Acknowledging this difficulty, the commission notes that perpetrators of child sexual abuse "manipulate people, processes and situations to create opportunities for abuse" which may be more readily available when institutions have not established a culture that prioritises child safety (Royal Commission, 2014a, p.8).

The hearings held by the commission were organised into thematic case studies, generally focusing on a particular institution or individual. In comments at the start of the second case study, Justice McClellan the commission chair noted that the role of the public hearings was intended to be rather different to the private sessions being conducted concurrently:

> "Public hearings have a different purpose to Private Sessions. Their primary role in this Royal Commission is to provide an opportunity for the development of case studies in which there can be an in depth public examination of the response of particular institutions to the sexual abuse of children within the care of that institution. In addition they enable us to publicly tell the story of some individuals who were abused so that we can explore the circumstances in which abuse may have occurred and the consequences for individuals who have been abused."

Case study 2: Jonathan Lord

Mr Lord was an employee of YMCA's Caringbah branch, providing recreation and child care services in Caringbah, just south of Sydney CBD. These services included five "Before and After School Care" centres located in five schools in the region.

For about two months in 2009, before joining YMCA Caringbah, Mr Lord was a "camp counsellor" at a YMCA camp in Virginia, U.S., a role that ended when he was found alone in a cabin with a boy in suspicious circumstances. A month after being dismissed from the camp in Virginia, Mr Lord applied for a casual position with YMCA Caringbah working with children. He claimed falsely that he come home early from the Virginia camp "because of a personal family matter." YMCA Caringbah did not check his story.

Mr Lord subsequently became a permanent employee at YMCA Caringbah, as "Childcare Coordinator" at one of the school-based Before and After School Care centres. He also worked at YMCA's creche for a morning or two each week.

Although several witnesses identified concerns about some of Mr Lord's behaviours towards children at YMCA, those behaviours were generally capable of an innocent explanation. However, with hindsight it was clear that Mr Lord was highly manipulative and was engaged actively in grooming both his young charges and their parents. His offending included serious molestation of more

than a dozen children during excursions and other events operated by YMCA. A summary of the specific charges can be found at paragraphs 24 to 50 of the opening of Senior Counsel assisting the commission (Furness, 21 October 2013). The outcome, after negotiations to finalise the charges to which Mr Lord would plead guilty, was a prison sentence of 10 years (6 years non-parole) for 13 offences committed against 12 children.

Senior Counsel assisting the commission described Mr Lord's popularity under the heading "Grooming" (Furness, 21 October 2013):

> "72. The Royal Commission will hear evidence that during August 2009 until September 2011, Jonathan Lord was highly regarded by his colleagues, the parents of the children at the various centres and, apparently by the children themselves. Jonathan Lord was described by his colleagues as a 'great person to work with' and 'the kids loved him. He was really positive and popular. It was more than a job to him, and he put in a lot of effort.' Jonathan Lord would 'always make a big effort to make the afternoon tea exciting. It was always extravagant and he would buy or bake cakes for the kids.'"

Much of the evidence for this hearing was concerned with the adequacy of YMCA's policies, the aftermath of the disclosures, and how YMCA dealt with the parents of the abused children. YMCA's actions included requiring its staff to sign confidentiality agreements purportedly on "Police instructions" but this was denied by NSW Police. The confidentiality agreements included a provision that if an employee made an unauthorised disclosure of information about the investigation, he or she "may be subject to disciplinary action, including termination."

Note: This summary is drawn from the opening address of Senior Counsel assisting the commission, and from witness transcripts.

Final Report of the Royal Commission

The Royal Commission's final report comprising 17 volumes was presented to the Governor-General of Australia on 15 December 2017. The report catalogues the evidence heard by the Royal Commission from 1,302 witnesses during 57 public hearings spread over 444 hearing days.

The Royal Commission received 1,388 submissions to issues paper and consultation papers, and reviewed 1.2 million documents. In addition, the Royal Commission held over 6,960 private sessions with survivors of child sexual abuse (Royal Commission, 2017, p.4).

In its final report the Royal Commission makes a wide range of recommendations, focused especially on (in the commission's words):

- preventing abuse or, at the very least, identifying it as early as possible
- improving the way perpetrators are investigated, prosecuted and sentenced
- improving survivors' access to justice and ongoing support.

Further information about the Royal Commission, including the whole of its final report, can be found online at:

https://www.childabuseroyalcommission.gov.au

Appendix 2 – Table of Witness Transcripts

This is a list of all witnesses who gave evidence in the Royal Commission's case study number 2 that related to Jonathan Lord. The last column indicates the witnesses whose evidence has been analysed in this thesis.

Identifier	Name	Category	D.A.
AN	"AN"	Victim's Parent	Yes
AX	"AX"	Victim's Parent	Yes
AS	"AS"	Victim's Parent	Yes
AU	"AU"	Victim's Parent	Yes
AZ	"AZ"	Victim's Parent	Yes
AT	"AT"	Victim's Parent	Yes
AW	"AW"	Parent of enrolled child	No
AD	Alicia Dellaca	Current or former YMCA employee	Yes
DO	Danielle Ockwell	Current or former YMCA employee	Yes
MB	Michelle Bates	Current or former YMCA employee	No
SO	Sheree Ockwell	Current or former YMCA employee	No
CB	Carine Beer	Current or former YMCA employee	Yes
CS	Chloe Starr	Current or former YMCA employee	Yes
ET	Erin Turner	Current or former YMCA employee	No
CC	Catharine Clements	Current or former YMCA employee	No
AM	Ann Mary Nolan	Current or former YMCA employee	No
JB	Jacqui Barnat	Current or former YMCA employee	No
LW	Liam Whitley	Current or former YMCA employee	No
PH	Phillip Hare	Current or former YMCA employee	No