ON THE ORIGIN OF COGNITIVE DISSONANCE

A thesis presented in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy in Psychology at
Massey University, Manawatu, New Zealand

Peter Clemerson

2016

For Navina, in gratitude
for her loving support and confidence

Abstract

Since Festinger (1957) published his monograph describing Cognitive Dissonance (CD), a wealth of research has led to a widely accepted understanding of its nature and effect on human behaviour. Holding two conflicting cognitions in mind simultaneously results in an aversive feeling which is alleviated when action is taken to resolve the conflict. In many ways, it acts like an aversive emotion though it is not widely regarded as one. Emotions are accepted as products of humanity's evolution over many millions of years. Despite its occasional designation as 'adaptive' and even rarer comparison to emotions, research establishing CD as a product of Darwinian selection pressures is still lacking. In this research, three main hypotheses based upon Darwinian considerations were developed which predicted differential sensitivity to contradictions according to semantic categories. Further minor predictions were made to test relative sensitivities within categories. Verification of the main hypotheses provides clear evidence for a Darwinian explanation for the existence of CD.

Two studies to test the hypotheses were based upon the contradiction paradigm. Over four hundred adults from university premises and shopping mall food courts volunteered to read short stories on a laptop computer screen. Half the stories contained a line contradicting an earlier one. These contradictory stories embraced nine semantic categories of contradiction. As participants successively pressed the space bar to display each story line, their response times were recorded. Predictions specified participants' relative sensitivities to each category of contradiction, indexed by the differential response times of contradictory lines.

Analyses of the response times of all participants combined and male and female participants separately produced confirmations of the main predictions, often with large effect sizes, and a mixture of confirmations and disconfirmations of the minor ones. Some interesting differences between male and female participants emerged for which tentative explanations are offered. The data were further analysed to verify the assumptions upon which the contradiction paradigm is based. In summary, it is claimed that this research establishes CD as a product of Darwinian evolution.

Acknowledgements

I am happy to express my gratitude to the following people:

Dr. Stephen Hill, Massey University Manawatu campus, primary supervisor, for careful and helpful guidance throughout the lengthy period of the research.

Dr. Steven Humphreys, Massey University Wellington campus, initial secondary supervisor, for ensuring that issues of relevance were properly explored.

Dr. Andrew Towers, Massey University Manawatu campus, final secondary supervisor, for ensuring that research disciplines were adhered to and for suggesting stimulating references.

Drs. Alan Winton and Rachel Pond and Ms. Gillian Craven, Massey University Manawatu campus, for allowing me to attend their lectures to recruit their students as participants.

Mr. Malcolm Loudon, Massey University Manawatu campus, for writing data collection programs.

Mr. Harvey Jones, Massey University Manawatu campus, for technical assistance in the provision of SPSS software.

Ms. Tracey Geoghegan, Marketing Coordinator, Johnsonville Mall, DNZ Property Fund Ltd. for permission to approach members of the public for data collection purposes.

Ms. Simone Hadley, Shopping Mall Manager, Precinct Ltd. Wellington, for permission to approach members of the public for data collection purposes.

Ms. Navina Clemerson, for challenging discussions and proof-reading assistance.

Mr. Rowan Clemerson, for computer technical assistance.

Ms. Tamsyn Clemerson-Philips, for participant recruitment.

Mr. Daniel Clemerson-Philips, for participant recruitment.

All the participants, numbering in excess of a thousand, who kindly provided their time to this research without reward and often with enthusiasm.

Contents

Table of Contents

List of Appendices

List of Tables

List of Figures

List of Abbreviations

ACC	Anterior Cingulate Cortex
BAS	Behavioural Approach System
BIS	Behavioural Inhibition System
BP	Before the Present
CAT	Computer Aided Tomography
CD	Cognitive Dissonance
DLFC	Dorsolateral Frontal Cortex
DNA	Deoxyribonucleic Acid
EEA	Environment of Evolutionary Adaptation
EEG	Electro Encephalograph
fMRI	Functional Magnetic Resonance Imaging
kya	Thousands of years ago
LFC	Left Frontal Cortex
MMT	Meaning Management Model
MRI	Magnetic Resonance Imaging
mtDNA	Mitochondrial Deoxyribonucleic Acid
mya	Millions of years ago
PET	Positron Emission Tomography
RFC	Right Frontal Cortex
RT	Response or Read Time
SRR	Survival, Reproduction and Reputation related
SSSM	Standard Social Science Model of the mind
TMT	Terror Management Theory
WM	Working Memory

Chapter 1 - Introduction

The woman walked through the station concourse towards platform 1 for the Johnsonville train, as she had done many times. A railway staff member intercepted her and asked ”Johnsonville?”. “Yes”, she replied. He waved her towards platform 3. She walked some way down platform 3, boarded the train and sat down. The illuminated sign gave no destination. The carriage was empty. She sat there alone. Feeling progressively uneasy, she stayed a while. Despite the message from the railway member of staff, no-one joined her. Eventually, she got out and walked back to the platform 3 entrance where the overhead sign now showed ‘Johnsonville’. She walked down the platform once again, saw a friend in the train, boarded a second time and sat down beside the friend, relieved. “That was uncomfortable” she said.[1]

*

“The late Saudi King has been lionised by politicians around the world. … Christine Lagarde, the head of the International Monetary Fund, even hailed the monarch as ‘A strong advocate for women’. That last eulogy ought to furrow brows. After all, when it comes to gender rights, Saudi Arabia’s absolute monarchy is one of the most heavily criticised regimes in the world.” (Tharoor, 2015).

*

Mary knew that her friend June was faddish about food and was always going on about the latest diets and ingredients that should or should not be eaten. Now it was supplements and fish. She was sure they were good for you and was advocating them to all her friends. But Mary had recently heard from another friend that supplements and special diets were useless. Some specialists were saying they could be harmful. The advice was so contradictory. She was confused and didn’t know what to do (Nagler, 2014; University of Colorado Cancer Centre, 2015; Whoriskey, 2015).

*

The passages above illustrate a state of mind first formally described in 1957. Festinger (1957) introduced the concept of cognitive dissonance (CD), drawing attention to the psychological discomfort caused when two thoughts or beliefs, more

[1] This incident occurred in February 2015, the second boarding and remark observed by the author. The earlier part was recounted by the woman to the author at his request.

technically, cognitions or propositions[2], are held in mind simultaneously despite being mutually contradictory. He proposed that this discomfort is the motivation for some form of activity which might resolve the contradiction and eliminate or lessen the discomfort. Festinger (1957) indicates that the discrepant propositions, or "elements of knowledge" (Harmon-Jones & Mills, 1999, p. 3), leading to cognitive dissonance, are expressible in short sentences in direct opposition to each other, such as 'This drink is pleasant' and 'This drink is unpleasant'.

A rich literature was developed during the decades following Festinger's (1957) lead, covering the different circumstances in which CD may be experienced and the means to mitigate it (Cohen, 1962; Festinger & Carlsmith, 1959; Sherman & Gorkin, 1980; Zanna & Cooper, 1974; Zana, Higgins, & Taves, 1976). This research illustrated the role of CD in common human actions such as changing opinions, beliefs, attitudes, and everyday behaviour but also in persistent adherence to previous beliefs, despite the availability of opposing information (Simon, Greenberg, & Brehm, 1995). In parallel with this development, a separate literature was developed which explored people's responses to other and apparently unconnected forms of discrepant information, such as fears of mortality in a life affirming species. A synthesis of these disparate lines of research is now emerging and it is within this developing concordance that CD will be examined.

Although Festinger (1957) notes that dissonance is "uncomfortable" (p. 3), he did not describe in detail the phenomenology of CD. However, its aversive subjective feeling is well captured in the phrase "elevated feelings of discomfort, i.e., feeling uncomfortable, uneasy, bothered" (Devine, Tauer, Barron, Elliott, & Vance, 1999, p. 309). Harmon-Jones (2000) characterises the CD feeling of discomfort as an "adaptive" response (p. 189) although no indication is given whether an evolutionary or more general sense is meant nor, if the former, how it may have evolved. However, he is forthright in his view that this aversive feeling is similar to a negative emotional state (p. 185) and implies thereby that its existence, like that of emotions, can be regarded as having an evolutionary explanation. In support of Harmon-Jones (2000), CD will be viewed as an aversive emotion because it can first trigger and then sustain particular

[2] The term proposition is used as in logic: a statement in which something is affirmed or denied, having a truth value of either true or false (Mordecai, 2012).

action until it is alleviated. Conventional emotions (Ekman, 1993; Keltner & Buswell, 1996; Plutchet, 2003) and bodily feelings, such as hunger and pain, are generally explained as products of natural selection (Lazarus, 1991; Nesse, 1990, Rolls, 2005). They are regarded as conscious or barely conscious experiences that act as triggers to decision making and/or sustainers of actions which promote the preservation of life and bodily integrity (LeDoux, 2000; Mobbs, Hagan, Dalgleish, Silston, & Prévost, 2015), the acquisition and retention of life-preserving possessions (State, 2004), mate or mates (Bartels & Zeki, 2000; Fisher, Aron, Mashek, Haifang, & Brown, 2002; Gonzaga, Keltner, Turner, Campos, & Altemus, 2006) and a reputation for prosociality (Bowles & Gintis, 2003; Trivers, 1971; Wedekind & Milinski, 2000). These survival, reproduction, and reputation (SRR) related feelings, emotions and decisions aid the passing of genes from one generation to the next. These feelings and emotions can be divided into two categories: positive/pleasant and negative/aversive, with a majority in the second category according to Averil (1980). A simple examination reveals that the aversive ones either prevent, or motivate the cessation of, actions deleterious to wellbeing, or motivate actions promoting wellbeing (Aunger & Curtis, 2013; Nesse, 1990). Action ceases once alleviation has been obtained. Given the aversive quality of CD and the role of aversive SRR-related emotions in triggering behaviour, several questions arise: (1) in what other ways does CD resemble conventional emotions, (2) can CD be considered a product of natural selection, and (3) can it perform the same function of promoting gene transmission? The current research endeavours to provide answers to these questions.

An accurate assessment of the environment is usually a necessary pre-requisite to wellbeing-promoting actions (Plutchet, 2003; Rolls 2005). We are frequently exposed to contradictory information about it, (e.g., particular foods are reported by some to be beneficial and by others to be harmful) which is a problem when determining which of two, or more, actions are to be pursued. It is hypothesised that natural selection has equipped humanity with a separate feeling, Cognitive Dissonance, that deflects people from a potentially damaging path, namely, failing to ascertain which of two contradictory SRR related cognitions is true and which false. Harmon-Jones (2000) makes the plausible assumption that action based upon false information will on average diminish wellbeing. Referring to assessments of environmental reality, Festinger (1957) hints at the role of natural selection when he notes that “the reality that

impinges on a person will exert pressures in the direction of bringing the appropriate cognitive elements into correspondence with that reality" (p. 11), although he does not say how such pressures might originate. While aversive emotions trigger actions to minimise harm in adverse environmental circumstances, CD's function is hypothesised also to minimise harm by influencing people's decision-making away from false views of the environment and consequent harmful actions by motivating action to obtain information that more closely reflects reality.

Testing the Adaptive Nature of Cognitive Dissonance

The hypothesis that CD is an adaptive product of natural selection can be tested by assessing the degree to which CD is felt when contradictions in SRR-related and non-SRR-related domains are presented and compared. In order to aid the passing of the relevant genes from one generation to the next, it is proposed that the resolution of some categories of contradictions is more important than the resolution of others. For example, resolving contradictions of SRR relevance, such as conflicting accounts of the intentions of a weapon-carrying neighbour, would be much more important than resolving contradictions of no adaptive significance, such as those relating to car parking infringements. It is hypothesised that the normal selection pressures associated with natural selection have led to the evolution of a greater sensitivity to contradictions of SRR relevance in order to increase the probability of these, relative to others, being noticed and resolved. Reproductive benefits would follow from the extra information obtained as part of a dissonance mitigation process. The discovery of this category-related greater sensitivity has therefore been made the goal of some experiments. Its existence would provide evidence supporting the claim that CD is a product of natural selection. As most contradictions are encountered verbally, its evolution is likely tied to the evolution of humanity's linguistic abilities. Despite the publication of some research that sought to demonstrate CD in non-human animals (Egan, Santos & Bloom, 2007, but see Chen & Risen, 2009 and Egan, Bloom, & Santos, 2010), implying a natural selection-related origin, no research has been conducted which explicitly investigates or explores how the pressures of Darwinian natural selection may have operated to produce CD as an important and motivating human attribute. Our understanding of CD remains incomplete until an account of its origin can be provided. The purpose of the research reported here was to supply one.

Chapter 2 - Theoretical Background

Overview of this chapter

This chapter is divided into five sections. The first reviews the nature of Cognitive Dissonance with an emphasis on those psychological characteristics of relevance to the theme of this thesis, namely, that evidence can be produced which shows CD to be a product of natural selection. The second section outlines a modular view of the mind and brain. The third section returns to CD drawing upon the modularity view of the mind/brain to identify a specific attribute of CD that, if experimentally found, would provide evidence for the role of natural selection in the evolution of CD. The fourth section takes a wider view of CD, proposing that it is one of several psychological phenomena that are all manifestations of a single underlying mechanism that evolved to detect cognitive discrepancies. The implication is that evidence that throws light on the evolution of one of them, for example, CD, throws light on them all. The fifth section develops an experimental rationale and makes specific testable predictions. The predicted outcomes of these tests are proposed as evidence for the evolution of CD by natural selection. The implications of the results for the other related phenomena are discussed in the General Discussion at the end of the thesis.

The Theory of Cognitive Dissonance

Festinger's Definition of Dissonance.

Festinger (1957) defined relationships between pairs of elements of knowledge, or cognitions, as dissonant, consonant or irrelevant. Dissonance refers to the "nonfitting relations between cognitions" (p. 3), a semantic clash. The cognitions are inconsistent or contradictory either explicitly or implicitly. He preferred the term 'consonance' to 'consistency' and proposed that "if either one [cognition] follows from the other, then the relation between them is consonant" (p. 15). If neither "follows from the other, the relation between them is irrelevant." (p. 15). He provided a more formal definition of dissonance, "These two elements are in a dissonant relation if, considering these two alone, the obverse[3]of one element would follow from the other" (p. 13).

Festinger's (1957) concept of cognition was very broad. "By the term cognition … I mean any knowledge, opinion or belief about the environment, about oneself or about ones behavior" (p. 3). It included what one is, does, feels, wants, desires. Also included were environmental facts about the world, where one lives, what is where, and what leads to what. He laid some stress on what things are satisfying or painful or important or inconsequential. Included also in his list are beliefs, values and attitudes. He coined the term "knowledges" for all the above but as this term has not been generally adopted, the term 'cognition' will be used. He emphasised reality as "the single most important determinant of the content of these elements". Of particular relevance to the current research is his statment that "It would be unlikely that an organism could live and survive if the elements of cognition were not to a large extent a veridical map of reality" (p. 10).

According to Festinger (1957), the magnitude of dissonance between one (set of) cognition(s) and another depends on the number and importance of the cognitions that are consonant and dissonant with the first. He formalised the magnitude as proportional to the number of dissonant cognitions divided by the sum of the consonant

[3] It is here assumed that Festinger is using the word 'obverse' in the sense that it was used in logic: "The inference by which we deny the opposite of an affirmation" (Collin's New English Dictionary, 1956). The Concise English Dictionary (1950) provides "counterpart of a fact or truth".

and dissonant ones. This ratio suggests a means by which the dissonance could be reduced: recruiting more consonant cognitions so that the denominator in the ratio becomes larger or reducing or invalidating the dissonant ones so that the numerator becomes smaller at a greater rate than the denominator does.

Festinger (1957) devotes little space to the phenomenology of dissonance but writes "[T]he existence of dissonance, being psychologically uncomfortable, will motivate the person to try to reduce dissonance and achieve consonance" (p. 3). He noted that the previous sentence can be generalised. "In place of 'dissonance', one can substitute other notions similar in nature, such as 'hunger, 'frustration', or 'disequilibrium'" (p. 3). These examples refer to the phenomenological states experienced by a hungry or frustrated person. Although he previously defined dissonance in terms of a semantic opposition between two cognitions, he here appears to include the phenomenology of 'dissonance' within its definition. A more extensive description appears in Devine, Tauer, Barron, Elliott, and Vance (1999, p. 309) where they write dissonance "should be revealed as elevated feelings of discomfort, e.g., uncomfortable, uneasy, bothered"[4]. They provide a diagram, reproduced below, which illustrates their view of Festinger's (1957) process model of dissonance.

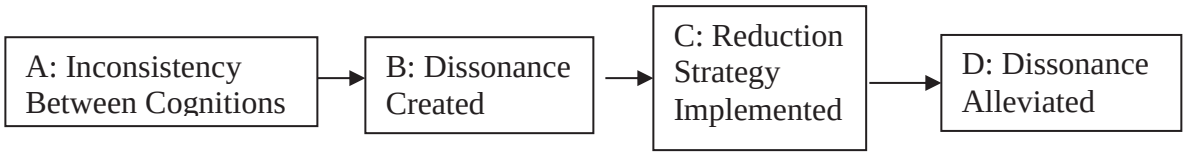


Figure 1. Schematic Diagram of Festinger's Theory of Cognitive Dissonance from Devine, Tauer, Barron, Elliott, and Vance (1999) p. 298.

Devine et al. (1999) explain the diagram. "In Festinger's model, inconsistency between cognitions (A) leads to dissonance (B), an uncomfortable psychological tension and arousal state that the person will be motivated to reduce. According to the model, the motivational properties of dissonance will lead to a dissonance reduction strategy" (p. 297). The first sentence *defines* dissonance in phenomenological terms and its unpleasantness is the motivating force to reduce it. When discussing the degree to

[4] Some researchers mention 'guilt' and 'regret' in connection with the cognitive discrepancies that trigger these emotions.

which two cognitions are discrepant, Brehm and Cohen (1962) frequently use the phrase "arousal of dissonance", strongly implying the term 'dissonance' refers to a phenomenological state. They report an experiment in which "[T]he magnitude of dissonance was measured in terms of increased liking for the vegetable" (p. 199) reinforcing the acceptance within the academic community that 'dissonance' refers to a phenomenological state as well as to discrepant cognitions. Given Festinger's use of the phrase "being psychologically uncomfortable", his comparison of dissonance with hunger and frustration, and the research community's expansion of these phenomenological allusions, the term 'dissonance' appears to refer to either or both discrepant cognitions and to the associated phenomenology. When reporting the older literature's use of the term dissonance, if a distinction between the phenomenological experience and the cognitions causing it is required, an explicit statement to this effect is made below. Contra Festinger (1957), Harmon-Jones, Harmon-Jones, Serra, and Gable (2011) differentiate between these two facets by naming the phenomenological experience as "dissonance", the causal cognitive conflict as "cognitive discrepancy", the reduction or elimination of dissonance as "dissonance reduction" via a change of cognition(s) they term "discrepancy reduction" (pp. 395-396). This is useful and these terms are adopted later. However, it is a relatively new proposal and does not therefore reflect common practice in the older CD literature.

Sources of Dissonance.

Festinger (1957) drew attention to several different circumstances in which dissonance may be generated. So that the variety of such circumstances may be appreciated, four examples drawn from his work are given below. In some cases, the conflict may be explicit but in others it is not.

1. A logical inconsistency

Festinger (1957) gives the example of a man simultaneously believing that humanity will reach the moon but also believing that humanity will never build a device that could leave the earth's atmosphere. "These two cognitions are dissonant with one another" (p. 14).

2. A conflict with cultural mores

If a person picks up a bone with their hands at a formal dinner and gnaws at it, it can be expected that a western observer would regard this behaviour as conflicting with, and therefore dissonant with, formal dining etiquette. It is worth noting that the observer's discomfort would be felt as soon as the contravening act is observed. An unexpressed expectation has been violated. There does not need to be a consciously expressed thought for the discomfort to be felt. Verbalisation is not necessary although the nature of the conflict can be expressed verbally very simply as a syllogism.

In circumstances Y, nobody does X (the rule)
In circumstances Y, A does X
A contravenes the rule.

3. Inconsistency between one cognition and a more encompassing one

A voter might normally vote for a particular party but in a particular election, s/he might choose to vote for a candidate of a rival party because of the high quality of this candidate in comparison with that of the candidate of his/her 'normal' party.

4. Conflict between current and past experience

Festinger (1957) gave the improbable example of someone who was standing in the rain yet observed that he is not getting wet (Festinger 1957 p. 14). Aronson (1999, pp. 109-110) contested this example. A more realistic example might be the case of someone who is normally polite and kind suddenly behaving very rudely or unkindly towards someone, dissonance arising in the mind of a third party observer.

Not all cases of dissonance generation may be so direct. In tests of textual incoherence, Huitema et al. (1993) used examples such as 'Dick had a week's vacation due and he wanted to go to a place where he could swim and sunbathe' followed by 'He went to his local travel agent and asked for a plane ticket to Alaska.' They found that

such implicit contradictions were noticed even when separated by other text. Clemerson (2009) also found that implicit contradictions similar to the Huitema et al. (1993) example were noticed by experimental participants but less frequently than if they were expressed explicitly.

Aronson (1969) noted that despite an apparent logical opposition of two cognitive elements, they may not give rise to dissonance. He provided the example of a married couple, one of whom is very clever and the other rather dull. He regarded opinions accurately reflecting these attributes of the two people as an opposition and noted that it might be expected to give rise to dissonance. However, it need not. In such circumstances, he posited that there are underlying cognitions that prevent the dissonance arising. In this case, it is widely appreciated that people base their choice of marriage partner upon a very wide range of criteria. While there may be differences in one attribute, similarities in others prevent the generation of dissonance. He suggested therefore that Festinger's comment that dissonance arises if "considering these two [elements] alone, the obverse of one element follows from the other" (Festinger, 1957, pp. 260-261) needs the qualification that dissonance will arise only if there do not exist underlying cognitions of the type he identified. The generation of such an underlying cognition has been identified as a dissonance reduction technique by other researchers (see below).

In a discussion of Festinger's (1957) definition of dissonance, Brehm and Cohen (1962) proposed that many of the variables in such cognitions are continuous rather than discrete, thereby weakening the application of Festinger's definition. The opposition of the pair "X owns lots of shares" and "X is not rich" is dependent upon personal interpretations of the words "lots" and "rich". They proposed instead the amended definition of dissonance; "X follows from any variation from Y in the direction of being obverse" (p. 16). Brehm and Cohen's (1962) definition allows a graded phenomenological experience to follow a single contradiction with the intensity of that experience being related to the degree to which two propositions are interpreted as contradicting each other, provided the relevant terms are capable of being regarded as continuous. While gradations in the intensity of the phenomenology of dissonance were proposed by Festinger (1957), it was not an aspect of CD that was pursued by the early

experimenters. As an issue of relevance to this research, the intensity of dissonance is taken up again in later sections.

Experimental Paradigms.

Because Festinger (1957) included behaviours, perceptions, attitudes, beliefs, and feelings as well as the consciously held propositions of the type illustrated previously in his definition of CD, a wide range of experiments were conducted during the following decades to investigate the nature of CD, the limitations of the theory and its relationship to other theories of motivation. Some of these experiments are briefly described later in order to provide a portrait of CD relevant to the current research. So that these experiments may be better understood, a brief description of the five experimental methods employed by the majority of CD experimenters is provided below.

Belief Disconfirmation.

Dissonance occurs when people are exposed to information which contradicts their beliefs. If beliefs are not changed, dissonance can be reduced either by seeking the support of others who hold the same belief or by converting others to that same belief, thereby obtaining their support, i.e., consonant cognitions. Festinger, Riecken and Schachter (1956) observed precisely the latter when a religious prediction failed to materialise creating cognitive dissonance between the prediction and reality. After generating a new underlying explanatory cognition, adherents embarked upon a new campaign to obtain recruits.

Effort Justification.

Aronson and Mills (1959) report an experiment in which women were ostensibly inducted into a discussion group. There were three conditions; an intensely embarrassing induction procedure, a mildly embarrassing one and a control. Although no group existed, participants did listen to a specially recorded dull and boring discussion. On discovering the apparent nature of the group, participants in the intensely

embarrassing condition "experienced dissonance between going through an unpleasant experience and taking part in worthless uninteresting discussions" (p. 180). They assessed the group more favourably than the participants in the other conditions at statistically significant levels. These more favourable assessments justified the embarrassment experienced and diminished the cognitive dissonance generated by the contrast between their expectations of the group and its disappointing nature.

Induced or Forced Compliance.

In a frequently used paradigm, participants are induced to engage in a counter attitudinal behaviour under conditions of high or low choice and their consequent attitude change is measured. On completion, participants in the high choice condition are found to have favourably altered their attitude, the change being interpreted as a means of dissonance reduction. Participants given little or no choice do not alter their views, any negative feelings of the participants being interpreted as a consequence of their coercive situation.

Free Choice.

When a choice is made between two items or options previously closely evaluated, a conflict is presumed to arise between the negative attributes of the chosen item or option and the positive attributes of the rejected one. Typically, experimental participants first evaluate a set of options or items, for example, a set of pictures, these being the 'before' evaluations. They are subsequently required to choose between two items that are either 'very different' or very nearly the 'same'. In the 'same' condition, participants are assumed to experience dissonance due to the conflict indicated above. The dissonance is hypothesised to be dissipated by re-assessing the options and spreading the evaluations so that they are no longer as close as before. These are the 'after' evaluations. Participants in the 'very different' condition do not spread the evaluations and are presumed not to have felt CD. The merits of this method of CD detection are contested by some authors (Chen, 2008; Chen & Risen, 2010) and questioned by others (Coppin, Delplanque, et al. 2010) but it continues to be used by other prominent researchers (Harmon-Jones, Harman-Jones, & Levy, 2015).

Hypocrisy studies.

In hypocrisy studies, participants in a 'hypocrisy' condition advocate a position with positive consequences and are reminded of their own past failures to live up to this position. Participants in other conditions either do not advocate the position or are not reminded of past failures. A dissonance related dependent variable, such as change of behaviour, is measured with the participants in the hypocrisy condition manifesting the greatest change (Fried & Aronson, 1995).

The Psychological Characteristics of Cognitive Dissonance.

In the following section, CD is portrayed via brief descriptions of the experiments that provided evidence for various facets of CD. Psychologically, the most notable is CD's aversive nature. The factors influencing the intensity of the feeling are covered because of the relevance of these to the current thesis. This is followed by a review of the CD reduction strategies used in experiments and everyday life. A model of CD is presented which builds on these discoveries and provides an abstract set of reasons why CD may exist. However, the model does not provide proximate reasons for CD's existence or an explanation how CD may have evolved into existence, the matter of the current research.

The Aversive Nature of Dissonance.

According to Zanna, Higgins, and Taves (1976), no investigators during the early years of cognitive dissonance research "convincingly established the aversive nature of dissonance" (p. 530), as Festinger (1957) had suggested. Zanna et al. (1976), employing the induced compliance technique, reasoned that if the arousal was aversive and the opportunity to misattribute this to various causes was provided, participants who were able to attribute their arousal to an external aversive cause would undergo less of an attitude change than participants unable to do so. They required participants to generate a feeling of discomfort by writing counter-attitude essays advocating a course of action with which they disagreed. Participants took a pill prior to writing their essays.

Participants were informed that as a result of the pill-taking, they would feel “tense” or “pleasant” or there was “no effect” or were provided with “no information”. The authors reasoned that if dissonance was a state of general arousal rather than aversive, then participants in both the “tense” and the “pleasant” conditions would be able to misattribute their arousal to the pill and both undergo similar and less attitude change relative to the “no effect” group. If the dissonance was aversive, the “tense” group would be able to misattribute their discomfort to the pill and undergo little attitude change. Of the four groups, the “tense’ group did indeed undergo the least attitude change. The authors interpreted the “tense” group’s lowest attitude change, according to predictions, as demonstrating that the dissonance-induced state of arousal was aversive.

Elliot and Devine (1994) used a Dissonance Thermometer, consisting of 24 self-report items measured on a Likert scale, to measure the affective states of participants. *Discomfort* measured how uncomfortable, uneasy, and bothered the participants felt. A *Negself* index measured disappointment, annoyance, guilt, self-criticism and a *Positive* index measured how good, happy, optimistic, and friendly they felt. Participants wrote an essay in response to reading a passage advocating a 10% increase in student fees. The two most relevant conditions were defined as (1) a Counter Attitude High Choice Affect-Attitude condition in which participants first agreed (High Choice) to write a counter attitudinal essay, were then tested for, first, their affective state and afterwards their attitude, and then wrote the essay, and (2) a Counter-Attitude High Choice Attitude-Affect condition in which they first agreed to write the counter attitudinal essay, wrote it and were then tested first for attitude and then for affect. The condition (1) group showed very little change in attitude in comparison with the prior assessment and manifested the highest measure of *Discomfort* in anticipation of writing the counter attitudinal essay. Attitude change in condition (2) was high relative to all other conditions, demonstrating the induced compliance effect. These results permit an interpretation that the writing of the essay represented a dissonance reduction strategy which resulted in a reduced level of discomfort. The Dissonance Thermometer included the *Negself* and *Positive* indices. The experimenters note “No significant effects were revealed in ANOVAs using the *NegSelf* and *Positive* indices, suggesting that *Discomfort* was the distinct affective consequence of the dissonance induction” (p. 387).

Motivational Characteristics.

Festinger (1957) hypothesised that "…dissonance, being psychologically uncomfortable, will motivate the person to try to reduce the dissonance and achieve consonance" (p. 3). Many experimenters demonstrated CD as a motivational factor in attitude change (Brehm & Cohen, 1962) while Pallak and Pittman (1972) tested the hypothesis that the general motivational properties of dissonance were similar to those of other non-cognitive drives. Drawing on learning theory literature, they proposed that if dissonance can be considered a motivational construct similar to other drives, experimental participants should perform less well on an unconnected task involving perceptual competition when in a dissonance state. Using the induced compliance method, Pallak and Pittman (1972) required participants to write counter attitudinal essays in high and low choice and justification conditions and found that high choice low justification participants made more errors in a Stroop test than participants in other conditions, supporting their hypothesis that dissonance has motivational properties similar to others explored within the learning theory literature, such as anxiety (Waterman, 1969).

Scher and Cooper (1989) conducted experiments using the induced compliance method with a view to demonstrating that it is not conflicting cognitions per se that induce dissonance. They claimed that the dissonance was caused by a belief in the participants' minds that aversive consequences will follow from the experimental actions. However, Harmon-Jones, Brehm, et al. (1996) also using an induced compliance method, were able to demonstrate that the motivational effects of CD operate even in the absence of aversive consequences.

Festinger (1957, p. 129) proposed that in circumstances where someone held a view despite increasing evidence and arguments favouring a contrary position, and thus experiencing accumulating dissonance, information supporting the contrary position may be sought, increasing dissonance to the point where the person is motivated to abandon the first view. The dissonance is thereby dissipated.

Factors influencing the intensity of dissonance.

Although Festinger (1957) asserted that "if two elements are dissonant with one another, the magnitude of the dissonance will be a function of the importance of the elements." (p. 16), as late as 1999 Devine, Tauer, Barron, Elliot, and Vance (1999) claimed that "… empirical attempts to explicate the role of attitude importance are scant" (p. 312). Devine, Froning, and Elliott, (1995, unpublished study cited in Devine et al., 1999) therefore conducted an experiment to test the view that when attitudes were of high importance, attitude change would be less than when attitudes were of low importance. Using an induced compliance experiment, they found that in high choice, high importance conditions, participants showed high discomfort but did not manifest attitude change. Aronson (1969, 1991) suggested that high dissonance can be expected in cases when the contradiction pairs are self-related and Simon et al. (1995) suggested that inconsistencies central to basic needs such as survival would arouse particularly high levels of dissonance, although they did not test this conjecture. Using the hypocrisy procedure, MacGrath (2011) found that participants who assigned a high importance to animal welfare issues experienced higher levels of dissonance than those with a low importance attitude and manifested this higher level of dissonance in their choice of dissonance reduction method.

In a more complicated experiment, Wakslak (2012) used an induced compliance experiment requiring student participants who had previously indicated their opposition to the introduction of a new examination, to write a counter attitudinal essay under conditions of high and low importance. Her examination introduced high and low level construal (Trope & Lieberman, 2010) as another variable. Construal Level Theory distinguishes between abstract assessments of an object or issue (high level) and concrete assessments (low level). Wakslak (2012) investigated a hypothesised three-way interaction between high and low choice, high and low importance and high and low construal levels. After writing the essay, high level construal participants were asked to answer abstract 'why' questions on health maintenance while low level construal participants were asked to answer questions about how good health may be maintained. Immediately afterwards, in high or low construal mind sets, participants' attitudes to the examination were reassessed. The greatest change in attitude took place among low construal low importance high choice participants while the least change in

attitude took place among high construal, high importance low choice participants. Wakslack interprets the results to indicate that both construal mind set and topic importance affect attitude change. Starzyk, Fabrigar, Soryal, and Fanning (2009) also investigated other factors besides topic importance as predictors of dissonance and found that the salience of the topic importance during the experiment influenced the intensity of dissonance. When importance was not made salient, high choice participants engaged in dissonance reduction techniques in proportion to the importance of the topic. When high importance attitude was made salient, high choice participants changed their attitudes less with increasing importance and showed no evidence of engaging in dissonance reduction techniques. The results of the Starzak et al. (2009) experiment indicate that although topic importance influences the generation of dissonance, its salience is also relevant.

Research has also indicated that other factors can influence the intensity of dissonance. Cialdini, Trost, and Newsom (1995) constructed a Preference for Consistency (PFC) Scale, which has been widely used to explain variance in consistency-based social phenomena (Guadagno & Ciadini, 2010). They demonstrated that participants with a personality disposition inclined to high PFC scores will show greater attitude change, indicating higher dissonance, than participants with low PFC scores, when presented with conflicting cognitions. Hoshino-Browne, Zanna, et al. (2005) demonstrated the relevance of cultural factors such as the individualism of Western cultures in contrast with the communalism of Eastern cultures. The experiments outlined in this section together with others reviewed in Harmon-Jones, Amodio & Harmon-Jones (2009) indicate that although topic importance influences the generation of dissonance, other variables moderate its effect.

Dissonance Reduction Techniques.

As an aversive experience, CD has been found to motivate recourse to one of a number of dissonance reduction techniques. The one adopted has usually been determined by experimental method. Experiments have demonstrated the effectiveness of the following:

1. Changing a prior cognition, e.g., an attitude or belief, such that it becomes less opposed to the newly presented or generated one. This capacity for change of

belief is exploited in the Free Choice and Induced Compliance research paradigms (Elliott & Devine, 1994).

2. Reducing the importance of the counter attitudinal behaviour and/or the issue, a process named trivialization by Simon, Greenberg, and Brehm (1995).
3. Subsequent to dissonance having been experienced when receiving information contradicting prior beliefs, bolstering the latter by generating consonant cognitions via ones own decisions or actions (Sherman & Gorkin, 1980).
4. Self-affirmation by obtaining consonant cognitions of oneself from self or others (Tauer, Devine, & Elliott, 1998, cited in Devine et al. 1999).
5. Misattribution of the discomfort to some other cause, provided a plausible one exists (Zanna & Cooper, 1974).
6. Denial of responsibility for counter attitudinal behaviour (Gosling, Denizeau, & Oberlé, 2006).
7. Restructuring cognitions not directly related to a specific attitude-behaviour discrepancy, for example, generating some background belief or opinion which lowers the discrepancy (Leippe & Eisenstadt, 1999).
8. Selective exposure which consists of seeking information when in a state of dissonance (Cotton, 1985: Erlich, Guttman, Schonbach, & Mills 1957).

Festinger (1957) suggested that behavioural practices are the most resistant to change implying other strategies among the many available will be adopted in preference. Simon et al. (1995) note that people use whichever dissonance reduction strategy is the easiest in laboratory conditions, this usually being the first one offered. However, Macgrath (2011) suggests that CD laboratory experiments may have caused some people to adopt vegetarianism, a major behavioural change.

An Action Model of Cognitive Dissonance.

In order to provide a comprehensive description of CD and the manner in which it manifests itself, Harmon-Jones and Harmon-Jones (2002) described an action-based model of cognitive dissonance which accepts Festinger's (1957) proposal that the magnitude of dissonance generated by a new cognition is related to the number and importance of inconsistent cognitions already held relative to the number and importance of consonant ones. The negative affect, termed a 'negative emotion' in

Harmon-Jones and Harmon-Jones (2002), can trigger some form of cognitive discrepancy reduction action. However, Harmon-Jones and Harmon-Jones (2002) note that Festinger (1957) was silent about why cognitive inconsistency was aversive. Their model draws on the work of Cacioppo and Berntson (1994) which suggests that cognitions serve as action tendencies. An action oriented state is one "in which intentions are formed to execute behaviors associated with a commitment" (Harmon-Jones, Schmeichel, Inzlicht, & Harmon-Jones, 2011, p 22). The model proposes two types of motivations:

1. A proximal motivation to reduce the dissonance that arises in the presence of the two conflicting cognitions.

2. A distal motivation which stems from the need to take effective action, implicitly to promote the wellbeing of the organism, which in the presence of two conflicting cognitions, and their implied conflicting actions, results in inaction.

The CD provides the proximal motivation to undertake "cognitive adjustment" (Harmon-Jones & Harmon-Jones, 2002, p. 712) or "psychological work" (Harmon-Jones & Harmon-Jones, 2008, p. 1518) to change or make cognitive adjustments which reduce both the negative affect and the cognitive conflict thereby re-enabling effective action to be taken. The overarching distal motivation provides an approach-motivated process (Harmon-Jones et al., 2011) as opposed to withdrawal orientation, which is required for the effective action, once the dissonance reducing psychological work has taken place.

The model is illustrated in Figure 2 below, part 2a illustrating the action tendency in the absence of conflicting cognitions and part 2b in their presence.

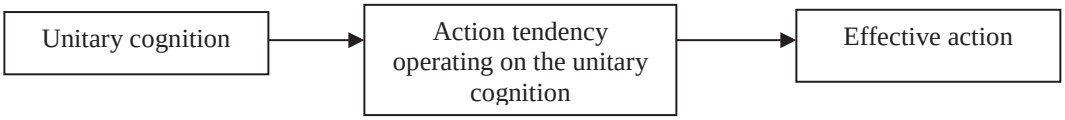


Figure 2a. An action tendency in the absence of conflicting cognitions. Based upon the text of Harmon-Jones and Harmon-Jones (2002).

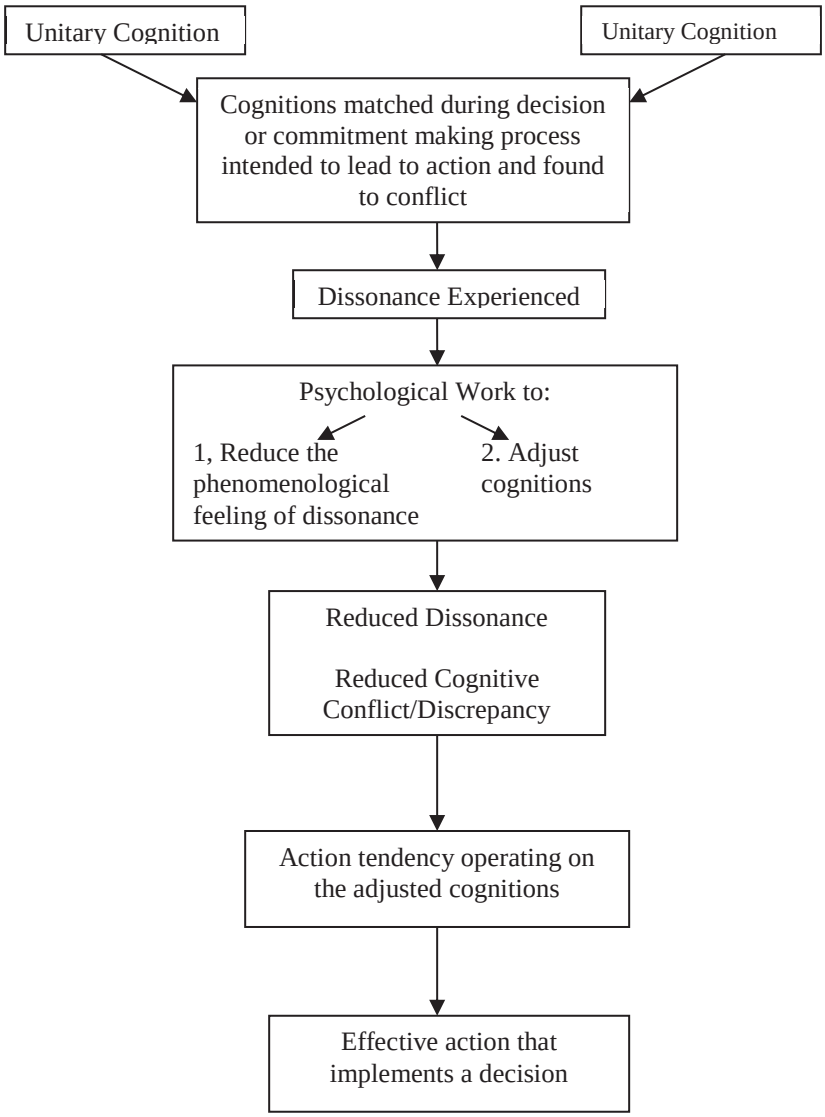


Figure 2b. An action tendency in the presence of conflicting cognitions. Based upon the text of Harmon-Jones and Harmon-Jones (2002).

Summary.

Phenomenologically, CD is an aversive feeling, the intensity of which depends upon a number of factors besides the cognitive conflict and its importance mentioned by Festinger (1957). The large corpus of experimental results demonstrates that the degree to which dissonance is generated may be influenced by the continuous or discontinuous nature of the conflicting cognitions, the likelihood of aversive consequences, the

relevance of the contradiction to self, cultural values, salience, construal level, and personality dispositions. It is broad in scope, capable of embracing an apparently unlimited range of topics, and motivational in its effect, prompting people to take action in the form of one or more of a variety dissonance reduction techniques including seeking more information, thereby enabling effective action. Festinger (1957, pp. 129-131) suggested that if the feeling of dissonance is already intense, the sought information may be in opposition to an original position, in order to trigger a change of attitude.

Comparison of CD with Conventional Emotions

Conventional emotions are widely regarded as products of natural selection (Ekman, 1992; Lazarus, 1991; Plutchick, 2003). In this section, they are described from an evolutionary point of view, the selection pressure for their evolution being the survival and reproductive benefits they confer. Arguments are developed to show that CD is functionally and in some cases, phenomenologically, similar to emotions. By analogy, it is suggested that the selection pressures and benefits that governed the evolution of conventional emotions are also those that governed the evolution of CD.

Lazarus (1991, Ch. 3) in his description of the motivational principle of emotions suggests the following:

1. Motivation can be regarded as the mobilisation of mental and behavioural effort either to achieve a goal or to prevent its thwarting,
2. The motivation is reactive to demands, constraints, and resources present, or potentially so,
3. The motivation incorporates a cognitive element in that to be reactive, the environment must be perceived, assessed against background knowledge and evaluated for implications for personal wellbeing.

According to Lazarus (1991), the sequence of steps within emotional processing are:

1. Cognitive assessment of the actual or potential environment,
2. Assessment and evaluation of the implications for personal wellbeing and consequent goal setting,
3. Generation of the motivation or compelling urge to action,
4. The mobilisation of the mental and behavioural effort to attain the goal.

The sequence of steps itemised above can be regarded as a transaction; it is finite and results in something being achieved. The time scales are usually short; as short as a few hundred milliseconds in the case of fear and fleeing, but longer in other cases. The emotion is usually dissipated by action. Plutchik (2003) formalised the above notions by suggesting that all emotional experiences are characterised by a sequence of related activities. Because the current research proposes that CD evolved through natural selection for the same reasons as emotions did, namely, to produce survival and reproductive benefits, Plutchik's activities appear in Table 1 below with equivalent entries for CD.

Given the similarities between emotion and CD columns in Table 1, CD is hypothesised to be acting identically to a conventional emotion. It may lead to actions of both a mental and physical nature which produce benefits which, proximally or distally, aid survival or reproduction. It is suggested that the Consequent Behaviour of CD may include (1) deciding that the conflict is resolvable via the acquisition of more information and (2) the determining via questioning which of the competing cognitions most nearly reflects reality. The effect is to put the experiencer of CD into a position to make a better based decision, and the function is to produce the benefit of that better decision. This view of CD is supported by Graesser and McMahen (1993) who carried out experiments which showed that contradictions introduced into texts prompted more questions from their participants than other alterations. "Existing research on question asking has uniformly embraced the notion that clashes between stimulus input and world knowledge are very much the essence of question generation" (Otero & Graesser, 2001, p. 144). This view reflects that of Erlich, Guttman, et al. (1957) and Cotton (1985) who recognised that obtaining new information constituted a dissonance reduction technique. Contradictions as triggers for informative inquiry are exploited widely in police forces (Kassim, Leo, et al. 2007) and advocated for research in the humanities (Miller, 2011). The topic of noticing anomalies in texts is also discussed in Appendix A - Discourse Comprehension.

Given the above, it is proposed that the forces that have been responsible for the evolution of emotions have been similarly responsible for that of dissonance, even if it is often experienced more weakly. While in the last part of this chapter, proposals are

Table 1

Plutchik's definitions of emotion related activities, adapted from Plutchik (2003, p. 109) and proposed equivalents for CD.

Characteristic	Plutchik's Definition for Emotions	Equivalent for CD
A stimulus	Sense data; threat, obstacle, loss of object, new territory, etc.	New sense data (Festinger 1957) usually in verbal form.
A cognition	A recognising and/or evaluative cognitive reaction to the stimulus, expressible in propositional terms although no such conscious cognition might be experienced, which initiates the emotional experience.	As for Plutchik's definition for an emotion except that the evaluation consists of the recognition of a cognitive discrepancy (Festinger 1957; Harmon-Jones et al. 2011).
An emotion	A subjective enduring mental state, aversive or pleasurable, that moves the experiencer to action.	The aversive feeling of dissonance (Elliot & Devine, 1994; Festinger, 1957).
Consequent behaviour	The action that takes place following the start of the emotional experience.	Dissonance reduction action as illustrated in the action model of CD in figure 2b. (Harmon-Jones et al. 2002).
An effect	The results of the action, of benefit to the actor.	The same as Plutchik's proposal for an emotion (Harmon-Jones et al. 2002).
A function	To produce the benefit which, proximally or distally, aids survival or reproduction.	The same function as Plutchik's proposal for an emotion is proposed in the current research for CD. It is also suggested in Harmon-Jones et al. (2002).

developed to test this view, the next part elaborates a feature of the mind that must influence whether Plutchik's 'stimulus' is followed by CD and action or not; its modularity and one form of modularity in particular.

Why the Mind must be Modular

The current research is based upon the hypothesis that the mind and brain have evolved through natural selection to be modular in two senses. Modularity of mind refers to any form of non-uniformity of cognitive processing and the modularity of the brain refers to the underlying anatomical and physiological heterogeneity, the corporeal infrastructure, upon which the modularity of the mind is based. In order to lay a foundation for the view of the mind as modular, various forms of non-uniformity of the brain are examined in support of the general principle of heterogeneity. Against this background evidence, an evolutionarily based form of modularity pertaining to cognitive dissonance is proposed.

The mind has not always been regarded as heterogeneous, in the ways implied above, as illustrated in the quotation below.

> Basing our view on the observed fact, that the most diverse types of man may adapt themselves to the same forms of life, I claim that, unless the contrary can be proved, we must assume that all complex activities are socially determined and not hereditary (Boas, 1916, p. 473).

Tooby and Cosmides (1992) called the view of the mind illustrated in the Boas quotation above the Standard Social Science Model (SSSM) of the mind. They cite several other prominent sociologists and anthropologists of the early and middle 20th century who supported this view (Durkheim 1895/1962; Geertz, 1973; Murdock, 1932). Tooby and Cosmides (1992) describe the SSSM as positing the mind as "the more or less passive recipient of culture and is the product of that culture" (p. 32) in which "innate aspects of behavior and psychological organisation are negligible, having been superseded by the capacity of culture" (p. 32).

In a strong criticism, Tooby and Cosmides (1992) drew attention to what they regarded as a fatal weakness of the SSSM. They note that the SSSM requires the contents of the mind to be independent of inherited influence and places no limits on

cultural variation. There is nothing to cause the mind to demonstrate any specialist processes other than those dictated by culture. Mental faculties must perform equally well in any way dictated by the local culture, an ability they call domain-general. Tooby and Cosmides (1992) note that the possible forms of culture are infinite and that therefore a domain general mind must be equally capable of meeting problems that vary to any degree in an unconstrained number of dimensions. Unlimited combinations of circumstance must be dealt with equally well. They claim that such a domain-general mind lacks the means to rapidly identify that much smaller set of circumstances in which speedy actions necessary for survival are promptly taken. A mind that lacked the specific mechanisms adapted to the problems of survival and reproduction would be selected against relative to one that possessed them.

The problems facing a domain-general mind solving a specific problem can be portrayed more concretely than in the abstract argument presented above. The dangers that threaten any organism in a competitive world are many and have to be combated in parallel. Information about threats is available in many forms, for example, sight, sound, and odour, and need to be assessed and processed simultaneously for maximum protection against the threats. In addition, within each of these channels, multiple signals are available for detection, for example, within vision, those of colour, shape, movement, direction, and distance. To extract maximum information about potential threats, any mechanism that can process many incoming signals in parallel is at an advantage relative to another that can process fewer.

There is also the problem of timing. In a competitive predator filled world, the problems of survival require short solution times. When predator and prey confront each other, whichever one can invoke specialist mechanisms which operate more quickly than those of the other is at an advantage. Such a situation sets up an arms race to develop the fastest mechanisms that the cellular nature of life forms can devise. The inevitability of such an arms race is illustrated by the existence of mechanisms with mechanical action times measured in units of 100 milliseconds in the human eye (Kwon, Shipley, Edirisinghe, et al. 2013) and orders of magnitude faster than this in the wings of insects (Scherer, 1995). In support of the survival and reproductive benefits derivable from linguistic abilities (Pinker, 1994), errors in word morphology can be detected in as little as 50-80 ms (Friederici, 2012).

The above arguments and examples and many others that could be cited support the conclusion that survival and reproductive benefits have been the drivers of specialisation across a wide range of domains. They are incompatible with a domain general mind. Examples of specialisations analogous to those described above can also be found within the human brain. Franz Joseph Gall in the early 19th century first proposed that there exist distinguishable cortical locations, each having distinct mental functions. Indeed, he proposed 37 of them, although later 19th century experiments carried out on animals by Pierre Flourens failed to verify this particular proposal (Milner, Squire, & Kandel, 1998). The 19th century discoveries of Broca and Wernicke that the areas in the left hemisphere, now named after them, are active during speech are generally accepted as the first to establish that particular localities specialise to some greater or lesser extent in different types of psychological processing. By 1881, Munk had established that visual activity was concentrated in the occipital lobe (Hoyt, 2003). The CAT, PET, MRI, and fMRI technologies developed in the later decades of the 20th century have enabled particular functions to be associated with particular areas of the brain with ever greater precision.

While critics of the early PET and fMRI work claimed that the results were being over- and contradictorily interpreted (Uttal, 2001) subsequent refinement of equipment and experimental methods have established the value of the technologies beyond dispute. Indeed, of particular interest to cognitive and evolutionary psychologists, Just, Cherkassky, Aryal, and Mitchell (2010) describe areas of the cortex which are activated when such words as *screwdriver*, *lettuce* or *house* are read. From the activation displays of 40 such words, the cortical areas were identified in which the representation of such semantic categories as manipulation, eating and shelter-entry are found. Each concept representation is distributed over unique combinations of shared cortical areas, a cortical archipelago (after Kanwisher, 2010), suggesting that finer localised semantic discrimination may yet be discoverable. Kanwisher (2010) and colleagues have already demonstrated that the fusiform face area, the extrastriate body area and the parahippocampal place area largely devote their processing functions to the face, body, and place (locality), after which they have been named. Recent work by Federenko, Behr, and Kanwisher (2011), Fedorenko (2014), and Fedorenko and Thompson-Schill (2014) has established that multiple identified cortical areas, some

more and some less specialist, are recruited to process language. In their research on language comprehension, Mesulam, Thomson, Weintraub, and Rogalski (2015) note "that the left anterior temporal lobe needs to be inserted into the canonical language network with a critical role in single word comprehension and object naming" (p. 13). These researchers have collectively demonstrated a high degree of functional specialisation or cognitive heterogeneity of the brain.

The histological work of Golgi, Cajal and Brodmann in the late 19th and early 20th centuries identified a wide variety of differentially distributed cell types in the human brain (Finger, 1994; Jones, 1999). The work of cell categorisation has culminated in the identification today of 145 types of neuron, according to one system of classification (Vickaryous & Hall, 2006). Current work on the mouse brain (Grange, Bohland, et al. 2014) is identifying how different cell types are distributed in the brain and discovering their individual gene expression profiles. In sum, the cellular non-homogeneity of the brain manifests itself in multiple types of neurons, their differential distribution and differential gene expression, together providing further support for a modular view of the brain.

Recent work within the Connectome Project (Van Essen, Smith, et al., 2013) to trace the brain's 'long-range' communications networks provides convincing evidence of differential degrees of neural interconnectivity. The project is rendering a wiring diagram of the brain, illustrating the non-uniform density of connections between different regions of the brain. Comparisons across 1200 participants, which included mono- and dizygotic twins, revealed differences in brain circuitry related to particular behavioural capacities and functions. In experiments requiring participants to undertake a variety of task types, "Each task [type] engages a complex network of brain regions that are partially overlapping with those of other tasks" (Van Essen, 2015). Some interpersonal connectome differences have been found to have genetic causes (Thompson, Ge, et al., 2013). Other connectome studies are revealing the relation between particular variations in cerebral connectivity and psychological disorders (Korgaonkar, Fornito, Williams, & Grieve, 2014). Mišic and Sporns (2016) note that connectivity patterns between modules can be related to behaviour and cognitions but also note the highly dynamic nature of connections between them according to the task being undertaken.

Clune, Mouret, and Lipson (2013) have made computer simulations of neural nets. In an object recognition task, they applied selection pressures, in the form of algorithms to reduce the number, and therefore the cost, of connections between nodes in the network. Algorithms were of two types; one to minimise connection numbers, and therefore connection costs, while still targeting high performance and a second control algorithm devoted to maximising performance. They found that the application of the first algorithm causes the emergence of a modular network, defined as one with fewer connections between nodes but still able to complete the task. Such networks are significantly more evolvable, meaning able to change, than the networks having more connections between nodes produced by the control algorithm. They propose that cerebral network modularity has evolved because, compared to non-modular networks, responses to environmental changes can take place more quickly and connection costs are lower.

Conclusions on Modularity.

Three types of cerebral non-homogeneity have been briefly described above to support the view that the brain is modular. Functional specialisation is distributed over the cortex, impressive degrees of neuronal differentiation have been discovered and widely different degrees of linkage between regions of the brain have been found to exist. In the search for an explanation for humanity's numerous psychological attributes, which may be thought of as psychological modules, the different types of neurological non-homogeneity enumerated above have probably provided the main answers. However, no one to one equivalence is claimed. "In brain terms, it could be, and probably is, that macroscopic regions of brain tissue include neurons from multiple information-processing systems with multiple functions" (Barrett & Kurzban, 2006, p. 641). The mapping of psychological function to neural substrate is relatively new and complicated pictures are emerging. One reason for these complications may be the huge number of psychological functions or modules that have to be mapped on to the brain, a topic explored in the next section.

The Evolutionary Psychology Modularity Hypothesis.

The view that the physical features of the brain outlined above are the specialised modular products of natural selection would attract little opposition. The view that the mind is also a specialised modular product of natural selection has been contested, as was illustrated in the account of the SSSM. However, a modular view of the mind was famously presented at a meeting of the Royal Society of London and the British Academy in London in 1995. Titling her lecture "The Evolution of Social Behaviour: Patterns in Primates and Man", Leda Cosmides commenced by holding up a Swiss army knife and presenting it as a model of the mind (Mithen, 1996). Since that time, the Swiss Army knife as a metaphor for the human mind has been indissolubly linked to her and her husband John Tooby. Less metaphorically, they regard the mind as "a diverse collection of evolved mechanisms, many of which are functionally specialist, domain-specific, content-imbued, and content-imparting" (Cosmides & Tooby, 1992, p. 221). These specialised psychological mechanisms enable us to function in a highly variable environment. The mind "… can be likened to a confederation of hundreds of thousands of functionally dedicated computers." (Tooby & Cosmides, 1992, p. 39). Tooby and Cosmides (1992) and Cosmides and Tooby (1997) propose the following simple principles that underlie their view of the brain and the mind:

1. The brain is a physical system. Its circuits are designed to generate behaviour that is appropriate to the environmental circumstances in which natural selection shaped them.

2. Our neural circuits were designed by natural selection to solve problems that our ancestors faced during our species' evolutionary history.

3. Consciousness is just the tip of the iceberg; most of what goes on in the mind is hidden. Conscious experience can mislead someone into thinking that the circuitry is simpler than it really is. Most problems experienced as easy are very difficult to solve as they require very complicated neural circuitry.

4. Different neural circuits are specialised for solving different adaptive problems. There are a very large number of such circuits.

They propose the term module for "complex structures that are functionally organized for processing information" (Tooby & Cosmides, 1992 p. 33). Their concept and use of the word 'module' has been attacked by some psychologists who claim that mental activity can not be broken down into localised single or distributed 'bits' of the brain (Uttal, 2001) or can not be thought of as collections of little computers processing information in the format and manner of Von Neumann machines (Wallace, 2010). Tooby and Cosmides have not generally provided descriptions of or made specific predictions as to how their functionally specialised modules are realised by particular neural configurations, despite their conviction that they are so realised and that these modules are adaptations. In support of the Tooby and Cosmides adaptationist view, Buss, Haselton, Shackelford, Bleske and Wakefield (1998) proposed the following definition of an adaptation which covers the Tooby and Cosmides concept of modules although, of course, not all adaptations are modular.

1. An adaptation is an inherited and reliably developed characteristic that came into existence as a feature of a species through natural selection because it helped to directly or indirectly facilitate reproduction during the period of its evolution. Buss et al. (1998) propose that solving an adaptive problem is the function of the adaptation.

2. Each adaptation has its own Environment of Evolutionary Adaptedness (EEA) which refers to the cumulative selection processes that constructed it, piece by piece until it came to characterise the species. The EEA will differ for each adaptation and is usually described as the statistical aggregate of selection pressures over a particular period of time which were responsible for its emergence.

Despite an emphasis on modularity, Ermer, Guerin, Cosmides, Tooby, and Miller, (2006) do allow that a limited degree of content-free processing may take place but even then, "the inferential processing appears specialized for solving particular adaptive problems" (p. 197) particularly those of social exchange. They provide further examples such as intuitive or naïve physics and biology and theory of mind which reliably develop in the human mind in the absence of any form of associated

socialisation. They are regarded as reasoning instincts but operate only within the limited domains mentioned.

Shallice and Cooper (2011, 2012) provide an assessment of the parallel investigations of brain-based and information processing researchers which encompasses a range of views of the modularity of the mind, from those of the philosopher Fodor (1983), in which some degree of modularity is allowed, to the massive modularity view of Tooby and Cosmides (1992). There appears to be no consensus as to how the word modularity is used although there appears to be a broad consensus that the mind/brain is in some sense modular. The brain is neither functionally nor anatomically homogeneous. I propose therefore that modularity be regarded as a general term to refer to the ability of a single identifiable localised fragment or interconnected set of fragments of brain tissue to specialise in particular functions, no matter how widely or narrowly defined. One can thus refer to something as a module if it performs a function as wide in scope as vision or, more usefully, as narrow as merely biasing some particular processing more in one direction than in another. This latter notion is reinforced by Krill, Platek, Goetz, and Shackelford (2007) who suggest: “Psychological mechanisms can also be expressed as *cognitive biases* (italics added) that cause people to more rapidly attend to some pieces of information relative to others” (p. 234). If two similarly formatted pieces of information are not processed in a content-free manner but, rather, their processing is biased according to some differences in their content, then it is proposed that the mechanism causing the bias can and should be regarded as a module. This view of the term module permits its application to cognitive dissonance, and in particular to the issue of the importance of the topic of a cognitive discrepancy. If one objectively defined category of cognitive discrepancies can be demonstrated to be more important than others, in the sense that there is a greater sensitivity to this category of discrepancies relative to others, the mechanism responsible for this difference can be regarded as another example of modularity. The current research is aimed at demonstrating the existence of this form of modularity.

The Embodiment of CD

The action model of dissonance (Harmon-Jones et al., 2002) and the modular view of the brain/mind together imply that if CD is a product of natural selection, somewhere in the brain must reside the seat of CD. Gemba, Sasaki, and Brook (1986) detected transcortical signals in the Anterior Cingulate Cortex (ACC) when monkeys *(Macaca fuscata)* were "observed during a transitional stage of uncertainty in motor learning." (p. 223) and suggested that errors might be the trigger. Gehring, Goss, Coles, Meter, and Donchin (1993) carried out experiments with human participants having electroencephalogram measurements taken from midline and lateral scalp sites while they performed a visual identification task under time pressure. The position of the cerebral Error Related Negative signal detected when participants made errors in the task suggested that the ACC was likely to be an error detection locus. Botvinick, Nystrom, et al. (1999) proposed that the ACC detects and signals the occurrence of conflicts in information processing. They tested this hypothesis in an fMRI experiment using a flanker task which required participants to detect whether a target arrow, flanked on both sides by two distracter arrows, pointed to the right or the left, indicating with suitable button presses. Their study revealed a region within the dorsal ACC (BA 32) where peak activation was greater during incompatible trials in which the target and flanker arrows pointed in different directions, than compatible ones. It also showed that the activation was more strongly related to conflict than to selection for action, although the authors note that activation of the ACC might be important in triggering cognitive control.

Kiehl, Liddle, and Hopfinger (2000) noted the heterogeneous and complex neural architecture of the ACC, with its rich afferent and efferent connections to the limbic, association and motor cortexes, and note the likely equivalent functional breadth. In particular, they draw attention to the affective processing involved in error detection and correction. Kiehl et al. (2000) conducted an fMRI experiment in which they required participants to respond with a button press to the display of an X but not to respond to a K. There was substantial activation in the rostral anterior cingulate cortex and the left lateral frontal cortex during button presses made in error.

In more recent years, some early CD research has been replicated using fMRI techniques to demonstrate ACC activity during CD (Kitayama, Chua, Tompson, & Han, 2013; Van Veen, Kruk, Schooler, & Carter, 2009). In a very recent fMRI study, de Vries, Byrnes, and Kehoe (2015) developed a novel procedure to identify the location of dissonance but not that of dissonance reduction, using accounts of everyday experiences. Activations were widely distributed supporting Uttal's (2001, 2011, 2013) contention that "high level cognitive processes are associated with widely distributed activity in many parts of the brain" (Uttal, 2001, p. 25) as recalling every day experiences is itself likely to lead to widely distributed activations. The issue of widely distributed activity is briefly explored in Appendix A - Discourse Comprehension. Nevertheless, de Vries et al. (2015) note that relative to participants in control conditions, participants in a dissonance condition showed more activity in the ACC, the anterior insular cortex, inferior frontal gyrus, and precuneus.

In sum, experiments conducted over the last three decades appear to have identified the ACC as the seat of CD although some qualifications appear necessary. The earlier experiments, using less sophisticated equipment, identified the ACC possibly precisely because their equipment was able to record only the strongest signals. As the later more sophisticated equipment was introduced a wider distribution of activations was detected. It is unclear whether the newly detected wider distribution of activation is due to the generation of dissonance only or whether these wider signals are due to other facets of the experimental methods employed. Uttal, in his several works, legitimately draws attention to the difficulties of precisely locating a psychological phenomenon that is almost inevitably accompanied by others within the time interval of seconds measured by the fMRI techniques. Nevertheless, the series of experiments noted above, and others, consistently identify the co-occurrence of dissonance and activation in the ACC. The expectation generated from the review of CD and the evolutionary modular view of the mind that the seat of CD should be identifiable appears to have been realised. Details of the hypothesised selection pressure responsible for this evolution are discussed in a later section.

The action model of CD (Harmon-Jones & Harmon-Jones, 2002) implies not only that the seat of CD should be identifiable but the seat of the action tendency following dissonance should be also. For an account of this, a preliminary review of

emotional valence is necessary. Harmon-Jones (2004) discusses the positivity and negativity of emotional valence and notes three factors that are variously considered as defining it: the conditions that evoke it, the adaptive consequences and the emotion's subjective feel. Citing Lazarus (1991), he notes that this last factor normally leads to anger being considered to have a negative valence. He briefly reviews research covering the relative Left Frontal Cortical (LFC) activity and approach-related positive emotion on the one hand and relative Right FC activity as the neural correlate of negative withdrawal–related emotions on the other. He concludes that this relationship is an over-simplification, given that anger, while having a negative valence when considered subjectively "often invokes approach motivation" (p. 338). He reports a set of experiments which essentially confirm the view that the pairing of an emotion with the Left or Right FC activity should take place, not according to its valence as defined by subjective feeling, but according to its approach or withdrawal associated motivation. Harmon-Jones, Lueck, Fearn, and Harmon-Jones (2006) report an experiment which manipulated the degree to which participants might expect to take angry-approach related activity. They found that the greater the expectancy of action, the greater the relative LFC activity, confirming the view that relative LFC activity is related to the approach-motivation of an emotion, rather than the subjective experience it gives rise to. Harmon-Jones, Harmon-Jones, Serra, and Gable (2011) carried out a forced choice experiment while using EEG equipment and found that high choice participants manifested higher activity in the LFC than low choice participants even when acting pro-attitudinally. They conclude that commitment to action followed by approach-motivational processes are reflected in higher LFC activity.

In a general review article, Carter and Van Veen (2007) note that whenever any of a wide variety of mental representations are compared and conflicts detected, cognitive control mechanisms located in the Dorsolateral Prefrontal Cortex are recruited to initiate the appropriate action to reduce the conflict. In a more recent review of several experiments to test the Action-Based Model of Dissonance, Harmon-Jones, Harmon-Jones, and Levy (2015) note the conflict monitoring role of dissonance (Van Veen, Kruk, Schooler, & Carter, 2009) and its consistency with the Action-Based Model, as many forms of cognitive discrepancy imply consequential action.

The experiments mentioned above suggest that the evolved neurological substrates for both elements of the Action-Based Model of CD have probably been identified. When dissonance is experienced, with correlated activity in the ACC, the consequential action tendency and approach processes have correlated neurological activism located in LFC. This conclusion has been supported by the findings of other avenues of research quite separate from CD. It seems that CD can not be properly understood without a consideration of these. The next section is therefore devoted to a brief review of these researches in order to provide a fuller context in which CD should be viewed.

A Wider View of Cognitive Dissonance

In a wide ranging article, Jonas, McGregor, et al. (2014) propose a model of the mind reacting to cognitive discrepancies which unifies several areas of research previously conducted independently. They identify common elements of CD, Terror Management Theory, Control Theories, the Meaning Maintenance, the Unconscious Vigilance, and Reactive Approach-Motivation (RAM) models, and propose a theoretical framework they call Threat and Defense which combines them. They note that all the earlier theories incorporate some form of cognitive discrepancy, similar to that first identified by Festinger (1957) and propose that the brain treats all forms of discrepancy as a threat, leading to anxiety which motivates actions aimed at reducing this aversive form of arousal. The similarities imply that an evolutionary process that led to the existence of CD may also have led to the existence of these others.

Terror Management Theory (TMT) identifies a discrepancy between "the awareness of the inevitability of death in an animal instinctively programmed for self-preservation" (Jonas, McGregor, et al., 2014, p. 61) and describes "the broad influence of the problem of death on human social behavior and illuminates the processes through which concerns about mortality exert their influence" (p. 61). According to Greenberg, Solomon, and Pyszczynski (1997), TMT employs a dual-process model of defences against the anxiety surrounding mortality salience. Proximal defences can be used to remove conscious death-related thoughts from attention either by suppressing such thoughts with distractions or by pushing the problem of death into the distant future. Promising to diet or starting an exercise programme, or "focusing their attention on

whatever evidence there might be to support long life expectancy" (Greenberg et al., p. 841) might have the necessary distracting effect. Distal and unconscious defences employ symbolic cultural conceptions of oneself and reality, for example, positing forms of lengthy or even eternal existence, as ways of dealing with the unconscious knowledge of the inevitability of death. Relevant cultural concepts would include religious beliefs of life after death or the creation of secular institutions such as foundations or named buildings.

Control Theories emphasise the need for dominion over ones life and note the discrepancy between the sought-after degree of control and reality. Kay, Whitson, Gaucher, and Galinsky (2009) suggest that belief in personal control buffers individuals from the uncomfortable reality that randomness and chance often determine important life outcomes. "To insulate themselves from the disabling anxiety that perceptions of randomness incite, individuals psychologically imbue their social, physical, and metaphysical environments with order and structure." (p. 264). Fritsche, Jonas, Ablasser, et al. (2013), as advocates of group based control, emphasise collective action and the self as a member of a group. They demonstrated that people show increased support for in-group reactions, out-group derogation and pro-organizational behaviour when their sense of global personal control is threatened.

McGregor, Nash, and Inzlicht (2009) propose a Reactive-Approach Motivation (RAM) model. They conducted experiments in which high self-esteem participants were put in threat and no-threat conditions and noted that threat condition participants manifested high Alpha band (8-13 Hz) EEG readings in the left DLF cortex, relative to the right, associated with approach motivation. They propose that their results be "interpreted together with past research showing that individuals with high self-esteem react similarly to various threats with diverse outcomes which seem to share the common denominator of approach motivation." (p. 1005).

McGregor, Nash, Mann, and Phills (2010) conducted a multi-method set of threat-related experiments designed to demonstrate that approach-motivation is domain general in the sense of being invoked by a variety of threat types. In all experiments, uncertainty threats produced evidence of Reactive-Approach Motivation. In a comment that anticipates the conclusion of Jonas et al. (2014), the authors suggest a link between

the findings of other experiments reported in the cognitive discrepancy literature and the results of their experiments: “Perhaps foremost is the link to action” (p. 142).

Holbrook, Sousa, and Hahn-Holbrook (2011) propose a model of unconscious vigilance, which although similar to the Terror Management Model, does not necessarily involve mortality-salience. They carried out experiments in which threat manipulations unrelated to death or coalitional concerns evoked a worldview defence, that is, a polarisation of ratings of pleasant and aversive cultural attitudes. They propose that “subtle threats heighten our sensitivity to emotionally evocative stimuli” (p. 463) and suggest that unconscious vigilance may be an adaptation designed to “marshal identification of and responsiveness to background hazards or resources.” (p. 461).

Proulx and Inzlicht (2012) have researched a model of the mind quite close to Festinger’s (1957). Describing the Meaning Management Model (MMM), they define meaning as “expected relationships” (p. 318) and propose that “[V]iolations of how people understand their experiences arouse a common psychological discomfort.” (p. 320). The MMM combines meaning violations or discrepancies with activity in interconnected neurocogntive and psychophysiological systems which detect the inconsistencies and motivate compensatory behaviours, described as palliative as they tend to reduce the aversive arousal. The authors locate the relevant neurocognitive system in the anterior cingulate cortex and note correlated psychophysiological activity such as enhanced skin conductivity, increased heart rate, pupil dilation, and startle eyeblink response, physiological responses similar to those found by CD researchers (Croyle & Cooper, 1983; Elkin & Lieppe, 1986; Losch & Cacioppo, 1990).

Proulx and Inzlicht (2012) draw upon findings from “every corner of the psychological literature” (p. 318) and categorise the compensatory efforts as follows:

1. Assimilation: absorption of apparent discrepant experiences during early life in the building of schemata or expectations so that they better reflect reality (Piaget, 1937/1954). Proulx and Inzlicht (2012) use the same term to refer to the cognitive processing that takes place when undesirable ambivalence is experienced as described by van Harreveld, van der Pligt, and de Liver (2009) in their Model of Ambivalence-Induced Discomfort (MAID). Procrastination and denial of responsibility are first employed to diminish the aversive arousal of ambivalence but

if this is unsuccessful, unbiased and biased processing is employed to validate only one attitude in an attempt to produce a single understanding.

2. Accommodation: a change in schemata so that they allow for the discrepant experience, usually in the form of an alteration in one of the previous attitudes (van Harreveld, et al., 2009). This is the capacity for change of belief exploited in the Induced Compliance CD research paradigm.
3. Affirmation: the enhancement of commitments to previously endorsed views when a particular view or value is violated by discrepant experiences (Kay, Moscovitch, et al., 2010; Kay, Whitson, et al., 2009). This contrasts with the dissonance reduction technique of trivialization (Simon, Greenberg, & Brehm, 1995). Both techniques change the relative importance of a set of cognitions, the trivialization diminishing the rejected view and affirmation enhancing the retained one.
4. Abstraction: the generation of a pattern where none might exist or the finding of meaning in a set of events that are apparently unrelated and inconsistent - to see "signals in the noise" (Proulx & Inzlicht, 2012, p. 328).
5. Assembly: the creation of more powerful encompassing frameworks of meaning that make sense of the negative or discrepant experiences (Park, 2010), or to "make sense out of senselessness" (Proulx & Inzlicht, 2012 p. 239). This is similar to dissonance reduction by restructuring cognitions not directly related to a specific attitude-behaviour discrepancy or generating some background belief or opinion which lowers the discrepancy (Leippe & Eisenstadt, 1999).

Proulx and Inzlicht (2012) emphasise the common factor in these five perspectives: a cognitive change to reduce or eliminate inconsistency. In proposing their Meaning Maintenance Model, they regard the different processes listed above as describing the same general core phenomenon. In an article which combines much of the above with the output of nearly sixty years of research into cognitive dissonance, Proulx, Inzlicht, and Harmon-Jones (2012) draw attention to what they call a "potentially unifying inconsistency compensation perspective" (p. 285) in which any cognitive or perceptual inconsistency may give rise to an aversive arousal. This in turn is understood to motivate the various palliative efforts described by Proulx and Inzlicht (2012), taking specific forms according to the particular branch of research investigating them. Proulx, Inzlicht, and Harmon-Jones (2012) suggest that the generation of multiple theories, explanatory mechanisms, and compensatory processes "has meant that we

know remarkably little about why we engage in these efforts" (p. 289/290) and call for a "unitary discipline-straddling account of inconsistency compensation" (p. 290).

Jonas, McGregor, et al. (2014) may fairly claim to have answered this call in synthesising their Anxiety to Approach model of threat and defense. Their model unifies and incorporates the discrepancy detection (threat)→aversive arousal→approach-related palliative effort schema that characterises the earlier models but also incorporates a flexibility that permits a variety of responses. They recruit "the well known Behavioral Inhibition System (BIS)" (Corr, 2011, p. 889) which they claim evolved to regulate effective responses to impeded goals. The BIS processes threat-related discrepancies and initiates one or more of hypervigilance, anxiety arousal, avoidance motivation and inhibition of current behaviour. If an effective solution to the discrepancy is available, it is implemented immediately, thereby shortening the anxiety and vigilance of the BIS stage. If a solution is not available, for example, because the threat pertains to the future, measures could provide an indirect solution or be merely palliative. Of relevance to this research, they note that these basic processes can include "an increased vigilance for new information" (p. 230). Citing a wide range of studies, they locate the discrepancy detection mechanism in the anterior cingulate cortex (ACC), behavioural inhibition in the ACC and the right prefrontal cortex (PFC), emotional regulation in the left and right PFCs and the BAS in the left PFC. They conclude that all the threat and defence perspectives commence with the detection of motivationally relevant discrepancies which activate the BIS. Either immediately or after an interval, approach-related responses occur in the BAS which mute the BIS generated responses.

The discrepancy/threat/anxiety/approach model proposed by Jonas, McGregor, et al. (2014) provides a unifying explanation for a wide range of discrepancy-initiated psychological experiences, each with its own school of research and history of discovery. This unity together with their common anatomical infrastructure implies very strongly that there is a single evolutionary explanation for their existence. If an explanation for the evolution of one of these very similar cognitive discrepancy mechanisms can be verified through experimentation, it is suggested that it illustrates the evolution of them all. The next section proposes such an explanation for CD and lays the groundwork for a discussion in Chapter 7 - General Discussion which proposes an explanation for the evolution of them all.

Experimental Considerations

In the discussion of the factors influencing the intensity of dissonance, accounts of experiments were provided which, with qualifications, verified Festinger's (1957) view that the strength of dissonance is related to the importance of the topic of the conflicting cognitions. If different categories of cognitions have different degrees of relevance to our survival and reproductive success, as is proposed in the current research, it is predicted that cognitions pertaining to short term survival, material welfare, reproductive success and related topics should be much more important to us in the sense that Festinger (1957) meant and therefore more powerful in their potential contribution to dissonance than others of little relevance.

The description of Festinger's (1957) pairing of cognitions given earlier drew attention to three categories of pairs of cognitions; consonant, neutral and dissonant.

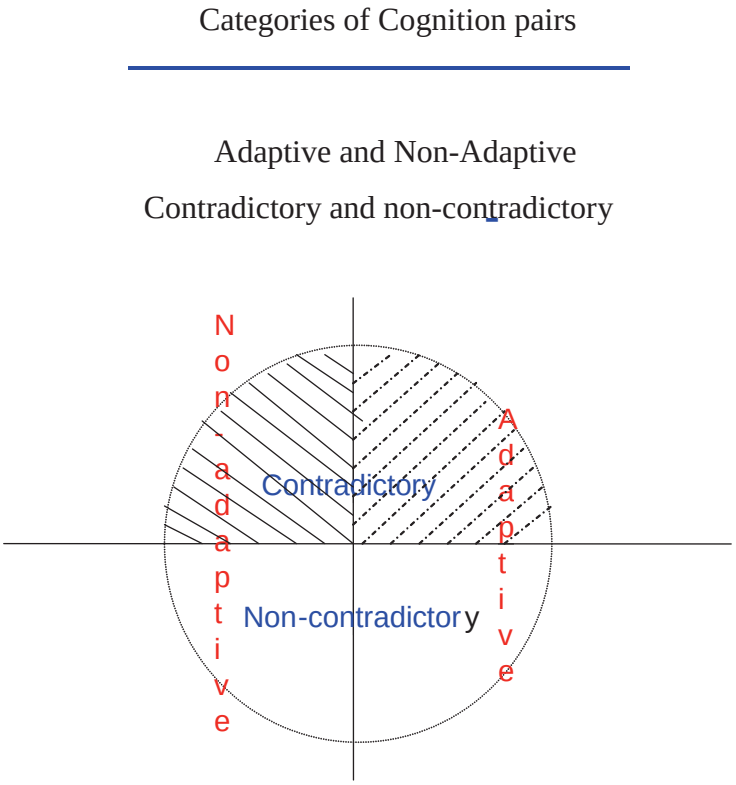


Figure 3. Schematic diagram of a 2 by 2 division of pairs of cognitions.

The first two can be combined into a non-contradictory category, while the last constitutes a contradictory category. A further categorisation can be introduced; pairs that are of adaptive relevance and those that are not. These concepts are illustrated in Figure 3.

The cognition pairs of interest are those in the top right quadrant, those that are both adaptive and contradictory. In more detail, important survival related sets of conflicting information could pertain to the abundance, condition, availability, location, and direction of predators and sources of food, water, and shelter. Similarly, reproduction-related conflicting information could pertain to the sexual attractiveness, parental inclinations, fertility, and fidelity of potential and actual mates. A contradictory set of statements about a woman's treatment of her own children or other children in the family would be of immensely greater importance to a potential future partner than contradictory views about her liking for particular pretty stones. It would be important for him, and his potential progeny, to notice the first contradiction but of little consequence if he did not notice the second.

Paleoanthropological studies (Karkanas et al., 2007; Marth, Schuler, et al., 2003) suggest that during the period in which humanity evolved the psychological attributes which characterise it today, survival was always at risk and became progressively more dependent upon co-operation and reciprocity as these human attributes evolved. In these circumstances, the extent to which a particular person was the beneficiary of others' actions depended upon the extent to which those others could depend upon the cooperation, honesty, and reciprocity of that person. One's own reputation for cooperation, honesty, and reciprocity would strongly influence how much benefit one received from others. It follows that ascertaining which of two contradictory cognitions pertaining to someone else's prosocial or reputational attributes was true would, on average, be adaptive as this determination prevents the provision of benefits to those who were unlikely to return them. In contrast, attention paid to conflicting information irrelevant to adaptive topics would be maladaptive. It would lead to unnecessary energy expenditure at best but occasionally might lead to danger.

It is hypothesised that attention paid to conflicting information in the three adaptive categories discussed above following the experiencing of CD, would on average have consequential reproductive benefits. These benefits constituted a selection

pressure for the development of a capacity for CD, acting marginally and on average over thousands of generations. A sensitivity bias towards noticing adaptive contradictions relative to Non-Adaptive ones can be regarded as evidence of this selection pressure. The current research has been aimed at detecting such a bias. Its existence would provide evidence for the role of natural selection in the evolution of Cognitive Dissonance.

The everyday discrepancies of the type explored above and the non-CD discrepancies in Jonas, McGregor, et al. (2014) are mostly word-based, although their model allows perceptual discrepancies to have a similar effect. While it cannot be known how frequently perception-based discrepancies were met in our pre-linguistic past, it is hypothesised that as our capacity for language evolved, the opportunities for discrepant experiences would have multiplied. It is suggested that during that period of our evolutionary past when we were evolving our capacity for language, the development of language and the capacity for CD would be mutually reinforcing. A listener's noticing a contradiction between someone else's statement about the environment and previous knowledge or other statements would be beneficial only if the listener, feeling CD, was motivated to resolve the discrepancy and sought more and better information upon which to base subsequent action. The greater the capacity for CD, the more likely that further information would be sought and the benefits obtained. CD imposes a selection pressure for a greater capacity for language. In short, CD drives language. At the same time, the more advanced someone's capacity for language, the more capable they would be of noticing a contradiction and the more frequent the opportunities for CD to drive further questioning. Better language would drive CD. It is appreciated that for a greater sensitivity to contradictions according to topic to have evolved, it is probable that our language capacity developed over an extended period of time. Neither a language capacity nor a bias within it is likely to have arisen as a saltation (Tallerman, 2014). The research-based evidence for this lengthy period of language evolution has been collected and is presented in Appendix B – The Evolution of Linguistic Abilities.

In the foregoing sections, the following themes have been emphasised: dissonance as defined by Festinger (1957) is mostly experienced as a result of cognitive discrepancies conveyed by words, the capacity for CD and its neurological substrate

likely evolved over a lengthy period contemporaneously with the evolution of language, and topics of adaptive relevance, survival, reproduction and reputation (SRR) are more important to our wellbeing than Non-Adaptive ones. It was hypothesised that the survival and reproductive benefits of resolution of adaptive cognitive discrepancies constituted the selection pressure for the evolution of a stronger sensitivity to these discrepancies than to Non-Adaptive ones. This greater sensitivity to adaptive contradictions can be regarded as the bias in cognitive processing mentioned in the section on the modularity of the mind. Experiments which demonstrate this greater sensitivity or bias provide evidence in support of the role of natural selection in the evolution of CD.

Experimental Predictions

Using the contradiction paradigm, described in Chapter 3 - Item Construction, Clemerson (2009) briefly explored the possible evolutionary origins of CD by presenting text which included contradictions, explicit and implicit, covering three topic categories: survival, reproduction and abstract. It was predicted that the contradictions related to the first two topics would be more important than abstract ones and therefore preferentially noticed. Analyses showed this to be the case for reproduction-related contradictions but the survival-related ones proved significantly less noticeable than either of the other topics. To explain this apparently anomalous result, recourse was made to Terror Management Theory (Pyszczynski, Greenberg, & Solomon, 1999), which proposes the emotional and behavioural responses to reminders of mortality described earlier. It was suggested that the survival-related contradiction drew participants' attention to their own mortality in the form of threats to life within a first person account of exploration in dangerous locations. When participants were presented with these mortality-related statements on a computer screen one line at a time, they were able to reduce the impact of the aroused aversive emotions by diverting their attention to the next line of text by quickly pressing the space bar.

Other evidence does suggest a greater sensitivity to survival-related matters than to others. Nairne, Thomson, and Pandeira (2007) and Nairne and Pandeira (2008) required participants to assess each element of a word list for relevance to a number of subjects, including survival. After a distraction activity, recall tests demonstrated a

significantly greater ability to remember the survival-related words than others. None of Nairne and colleagues' survival-related words included 'life' or 'death' or their synonyms. Rather, they consisted of words such as *juice*, *soccer* and *bear*. Two categories of Survival related contradictions can therefore be identified:
1. a life threatening set used in Clemerson (2009),
2. a non-life threatening set featuring a range of survival related possessions such as food, clothing, and shelter, similar in principle to those of Nairne et al. (2007, 2008).

Given the findings of Nairne et al. (2007, 2008) and Clemerson (2009) and the implications of Figure 3, a prediction (Prediction 1a) can be made that contradictory pairs of cognitions in a Survival group comprised of the two categories of survival-related contradictions noted above will be more noticeable than contradictory pairs in a Non-Adaptive group. Because of the immediacy of threats and the explicit phrasing used in the life threatening case, it can also be predicted (Prediction 2a) that the Life Threatening contradictions will be more noticeable than the non-life threatening contradictions. However, it must be noted that the counterintuitive result obtained in Clemerson (2009) hints that the opposite prediction should be made.

Contradictions pertaining to reproduction also fall into the adaptive category illustrated in Figure 3. As for Survival, a set of more specific topics can be identified as representative of the more general topic of Reproduction. In the current research, Child-Loss, Family, and Fidelity, drawn from the Evolutionary Psychology literature (Buss, 2005) were adopted. The prediction (Prediction 1b) was made that this Reproduction group of contradictions would be more noticeable than Non-Adaptive contradictions. Within this group, Child-Loss was regarded as the topic most likely to manifest the biggest contrast with the Non-Adaptive topics on the grounds that it most directly addresses the issue of the passage of genes into the next generation, in this case, in the form of a failure to do so. The Family topic tests attitudes to reproduction rather than reproductive success or failure. It was predicted (Prediction 2b) that Child-Loss contradictions would be more noticeable than those pertaining to the Family topic. Fidelity was not rated relative to the two other topics as the decision to include it in the experiments was taken at a later date and the predictions were not revised.

Although people are jealous of their reputations in many domains, it was hypothesised that two would have been of particular importance in evolutionary times; honesty, referring to refraining from theft, lying, slander, and promise breaking, and reciprocity, referring to ones preparedness to return favours. These topics were therefore chosen as representative of the more general topic of Reputation. It was predicted (Prediction 1c) that the Reputation group of contradictions will be more noticeable than Non-Adaptive contradictions.

It was also hypothesized that dishonesty referring to theft, etc. would damage a reputation more immediately due to the direct personal consequences to the victim than a failure to reciprocate a service. In the latter case, the victim has to detect a pattern of failure to reciprocate over a period of time. In addition, damage to reputation can be mitigated by a partial or occasional return of favours by the perpetrator, making it harder for a victim to notice the pattern. Because of the immediacy of the former, it was predicted (Prediction 2c) that Honesty-related contradictions would be more noticeable than the Reciprocity ones.

Two Non-Adaptive categories of contradictions were identified; Socially Important contradictions and Socially Unimportant ones. All people undergo processes of socialization which may enhance their sensitivity to the first category. It was therefore predicted (Prediction 2d) that the Socially Important contradictions would be more noticeable than the Socially Unimportant ones. Both topics were proposed as the basis for comparisons with the Adaptive ones.

Predictions were made pertaining to the ranking of the adaptive topics. As survival is a prerequisite for the passage of genes from one generation to the next, it was hypothesised that the survival contradictions would be more noticeable than the other adaptive contradictions. As reputation is a supportive attribute in the passage of genes into the next generation but is not a necessary one, the reputation group would therefore be the least noticeable. A prediction (Prediction 3a) was made that the ranking of noticeability for the three adaptive groups would be, first, Survival, second, Reproduction and last, Reputation. As evolutionary theory requires all psychological attributes to have as their ultimate purpose the passage of the genes for those attributes into the next generation, and as reproduction is the mechanism by which this is

achieved, a second and alternative prediction was made that the Reproduction group contradictions would more noticeable than the other adaptive groups. For the same reasons as for the previous prediction, the Reputation group was predicted to be the least noticeable. Out of a total of six possible rankings, an alternative prediction (Prediction 3b) was made that Reproduction will be the most noticeable, followed by Survival, with Reputation again, the last.

Chapter 3 – Study 1: Item Construction

In the contradiction paradigm (Lassonde, Surla, Buchanan, & O'Brien, 2012) or the inconsistency paradigm (Braasch, Rouet, Vibert, & Britt, 2012) participants are assessed for their ability to notice contradictions or anomalies between one line in a text and a later one. They are typically presented with lines one at a time on a monitor screen after pressing a keyboard key, and are assessed for having noticed a contradiction between two lines, either by answering questions after completing the passage (Keck, Kammerer, & Starauscheck, 2015; Otero & Kintsch, 1992) or by having their key depression times recorded and their reading or response times (RTs) calculated (Huitema, Dopkins, Klin, & Myers, 1993; Raap, 2008; Singer & Doering, 2014). Extended response times are interpreted as the contradiction or anomaly having been noticed. The latter variant was used in the current research, with all procedures employed approved by the Massey University Human Ethics Committee.

In Chapter 2 - Theoretical Background, it was predicted that people can be shown to be biased in their sensitivity to contradictions; they are more likely to notice adaptive contradictions, where adaptive pertains to survival, reproduction and reputation (SRR), than Non-Adaptive ones. In order to test these predictions using the contradiction paradigm, pairs of adaptive and Non-Adaptive contradictions had first to be created and tested for contradictoriness. Once this was achieved the pairs were embedded in short texts telling stories. These contradictory stories were subsequently intermixed with non-contradictory ones of identical format and used in two further studies to test the predictions.

Contradiction Construction

Method.

The adaptive topics chosen for the contradiction comparisons in the current research were based upon the evolutionary psychology literature (Buss, 2005) and are broad in scope. Seven topics were selected as representatively adaptive; two survival topics, Survival Life-Threatening and Survival Non-Life Threatening, three reproductive topics, Child-Loss, Family, and Fidelity and two reputation topics,

Honesty and Reciprocity. The Non-Adaptive ones were even broader and difficult "to pin down" (Brown and Yule, 1983, p. 68), for example, the Non-Adaptive Socially Unimportant topic. It was decided that within such broad groupings, a sub-topic would be selected which was semantically included within the main topic. Also, a single pair of contradictory sentences for each chosen topic or sub-topic was deemed insufficient to assess participants' sensitivity to contradictions pertaining to that topic. Three contradictory pairs of sentences were decided upon as providing adequate breadth, analogous to Huitema et al.'s (1993) use of three texts aimed at identifying anomalies similar to each other but with varying background material. As even a triplet of stories could in principle be very diverse in wording and meaning, it was decided that the stories, rather than only the contradictions, should be tested for a sufficient similarity to each other but different from those of other topics.

It was essential that the contradictions used in the topics' texts be of approximately equal salience. Different degrees of contradictoriness could operate as a confound when short texts or stories containing the pairs were read at a later stage. De Marneffe, Rafferty, and Manning (2008) produced a taxonomy of contradiction constructions. Three describe the constructions mostly used in the current research; those that employ antonyms in the sentence pairs, those that employ negation, for example, no, not, none, nobody, never, and similar, and number/date mismatches. These they classify as "easy" contradictions "because they can often be automatically detected without full sentence comprehension" (p. 1041). Their other categories are complex in the sense that substantial inferencing is required, something deliberately avoided in this research.

Despite the work of De Marneffe et al. (2008), there are no generally accepted methods used to measure contradictoriness nor any units. Mohammed, Dorr, Hirst, and Turney (2013) working in the area of machine understanding of texts, complain that "existing lexicons often do not classify opposites into different kinds" (p. 19) and "do not explicitly list word pairs that are not opposite yet have some degree of contrast in meaning." (p. 19). Santos, Lu, Lenci, and Huang (2014) agree that "currently, there is no effective measure to discriminate antonyms from synonyms because they share many common features" (p. 1). Mohammed et al. (2013) resorted to a crowd sourcing experiment to ascertain the degree to which people agreed on "the concept of

oppositeness and its different kinds" (p. 19). Their method of comparing words of apparent opposition is claimed to be an improvement on previous ones but cannot yet be said to be generally accepted.

Contradictory candidate pairs of sentences, largely employing the techniques described in De Marneffe et al. (2008), were generated for all topics. An example for the Socially Unimportant topic was,

"I love making papier-mâché from newspapers". and
"I hate making papier-mâché from newspapers".

Given the lack of general agreement on a definition of contradictoriness and units of measurement, the contradictoriness of candidate pairs was assessed using a process similar to that employed by Braasch, Rouet, Vibert, & Britt (2012), and Keck, Kammerer, and Starauschek (2015) who used Likert scales for assessment by panels of contradictoriness and reliability, respectively, of text pairs. It was discovered in Clemerson (2009) that judgements of contradictoriness may be made in relative terms. When a Likert scale of 1 – 7 is used to assess contradictoriness, a highly contradictory pair may be marked with a 7 and a subsequent pair that is judged to be only a little less contradictory is marked with a 1, even if participants have been requested to use the lowest mark only if they believe a pair not to be in any way contradictory. Candidate contradictory pairs were therefore mixed with a set of non-contradictory pairs in order to ensure that the lower end of the Likert scale was used only for what were judged to be non- or nearly non-contradictory pairs and the higher end of the Likert scale retained for those pairs judged to be highly contradictory. The sentence pairs were presented for assessment in batches to panels of 60 participants. As in Braasch et al. (2012), it was found that a few panelists appeared to have reverse scored the contradictions, giving a low mark to a sentence pair designed to be highly contradictory. To allow for this and individual differences in the assessments, only sentence pairs that were marked by the panels with an average score of 6 or more were accepted as sufficiently contradictory to be taken forward to the next stage of the research.

Participants.

In non-random convenience sampling, seven hundred and eighty participants, divided into 13 panels of 60 participants, were invited to assess pairs of sentences for contradictoriness. They were mostly undergraduate students at two university campuses in Wellington, usually recruited in groups of between 2 and 5. They were therefore mostly under 25 years of age and approximately evenly divided between males and females. Other participants were recruited as friends and associates of the researcher or as diners in shopping malls, similarly approximately divided between males and females. No more than 5 minutes of their time was requested which was almost universally provided.

Stimulus and Materials.

The stimulus consisted of A4 sheets of paper on which were printed a mix of up to thirty-four contradictory and non-contradictory pairs of sentences. The figures 1 to 7 were printed below each pair. An example sheet appears as Appendix C - Candidate Contradictory Pairs.

All pairs were composed to meet a presentation standard requiring both sentences to consist of between eleven and thirteen syllables and to be between six and eight cm long when printed in 12 point Times New Roman font. Although the length of the sentences was not important for this first stage, the standard was defined in order to meet a requirement that the lines of stories, in which some pairs were later to be embedded, all subtend approximately the same angle to participants' eyes and were easily read at a normal distance from the eye.

Procedure.

Participants, mostly in small groups, were approached and requested to participate at times when they were seen to be relaxing and not otherwise active. All were told that participation would require only a few minutes of their time. Very few refusals occurred. Participants were requested to read a pair of sentences and

immediately assess it for contradictoriness in the space below where the numbers 1 to 7 were printed and then progress to the next pair. They were asked to put a ring around the 1 only for sentences that they regarded as not in any way contradictory, put a ring around the 7 for those pairs that they assessed as completely or very contradictory and use the numbers in between according to their judgement for those pairs that were contradictory to some degree.

Each set of sentence pairs was assessed in this manner until 60 independent assessments had been collected. The marks were then transcribed to a spreadsheet and the average and standard deviation of the contradictoriness marks for each pair calculated. The SDs were calculated in order to assist in a later choice between pairs pertaining to the same topic which had also met the 6+ threshold; the smaller the SD, the more uniform the assessment. In the event of failure to meet the 6+ criterion, pairs were either slightly edited or abandoned, according to how near the acceptance threshold they were scored. In the former case, they were reassessed in a later round of testing with a separate and new panel of participants.

As three contradictory pairs of sentences were deemed necessary for each topic, a minimum of 27 contradictory pairs were needed to meet the 6+ threshold for the nine topics to be tested in the later stages of this research. However, the assessments continued beyond this point. More than three pairs meeting the 6+ threshold for each topic were sought. This permitted a later selection to be made based upon the SDs and a judgement of the degree to which they differed in meaning despite pertaining to the same topic. A total of 127 pairs were assessed via the 13 panels of 60 participants, with 61 pairs meeting the threshold. The 27 pairs selected combined higher contradictoriness scores and lower standard deviations, and within each topic the greatest diversity of meaning, as it had been decided that the stories would later be tested for sufficient similarity.

Story Structure.

There appears to be no consistency among contradictory paradigm researchers in the way that the discrepant sentences are placed in texts. Huitema et al. (1993) separated anomalous sentences by a variable number of lines up to a maximum of 5. Otero and

Kintsch (1992) maintained a fixed separation of 3 lines, although their sentences were longer than those employed in the current research. Rapp and Kendeou (2007) placed their discrepant sentences sequentially. In the current research twenty-seven stories of 11 lines each were generated into which the twenty-seven selected contradictory pairs were embedded as lines 3 and 10. Placing 6 intermediate lines between contradictory ones, rather than a maximum of 5 as in Huitema et al. (1993) was judged a more stringent test, one that topics such as survival and reproduction should be able to meet.

Results

The 27 contradictory pairs chosen for inclusion in the stories appear in Table 2 below.

Table 2

Contradictory Pairs - Marks and SDs

Topic Group		Topic	Sentence Pairs	Mean Score	SD
Adaptive	Survival	Non-Life Threatening	Finding warm shelter on cold nights will be difficult. Finding warm shelter on cold nights will be easy.	6.70	0.78
			Our survival is dependent upon good shelter. We'll survive perfectly well without shelter	6.15	1.09
			Our health will suffer without warm clothes in winter. We don't need warm clothing to stay healthy in winter.	6.16	1.24
		Life Threatening	Not even once was my life ever in danger. Many times, my life was in very great danger.	6.38	1.26
			I am unconcerned about losing my life I live in perpetual fear for my life	6.29	0.96
			I was frightened for my life throughout the journey Never during the journey did I fear for my life	6.79	0.76
	Reproduction	Child-Loss	I grieved a long time after the death of our child. I didn't grieve at all after the death of our child.	6.64	1.50
			I will quickly recover from my child's death I will never recover from my child's death	6.58	0.95
			Our first child died suddenly when still a baby We've not suffered the loss of any of our children	6.37	1.38
		Fidelity	I felt wretched when my partner had affairs. My partner's many affairs did not upset me.	6.49	1.18

		I've had no sexual affairs since I got married. I've had eight sexual affairs since I got married.	6.41	1.39
		My partner's affairs cause me no jealousy. My partners' affairs make me intensely jealous.	6.52	1.37
	Family	We look forward to having several children. We have decided not to have any children.	6.23	1.43
		We are firm in not wanting any children. We are both excited about having children.	6.08	1.28
		Our children will be my life's greatest purpose. Our children will provide no purpose to my life.	6.69	0.96
Reputation	Reciprocity	People remember whether you return favours People don't remember whether you return favours	6.29	1.41
		Returning favours ensures one a good name Returning favours won't ensure one a good name.,	6.52	1.16
		I always return kindness from others with kindness. I don't return kindness from others with kindness.	6.28	0.94
	Honesty	Cheating greatly affects our social standing. Cheating has no effect on our social standing	6.20	1.23
		I really mind if I'm accused of dishonesty. I don't mind if I'm accused of dishonesty.	6.79	0.77
		It matters a lot if people think I'm a crook. It doesn't matter if people think I'm a crook.	6.09	1.43

Non-Adaptive	Socially Important	I got my second parking ticket yesterday I've never had a parking ticket issued to me	6.30	1.48
		I got issued with a speeding ticket yesterday I have never been issued with a speeding ticket	6.56	1.13
		It's OK to drive through a red light if there's no traffic. People should never drive through a red traffic light.	6.61	0.84
	Socially Unimportant	The striped paper bag is full to the top. The striped paper bag is completely empty.	6.60	1.34
		I love making papier-mâché from newspapers. I hate making papier-mâché from newspapers.	6.60	1.21
		Sebastian had lots of wrapping paper. Sebastian had no wrapping paper at all.	6.75	0.90

Story Construction.

Twenty-seven stories were subsequently generated into which the twenty-seven selected contradictory pairs were embedded as lines 3 and 10. It was appreciated that various features of lines 1, 2, and 4 - 9 of the stories could affect the noticeability of a contradiction in line 10. Sudden changes of tense, mood, or theme could affect participants' likelihood, both positively or negatively, of noticing a contradiction. As in Otero and Kintsch's (1992), the stories present a progression in accordance with rules extracted from an examination of the Discourse Comprehension literature. A précis of this literature and the rules that were extracted from it appear in Appendix A - Discourse Comprehension. These rules were applied both at the time of composition of the twenty-seven contradictory stories and in subsequent peer reviews, the latter leading to some amendments; no previously selected stories were abandoned. The changes were made only to the lines in the 4 - 9 range. Six non-contradictory stories were generated, meeting the same presentation standard and are listed in Appendix D - Non-Contradictory Stories.

Discussion

The difficulty in generating pairs that met the 6+ threshold exceeded expectations. The 61 pairs that met the threshold required a total of 127 to be assessed. Many pairs were visually almost identical. The pair below met the threshold with a mark of 6.79.

I really mind if I'm accused of dishonesty.

I don't mind if I'm accused of dishonesty.

In contrast, the pair below obtained a mark of only 5.64.

We will have a family of several children.

We have decided not to have any children.

Despite the difference in appearance of the two pairs above, it was sometimes difficult to ascertain any very clear reasons why some pairs obtained a score of 6+ and others did not. Nevertheless, a subjective and detailed analysis appears in

Appendix E – Contradictory Pair Analysis, from which the rules below were extracted.

1. Ensure that the two contradictory sentences are very similar in syntax and appearance or form, particularly the second half. One method of achieving this is by employing the ‘easy’ constructions listed by De Marneffe et al. (2008).
2. If the contradictoriness is conveyed via adjectives, use only well recognised antonyms.
3. Despite the above, some dissimilarity of appearance can be tolerated if the sentences have high emotional valence.
4. Avoid the abstract.
5. Avoid more than one verb in the sentences.
6. Avoid opposing adjectival or adverbial phrases which imply contradictoriness rather than being explicitly contradictory.

Chapter 4 - Study 2: Story Testing

In Chapter 3 - Item Construction, it was mentioned that the triplets of stories for each topic would be tested to ensure that although exhibiting some diversity of meaning, they were sufficiently similar to each other and illustrative of the relevant topic and not some other. In this chapter the testing of the stories composed from the 27 selected contradictions is described. The stories were each composed (1) to conform to a presentation standard in order to be superficially indistinguishable from each other in appearance and (2) to have the greatest diversity of meaning while still being sufficiently similar to each other not to be regarded as a story pertaining to a 'wrong' topic.

Method

The assessment of the nine topic's triplets of stories was conducted via a closed card sorting test. The trays for the test were labelled either with the original topic name or a close paraphrase which reflected more precisely the sub-topic or theme of the stories. The participants were required to match each story to one of the labels of the trays. A threshold of not less than 90% correct placements across all stories and participants was adopted as high enough to consider the test a sufficiently good one of stories' similarity to allow them to progress to the next stage unchanged. Landis and Koch (1977) developed a family of Kappa-type statistics based upon Fleiss' (1971) adaptation of Cohen's (1960) Kappa Measure of Agreement. Landis and Koch (1977) suggested that the result of a Kappa-type calculation in the range of 81 – 100% should be regarded as "Almost Perfect" (p. 165). The detailed description of Landis and Koch's (1977) adaptation to the tray sorting test adopted for the current research appears as Appendix F - Tray Sorting 90% Correct Placement. Tullis and Wood (2004) note a diminishing return when employing card sorting tests to ascertain similarities as increasing numbers of participants are employed. After a test with 168 participants, they noted that in subsequent simulations, 20 to 30 participants gave results very close to those obtained for the 168. In the interests of caution, a figure of 30 participants rather than 20 was adopted for this study's sorting test. The number of stories, 27, and topics,

9, both meet Rugg and McGeorge's (1997) suggested maximum of 30 instances of "entities" to be assessed via card sorts.

It was also decided that the stories would be assessed for interest as there might be a high correlation between the noticeability of a contradiction and the degree of interest in the relevant topic. Although degrees of interest were not formally defined as objectives of the current research, it was predicted that a greater interest in the adaptive topics would be found than in the Non-Adaptive ones. It was conjectured that the evolutionary considerations that led to a prediction of bias in the noticeability of contradictions would also lead to a predictable bias in matters of interest (Carroll, 2004; Dutton 2009).

Participants

Thirty participants (17 females) were requested to participate in the test. Most were undergraduates, recruited on university premises, under the age of 25. Others were friends and acquaintances of the researcher.

Materials

The 27 stories were printed on small sheets of paper approximately 14 cm by 9 cm. Nine small paper trays, each labelled as in Table 3 below, were used into which participants were required to place the story sheets.

Table 3

Topics and Tray Labels used for the Tray Sorting Test

Topic Category	Topic	Tray label
Survival	Life Threatening	Life Threats
	Non-Life Threatening	Exposure
Reproduction	Family	Family
	Child Loss	Child Loss
	Infidelity	Unfaithfulness
Reputation	Honesty	Untrustworthiness
	Reciprocity	Reciprocity
Non-Adaptive	Socially Important	Driving Rules
	Socially Unimportant	Paper things

While some trays were labelled with the name of the topic because this was judged to be precise and a relatively close match to the stories, for example Life Threats, others were labelled with the sub-topic that was chosen at the time of contradiction composition and testing or a theme common to the stories as aids to understanding. Paper Things as a sub-topic of the Socially Unimportant topic is an example. While this may have been sensible in that particular case, in others cases, not using the label of the topic itself, such as in the case of Honesty, may have proved to be the opposite to some degree of what was intended, as described in the Results section below. A photograph of the trays and printed story sheets appears as Appendix G - Tray Sort.

Procedure

The nine labelled paper trays were placed in front of the participants who were then provided with a well shuffled pack of the 27 contradictory stories. Participants were instructed to read each story and identify a topic or theme from it. They were then to place the sheet into the labelled tray which, in their opinion, best matched the topic or theme of the story. They were informed that all the stories contained contradictions which were to be ignored as irrelevant to the purpose of this experiment. Participants were informed that as there were twenty-seven stories and nine trays, at the end of the

exercise, each tray should contain three stories. They were advised that if at any stage during the readings and placings, they wanted to revise a placement, they were free to do so. As the participants made their placements, the researcher remained in the vicinity to answer questions, but at a distance. On completion of each participants' placings, their placements were assessed for correctness.

A simple method of assessing interest was chosen in order to avoid the difficulty for participants of having to sequence all nine topics and the lengthy time that could be required. At the completion of their placement activity, twenty-four participants were requested to say which 3 of the 9 topics they regarded as of highest interest and which 3 topics they regarded as of lowest interest, the remaining 3 topics to be considered of medium interest. For each participant, the topics were awarded marks as shown below;
3 for each of the three topics of highest interest,
2 for each of the three topics of medium interest,
1 for each of the three topics of lowest interest.
The marks were subsequently totalled to arrive at an indication of relative interest across all topics and all participants.

Results

Categorisation Errors Rates.

As the panel of convenience numbered 30 people, a total of 30 X 27 (= 810) placements had to be assessed. Seventeen participants placed all 27 stories in the correct trays. The total number of misplacements by the other thirteen was 33, giving an error rate of 4%, a correct placement rate of 96%. One Reputation/Honesty/Untrustworthiness story which focused on cheating was placed in the Reputation/Infidelity/Unfaithfulness tray 5 times, one Reproduction/Fidelity story was placed in Reputation/Honesty/Untrustworthiness tray 4 times and another was placed in that tray twice. For each of these 11 misplacements there was of course a paired misplacement but assessment of these revealed no particular pattern. These 22 misplacements thus accounted for the majority. There was no pattern to the remaining misplacements. The cause of the cheating misplacements probably lies in the ambiguity

of the word ‘cheating’. In an early Item Construction trial, one of the contradictions contained the word ‘swindling’, intended to convey issues relating to business transactions. Some participants asked what ‘swindling’ meant. In subsequent trials, the word ‘cheating’ was substituted. However, the meaning of ‘cheating’ as commonly referring to adultery in American films has been adopted in New Zealand and may have led to the story being interpreted as about infidelity. Two other stories which included the word ‘affairs’, intended to convey infidelity, were correctly interpreted by 26 and 28 participants but the other 4 and 2 participants respectively, understandably if inaccurately, regarded them as conveying untrustworthiness. The substitution of Unfaithfulness and Untrustworthiness for Fidelity and Honesty respectively, intended to improve matching precision, probably had the opposite effect of lowering the correct placement tally. As the 90% threshold had been well exceeded, no amendment was made to any of the stories.

Interest Assessment.

Tables 4 and 5 below present the results of the expressions of interest.

Table 4

Survival		Reproduction			Reputation		Non-Adapative	
Life Threats	Exposure	Family	Child-Loss	Infidelity	Honesty	Reciprocity	Driving Rules	Paper things
2.21	1.75	2.29	2.46	2.42	2.42	2.25	1	1.21

Mean interest score awarded to each topic

The mean scores awarded to each topic category are shown in Table 5 below.

Table 5

Mean of interest scores awarded per participant to each topic category

Topic Category	Mean of scores awarded
Survival	1.98
Reproduction	2.39
Reputation	2.34
Non-Adaptive	1.11

Discussion

The “almost perfect” pass mark of 81% to 100% (Landis and Koch, 1977) straddled the figure of 90% originally adopted for the sorting test. The high correct placement rate of 96% actually achieved is therefore a gratifying result. However, it might have been higher had the original topic names of Honesty and Fidelity been retained.

No formal predictions had been made pertaining to the assessment of relative interest of the four topic groups. However, prior to the test, an expectation had been voiced that the Non-Adaptive topics would emerge as those of lowest interest. This expectation was met. Nevertheless, within the adaptive groups one unexpected result was obtained. The Survival scores were lower than the other adaptive ones, against expectations, due to the low Exposure marks. However, the greater interest expressed in the Life Threatening stories relative to the Exposure ones supported the prediction made in the Theoretical Background that life threatening contradictions would be more noticeable than non-life threatening ones, which carried a survival threat only by implication. The Child-Loss topic was awarded the highest average mark in the Reproduction category, supporting the prediction made in the Theoretical Background that the Child-Loss contradictions would be the most noticeable in the Reproduction Group. The Honesty topic was found to be of higher interest than Reputation, supporting the prediction made in the Theoretical Background that contradictions pertaining to honesty would be the more noticeable in this Reputation Group. The average interest marks awarded to the Non-Adaptive topics were both substantially less than the Adaptive categories, as noted above, but the higher marks awarded to the Socially Unimportant topic relative to the Socially Important one was against expectation.

As the assessment of interest was only a small scale incidental and relatively unimportant part of the research, the results do not warrant great attention. However, the large differences in the marks awarded to the adaptive topics relative to those awarded to the Non-Adaptive ones, does suggest an avenue for future research. A large scale

quantitative study of the correlation between a topic's adaptive relevance and its interestingness may vindicate the conjecture made earlier in this chapter.

The general results of the tray sorting tests are interpreted as indicating that the stories for each topic are semantically diverse within their topic but sufficiently similar to each other not to be thought of as representing a different topic. As a consequence of the scores obtained from the test and the judgement based upon them, the stories were taken forward to Study 3. This study commenced the testing of the predictions made in Chapter 2 - Theoretical Background

Chapter 5 - Study 3: Prediction Testing

Study 3 was designed to test some of the predictions made in Chapter 2 - Theoretical Background. The general hypothesis underlying the current research is that adaptive contradictions are more important and will therefore be more noticeable than Non-Adaptive ones. Three groups of adaptive contradictions have been identified: Survival, Reproduction and Reputation. Study 3 was therefore aimed at testing the main underlying hypothesis through comparisons of topics from all three of these adaptive groups with the Non-Adaptive group. The Non-Adaptive group, comprising the Socially Important and Socially Unimportant topics, provided the baseline with which the adaptive topics were compared. The Survival Group included both Life Threatening and Non-Life Threatening Survival topics. This enabled the testing of one of the most important predictions of the current research, Prediction 1a, comparing the Survival Group with the Non-Adaptive group. The inclusion of both Survival topics enabled them to be compared with each other, thereby assessing Prediction 2a. Comparisons of the adaptive groups, Reproduction and Reputation, were also made with the Non-Adaptive Group, with the Child-Loss topic chosen from the Reproduction Group on the grounds that such loss is a devastating experience with long term consequences (Murphy, 2008) and sensitivity to contradictions is likely to be high. From the Reputation Group, Honesty was chosen as a characteristic fundamental to the formation of trust and cooperation with others (Rubin & Somanathan, 1998; Somanathan & Rubin, 2004) and likely to contrast significantly with the Non-Adaptive topics.

Method

The experiment employed the contradiction paradigm (Lassonde, Surla, Buchanan, & O'Brien, 2012). The method provides an experimental means or instrument to detect whether an anomaly has been noticed. Participants were presented with contradictory statements within a passage of text. In the current research, the instrument was longer reading times for lines contradicting an earlier one. As is common in contradiction paradigm experiments, texts containing contradictions were interspersed with other texts which contained none (Singer & Doering, 2014). As in

Egidi and Gerrig (2006), the non-contradictory texts were used "to disguise the experiment's purpose and allow counterbalancing" (p. 1324).

Contradictory paradigm experiments are based upon three assumptions:
(1) the contradictory lines give rise to longer response times (RTs) than the non-contradictory ones,
(2) the longer RTs are due to the participants detecting the contradiction and,
(3) the participants may feel cognitive dissonance as a consequence.
The data collected for the current research permitted the testing of these assumptions. An analysis of the validity of these assumptions appears in Appendix H - Contradictory Paradigm Assumption Testing.

Participants.

A priori calculations for participant numbers using the software tool G*Power for ANOVA Repeated Measures with a small effect size $f = .1$, $\alpha = .05$, Power = .8, number of measurements = 6 (the number of contradictory stories read by each participant) indicated that over 100 participants would be required to obtain statistically significant results. This assessment can be set against practical cases. A rich literature exists of research covering the contradiction paradigm. It was used as a guide to establish participant numbers for the current research (McCready, 1996). These have varied widely, from as few as 24 (Braasch et al., 2012; Putnam, Wahlheim, & Jacoby, 2014) to as many as 161 (Keck et al., 2015). Where statistically significant effect size calculations have been presented, they have varied from small with *partial* $\eta^2 = .028$, (Keck et al. 2015), to very large with Cohen's $d = 1.13$ (Braasch et al. 2012) and *partial* $\eta^2 = .16$ (Rapp, Hinze, Slaten, & Horton, 2014). Nairne and Pandeirada (2007, 2010), whose work partly motivated the current research, used a different experimental paradigm, namely, testing memory for survival related words among word lists. With participant numbers at 150 (Nairne & Pandeirada, 2007) and 120 (Nairne & Pandeirada, 2010) statistically significant large effect sizes were obtained. In the current experiment, using the recruitment methods described below, it proved practical to recruit 201 participants, (105 females), more than the experimenters cited above, so that even if only small effect sizes were obtained, the results should be statistically significant.

Most recruitment took place at the campuses of Massey University in Wellington and Palmerston North and the campus of Victoria University of Wellington where volunteer recruits were mainly undergraduates. The managers of two Wellington shopping malls permitted additional participants to be recruited in their food courts. Most mall participants were from the top half of the education spectrum, as indicated by their achieved or current educational level (see below). Other participants were friends and acquaintances of the author.

Participants provided anonymous profile data: sex/gender, their mother tongue (English or other), continent of childhood and highest educational attainment. Sex/gender was captured as it was considered that evolution-oriented research may reveal differences between the two sexes. A rich literature exists proposing evolutionary reasons for differences between some psychological characteristics of the two sexes (Buss, 1989, 2007; Haselton, Buss, Oubaid, & Angleitner, 2005; Schmit & Buss, 1996) with some exploring a wide range of cognitive differences, for example, McGillicuddy-De Lisi & De Lisi (2002). The CD literature typically cites the number of participants by sex, apparently to indicate a rough balance, but studies that demonstrate a sex bias have not been found. As the CD literature has not proposed theoretical reasons why differential sensitivity to contradictions according to sex should exist, no precise predictions were made. Mother tongue - English or other - and first continent of residence were captured with a view to ascertaining whether the relevant participants' RTs were significantly different or not, their data to be analysed separately if substantial anomalies were detected.

Stimuli.

The stimuli were the eighteen stories, described in Chapter 4 - Study 2: Story Testing, pertaining to the six contradictory topics included in this Study (three stories per topic), and the six non-contradictory stories. The contradictory and non-contradictory stories conformed to the same format. Each participant was presented with a particular subset of 12 stories: the first was always non-contradictory and was followed by an alternating sequence of contradictory and non-contradictory ones. As a precaution against order effects, the position in which any one topic occurred in the

sequence of 12 stories was varied, six topic sequences being required to ensure that each topic occurred in each contradictory story position, a practice similar to that of Braasch et al. (2012). As three stories were generated for each topic, 18 different sequences of stories were required to ensure that all three stories for all six topics were displayed with equal frequency in each position. The concept, which is largely based upon the principle of the Latin Square, is more fully explained and illustrated in Appendix I – Topic Sequences. The full eighteen sequences for Study 3 appear in Appendix J - Eighteen Versions. All contradictory stories appear as Appendix K – Contradictory Stories.

Materials.

The software written for this research was similar in design to WebExp2 (Mayo, Corley, & Keller, 2005), a software tool for collecting timing data over the internet. In this experiment the main functions of the software were to display successively lines of stories as participants pressed the space bar of a keyboard and to record and store the times when they did so. The times that participants spent reading each line, their response times (RTs), were calculated by subtraction of one recorded time from the next. Although WebExp2 provides for stimuli to be presented in many formats, this experiment's software limited it to written text. Times were captured by JavaScript's getTime() function when displaying the first line of a story and again when the operating system detected the depression of the space bar. Inaccuracies with a mean on the order of 10 - 15 ms can occur due to variations in the monitor refresh times relative to key depressions and keyboard latencies (Plant & Turner, 2009). Typical RTs in this experiment ranged from 1,000 to 7,000 ms. The Means of participants' RTs are over two orders of magnitude larger than machine-introduced errors that are themselves averaged as part of the calculation of the RT means. Consequently, machine-introduced errors were assessed as too small to affect the results reported for this research. A replicated feature of WebExp2, its versions, allowed the eighteen different sequences of stories to be presented to successive participants. On completion of the experiment, a Comma Separated Variable formatted file of participants' RTs, together with date, time and personal profile data, was created and stored in the application's database for subsequent processing and analysis. The architecture of the software consisted of the application software written in Javascript. Server side software, managing the versions and results files, was written in Java and ran under the control of the open source server

and servlet container Apache Tomcat system, version 6.0.18. Participants' data were transferred to the author's PC for analysis via a memory stick. SPSS version 23 and Microsoft Excel 2016 software were used to calculate the Results.

Procedure.

On commencing, each participant first read instructions, (see Appendix L - Instructions) in which they were informed that the experiment had been approved by the relevant Massey University Ethics Committee and that they would be reading twelve short stories of eleven lines each. The story reading would be preceded by a short practice story and questions. In all cases, they were to read a line and press the space bar once they had finished reading that line. The program would then display the next line and so on until the end. Following each story, a page containing four simple questions about that story would be displayed, which were to be answered by typing very short answers, many with a yes (Y) or no (N). They were informed that they would be required to provide some anonymous profile data and that providing their email addresses was optional. If they chose to do so, they would later receive an explanation of the objectives and results of the experiments. They were also told that they may cease to participate at any time and that their informed consent would be assumed if they chose to advance to the practice story, which repeated the instructions.

Once participants had read these instructions, they pressed the space bar to provide the profile data. Thereafter, they proceeded to read the stories and answer the questions to completion. Afterwards, participants were given a brief description of the objectives of the research. All participants who supplied their email addresses have been sent an explanation of the objectives of the research and a summary of the results.

Results

Initial Data Analysis.

Normality Tests and Descriptive Statistics.

Homogeneity of Variance was tested by application of the F_{Max} test, the ratio of the largest topic variance to the smallest not exceeding 10 (Tabachnick & Fidel, 2013). The value of F_{Max} for Study 3 RTs is 2.07. Normality checks revealed that Kurtosis values were more than twice the Standard Error (Rose, Spinks, & Canhoto, 2014) indicating high Kurtosis. The RT data were positively skewed, with the Kolmogorov-Smirnov test showing significance levels under .05 for all topics. However, the Standard Error of Skewness declines to small values with large samples[5] undermining the utility of the test (Rose, et al. 2014). Descriptive Statistics appear in Table 6 below.

[5] The standard error of skewness is roughly inversely proportional to the square root of the sample size as indicated by its definition, $(6n(n-1)/(n-2)(n+1)(n+3))^{1/2}$ which approximates to $n^{-1/2}$ when n is large.

Table 6

Descriptive Statistics for Study 3 RTs - all participants. All measurements are in ms.

Topic	N	Mean	SD	95 % C.I.		Min	Max	Skewness		Kurtosis	
				Lower	Upper			Statistic	Std. Error	Statistic	Std. Error
Surv_Life_Thr	201	3096	1625	2826	3366	743	15809	2.411	0.172	3.112	0.341
Suurv_Non_Life_Thr	201	3269	2098	2977	3561	375	15317	2.588	0.172	9.608	0.341
Child-Loss	201	3613	2250	3300	3926	874	21118	3.31	0.172	19.256	0.341
Honesty	201	3062	1961	2789	3334	687	15491	2.972	0.172	13.953	0.341
Soc_Imp	201	2821	1681	2587	3054	597	14435	2.573	0.172	11.764	0.341
Soc_Unimp	201	2893	1564	2676	3111	606	15166	2.788	0.172	17.948	0.341

Note: Soc_Imp = Socially Important, Soc_Unimp = Socially Unimportant, Surv_Life_Thr = Survival Life Threatening, Surv_Non_Life_Thr = Survival Non-Life Threatening.

First Language Analysis.

To ascertain whether there was sufficient difference in RTs to justify separate analyses for English First Language and Non-English First Language participants, between-groups *t*-tests were conducted comparing the Line 10 RTs for both categories of participants for all topics. Descriptive Statistics and the results of the *t*-tests appear in Table 7 below.

Table 7
Comparisons of English First Language Participants' RTs with Non-English Participants' RTs for all Study 3 topics. Means and SDs are in ms.

Topic	Language	N	Mean	SD	Sig (2-tailed)	*t*	*df*
Child-Loss	English	190	3610	2260	0.943	0.072	199
	Other	11	3660	2187			
Honesty	English	190	2992	1785	0.305	1.078	10.25
	Other	11	4260	3872			
Soc_Imp	English	190	2814	1690	0.827	0.219	199
	Other	11	2929	1584			
Soc_Unimp	English	190	2787	1275	0.118	1.709	10.13
	Other	11	4734	3769			
Surv_Life_Thr	English	190	3058	1928	0.242	1.174	199
	Other	11	3763	2081			
Surv_Non_Life_Thr	English	190	3179	1929	0.192	1.398	10.29
	Other	11	4819	3863			

Note: Soc_Imp = Socially Important, Soc_Unimp = Socially Unimportant, Surv_Life_Thr = Survival Life Threatening, Surv_Non_Life_Thr = Survival Non-Life Threatening.

None of the *t*-tests reached conventional statistical significance. The 'Other' participants' data were therefore not removed from the main body of data for separate analysis.

Continent of Childhood Analysis.

To ascertain whether participants' continent of childhood acted as an independent variable, a one way between-groups ANOVA for all topics was conducted with continent of childhood as the independent factor. There was no statistically

significant difference in the RTs for any continent of origin for any topic. No participant data were therefore removed from the main body of data for separate analysis.

Education Level Analysis.

Prior to reading the stories, participants were asked to report their educational achievement to date according to whether they had achieved one of three educational levels, secondary, graduate or postgraduate, which were explained for participants during data collection and for analysis purposes as:

1. Secondary: completed secondary education but either had no tertiary education or had not yet completed first year of tertiary education,

2. Graduate: completed first year university education and could mostly be expected to graduate,

3. Postgraduate: had a postgraduate qualification or were currently engaged in post-graduate education.

The Descriptive Statistics appear in Table 8 below.

Table 8
Study 3 Descriptive Statistics for Mean Line 10 RTs according to Educational Status. All times in ms.

Education Level	N	Mean	SD
Secondary	67	3,300	2,121
Graduate	112	2,996	1,828
Postgraduate	22	3,353	1,911

A one way between-groups ANOVA was conducted to assess whether there were any statistically significant differences in mean Line 10 RTs for the three levels of education for each topic. No statistically significant difference was found. No further analyses of Study 3 participant data separated by educational level were conducted.

Outlier Policy and Application.

Once participants had commenced reading the stories, they mostly proceeded without interruption or diversion to completion. Occasionally, participants would be interrupted by external factors, ask questions or their attention seemed to wander. Ratcliffe (1993) describes outliers as resulting "from spurious events such as loss of attention, daydreaming, distraction and so forth" (p. 511). Hawkins (1980) defines an outlier "as an observation that deviates so much from other observations as to arouse suspicion that it was generated by a different mechanism." (p. 1). Accordingly, such observations have been made the basis of an outlier policy as extended RTs have been interpreted as unrelated purely to participants' reactions to contradictory lines. Tabachnick and Fidell (2013) describe an outlier as "not a member of the population from which you intended to sample" (p. 74). Tabachnick and Fidell (2013) additionally view an outlier as "a case (or a very few cases) that seems to be unattached to the rest of the distribution" (p. 74). Inspection of the Study 3 histograms in Appendix M - Histograms Illustrating Outliers typically reveals some unattached data points, in the worst cases with RTs in the range of 10 to 22 seconds. These RTs extend much further from the "pileup of cases near the mean" (Tabachnick & Fidell, 2013, p. 74) supporting the view that another factor is responsible for these extremes, a 'contaminant' (Hawkins, 1980). The histograms show all topics' RTs to be positively skewed, meeting an expectation for response times for this type of experiment (Ratcliffe, 1993; Roberts & Russo, 1999, p. 80). Illustrative histograms appear in Figure 4 below.

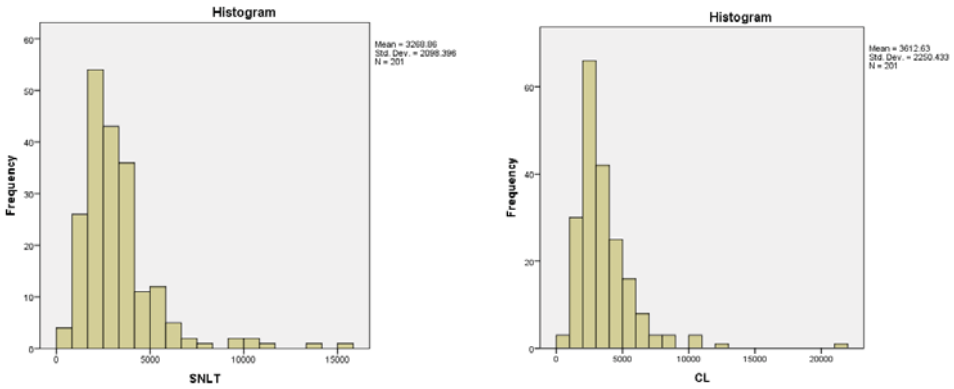


Figure 4. Sample histograms illustrating outliers and the +ve skew of RTs.
Note: SNLT = Survival Non-Life Threatening, CL = Child-Loss.

A large literature has been developed for the identification and treatment of outliers for a very diverse range of applications (e.g., Barnett & Lewis, 1995; Ben-Gal, 2010; Hawkins, 1980; Mosteller & Parunak, 1985). No universally applicable method has emerged. For example, in experiments analogous to the current one, Lassonde et al. (2012) and Rapp and Kendeou (2007) rejected all data points beyond 2.5 standard deviations from the mean. Huitema et al. (1993) rejected all RTs in excess of 4,000 ms. No explanation for any particular cut-offs were given. Baayan and Milin (2010) describe a-priori screening for outliers as a "widely accepted practice in by-subject and by-item analysis. It simply removes all observations that are at a distance of more than two standard deviations from the mean of the distribution" (p. 16). Histograms of Study 3 topics' data excluding outliers defined as those greater than the Mean + 2 SD, Mean + 2.5 SD and Mean + 3 SD showed that most of the unattached data points were eliminated if outliers were defined as data points with values equal to or greater than the Mean + 2 SD. The full set of histograms illustrating the effect of excluding outliers so defined appears in Appendix M - Histograms Illustrating the Effect of Outlier Exclusion. This cautious approach is supported by Roberts and Russo (1999) who state "Any scores which are greater than the mean plus two standard deviations or the mean minus two standard deviations may be deleted" (p. 80). Cardinal and Aitken, (2006, p. 67) also acknowledge the frequency of this approach. Stadtler, Scharrer, Skodzik, and Bromme (2014) adopted this practice in a contradictory paradigm experiment similar to the current research. Outliers have therefore been defined and eliminated according to this recommendation, which resulted in a loss of 3.9 % of the data, well below Ratcliffe's (1993) suggested maximum of 5%. It should be noted that a loss of a data point for one topic for a particular participant did not necessarily result in the loss of the remaining data points for that participant. However, some of the predictions in the Theoretical Background require planned comparison analyses which include several topics' RTs. For example, Prediction 1a required an analysis comparing both Survival topics' RTs with both the Non-Adaptive topics' RTs. In such cases where one or more of the relevant values were missing, listwise deletion did eliminate that participant's data from this comparison. However, other comparisons not using a missing data point were not affected. The analyses reported in the next section testing the predictions made in Chapter 2 - Theoretical Background have been conducted on data from which outliers as defined above have been eliminated.

Test of Normality and Descriptive Statistics of Participant Data Excluding Outliers.

Once outliers had been eliminated, the residual data were assessed for normality and 95% Confidence Interval calculated. Descriptive Statistics appear in Table 9 below and the 95% confidence intervals are illustrated in Figure 5 below Table 9.

Table 9

Study 3 Descriptive Statistics. All times in ms.

Topic	N	Mean	SD.	95 % C.I.		Min.	Max.	Skewness		Kurtosis	
				Lower	Upper			Statistic	Std. Error	Statistic	Std. Error
SLT	192	2810	1337	2619	3000	743	6850	0.895	0.176	0.25	0.349
SNLT	193	2949	1335	2759	3138	375	7183	0.796	0.175	0.42	0.348
CL	193	3295	1435	3091	3498	874	7549	0.798	0.175	0.12	0.348
H	193	2782	1273	2601	2963	687	6786	0.959	0.175	0.53	0.348
SI	192	2575	1157	2410	2739	597	6146	0.951	0.175	0.61	0.349
SU	196	2759	1209	2588	2929	606	5905	0.585	0.174	-0.31	0.346

Note: SLT = Survival Life Threatening, SNLT = Survival Non-Life Threatening, CL = Child loss, H = Honesty, SI = Socially Important, SU = Socially Unimportant

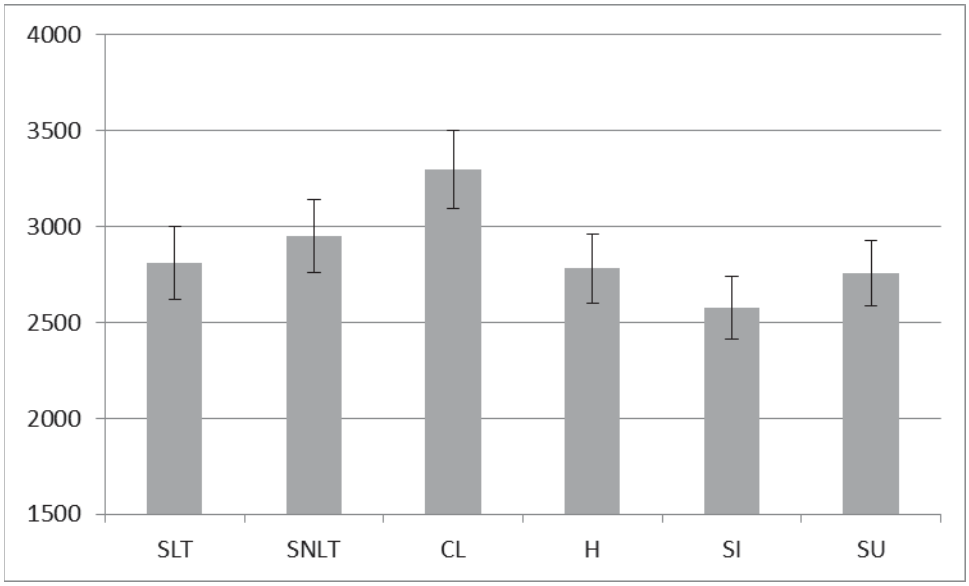


Figure 5. Histogram of Means of the Line 10 RTs with 95% Confidence Intervals

Note: SLT = Survival Life Threatening, SNLT = Survival Non-Life Threatening, CL = Child-Loss, H = Honesty, SI = Socially Important, SU = Socially Unimportant.

Tests of Sex/Gender.

It was anticipated that male and female sensitivities to contradictions could vary according to the topic although no particular predictions were made. To ascertain whether any such difference exists, an independent samples *t*-test was conducted for all topics comparing the Line 10 RTs of male participants with those of female participants. The results appear in Table 10 below.

Table 10
Descriptive Statistics and t-test results for Study 3 Line 10 RTs for male and female participants for all Topics. All times in ms.

	Sex	N	Mean	SD	Means diff.	t	Sig (2 Tailed)	*Partial η2*	
SLT	Female	104	2562	1186	541	-2.796	.006	.04	Small to moderate
	Male	88	3102	1450					
SNLT	Female	103	2673	1066	591	-3.062	.003	.05	Small to moderate
	Male	90	3264	1535					
CL	Female	104	3117	1303	384	-1.866	.064	.02	Small
	Male	89	3502	1558					
H	Female	102	2686	1174	204	-1.11	.268	Not applicable	
	Male	91	2890	1374					
SI	Female	102	2412	1138	348	-2.095	.037	.02	Small
	Male	90	2759	1157					
SU	Female	102	2627	1206	275	-1.595	.112	Not Applicable	
	Male	94	2902	1203					
			Mean of Diffs		390				

Note: SLT = Survival Life Threatening, SNLT = Survival Non-Life Threatening, CL = Child loss, H = Honesty, SI = Socially Important, SU = Socially Unimportant.

Male RTs were longer than female ones for all topics. The differences for both Survival topics and the Socially Important topic reached conventional statistical significance while that of Child-Loss approached it. Some of the effect sizes approached moderate size. This research is based upon the assumption that longer Line 10 RTs indicate the detection of contradictions in stories. One interpretation of the results above is that males generally detect all topic contradictions more readily than females. However, faster female reading speeds may be a factor (Berman & Bird, 1933) although the issue is contested (Moore, 1940). **T**he varying differences between the two sexes RTs' according to topic suggested a sex-based differential sensitivity to contradictions according to topic. One sex may be more sensitive to contradictions than the other for one topic with the differential sensitively diverging for another. A two way between-groups ANOVA to explore the impact of sex and story type on RTs showed no interaction effect when all six stories were assessed together, p = .653. However, a two way ANOVA conducted with just the Survival Non-Life Threatening and Honesty topics, the two showing the greatest and least difference in sex-based means, showed a greater interaction, F (2, 382) = 2.165, p = .142, although not meeting conventional significance. Figure 6 below shows the relevant plots.

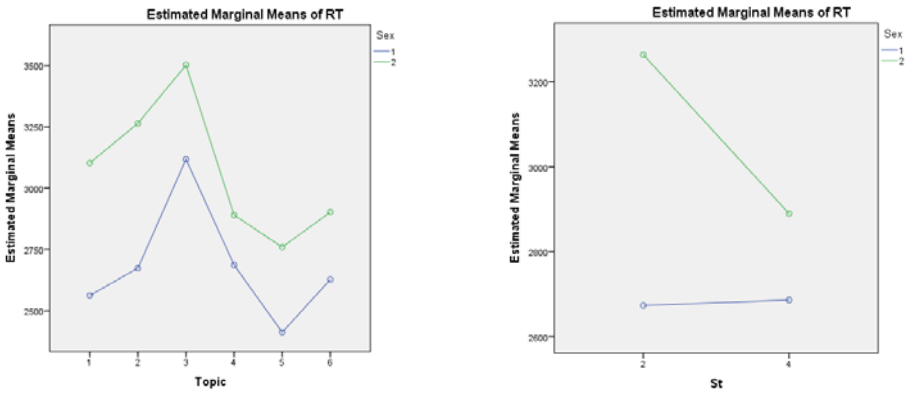


Figure 6 Study 3 SPSS plots of Means of Line 10 RTs for male and female participants for all topics and for Survival Non-Life Threatening and Honesty
Note: Upper lines shows male RTs and lower ones the female RTs. Topic 1 = Survival Life Threatening, 2 = Survival Non-Life Threatening, 3 = Child-Loss, 4 = Honesty, 5 = Socially Important and 6 = Socially Unimportant.

Although the analyses do not reach conventional significance, the results suggest that in addition to testing the predictions with all participants' data combined, exploratory

assessments using male and female data separately may be fruitful. Accordingly, the analyses, described below, have been conducted three times; with all participants' data combined and for the two sexes separately.

Use of Transformed Data.

Micceri (1989) refers to normal distributions in psychological research as "unicorns" (p. 156). Referring to real world data, he quotes Geary (1947) approvingly "normality is a myth: there never was and never will be a normal distribution" (p. 165) and draws attention to the normality assumptions that "permeate both textbooks and the research literature of the social and behavioral sciences" (p. 157). The non-normality of much real world data is confirmed by Blanca, Arnau, et al. (2013). Nevertheless, Micceri (1989) writes that "Adequate research is available to suggest that most parametric statistics should be fairly robust to both alpha and beta given light tail weights and moderate contaminations" (p. 165) and suggests that "for extreme to exponential asymmetry, one might expect at least the independent means *t* (given approximately equal Ns) and *f* to exhibit robustness to alpha, if not beta" (p. 167). This view is supported by McDonald (2014) who states that "parametric tests aren't extremely sensitive to deviations from their assumptions" (p. 144). However, there does not appear to be a general consensus on the limits to robustness and non-normality. West, Finch, and Curren (1995) regarded a value of skewness > 3 as highly non-normal. Brown (2015) cites a rule of thumb based upon Bulmer (1979) that observations with positive skewness between 0.5 and 1 are moderately skewed and values under 0.5 are approximately symmetric. According to this interpretation, all topics' RTs in the current research must be regarded as moderately skewed. However, Tabachnick and Fidell (2013) state that "In a large sample, a variable with statistically significant skewness often does not deviate enough from normality to make a substantive difference in the analysis" (p. 80). Lix, Keselman, and Keselman (1996) quoted Glass, Peckham, and Sanders (1972) as concluding that "skewed populations have very little effect on either the level of significance or the power of the fixed effect model F-test" (p. 273) except when group sizes are small. Bulmer (1979) refers to several empirical and theoretical investigations that have been carried out to determine the effect of non-normality on the distribution of *t* in small samples and concludes that the "distribution is remarkably insensitive to small or moderate departures from normality and that one need not worry

about this provided either that the non-normality is not extreme or that the number of observations is not too small" (pp. 152-153). Oberfeld and Franke (2013) warn against repeated measures analyses with non-normally distributed data if samples are less than 30, while Roberts and Russo (1999) warn that ANOVAs can be conducted only when all skews are in the same direction, as they are in the current experiment. Given that the current research has sample sizes for each topic after exclusion of outliers in excess of 190, which are large by the numbers normally seen in contradiction paradigm experiments, and skewness values that are only moderate, according to Brown (2015), parametric analyses have been conducted. However, Table 10 records positively skewed RT data even if it is only moderately skewed, according to Brown (2015). While the quotations above appear to support the application of parametric analyses, the vagueness of some comments might still provide room for a different opinion. In order to cater to such a view, two further sets of analyses have been performed and are reported below: (1) parametric tests on the data transformed by the square root transformation to establish better normality and (2) non-parametric tests where possible and meaningful. Trochim, Donelly, and Arora (2016) advise that care must be taken when interpreting the results of statistical analysis based upon transformed variables as the relationship between variables is no longer on the same scale as at the commencement of the analysis. Osborne (2002) notes that the intended effect of the square root transformation is to compress the upper part of the distribution more than the lower, thereby improving the skewness if the distribution is positively skewed. As the purpose of the tests is to ascertain differences between means of distributions, detecting significant differences between the means of the square roots of two or more variables may still provide useful insight into the differences between the untransformed variables, according to some authors (Dallal, 2012). Descriptive Statistics for the transformed variables appear in Table 11 below.

Table 11
Descriptive Statistics for Study 3 Square Root Transformed RTs

	N	Mean	SD	Skewness		Kurtosis	
				Statistic	Std. Error	Statistic	Std. Error
SqrtSLT	192	51.58	12.23	0.435	.175	-0.431	0.349
SqrtSNLT	193	52.92	12.19	0.200	.175	-0.108	0.348
SqrtCL	194	56.08	12.26	0.325	.175	-0.344	0.348
SqrtH	193	51.45	11.66	0.465	.175	-0.214	0.348
SqrtSI	192	49.52	11.07	0.403	.175	-0.058	0.349
SqrtSU	196	51.25	11.5	0.154	.174	-0.643	0.346

Note: SqrtSLT = Square root of Survival Life Threatening, SqrtSLNT = Square root of Survival Non-Life Threatening, SqrtCL = Square root of Child-Loss, SqrtH = Square root of Honesty, SqrtSI = Square root of Socially Important, SqrtSU = Square root of Socially Unimportant.

Skewness values for topics were substantially reduced by transforming the data, with all values below 0.5. According to Brown (2015), these variables may be regarded as approximately symmetrically distributed. Kurtosis values are all now negative with half being reduced absolutely and half increased absolutely. However, all have remained below two times the standard error.

Prediction Analyses.

***Planned Comparison Table*.**

Several of the predictions made in the Theoretical Background require planned comparisons between one or more topics' mean of Line 10 RTs and equivalent combinations of other topics' RTs. In order to facilitate an understanding of these, Table 12 below provides the SPSS Planned Comparison weightings used for each of the Study 3 predictions.

Table 12
Weightings used in SPSS Planned Comparisons for Study 3

Prediction	Description	Topic and Weighting					
		SLT	SNLT	H	CL	SI	SU
1a	Survival Group topics are more noticeable than the Non-Adaptive Group as measured by RTs.	1	1			-1	-1
1b	Child-Loss, as Study 3 representative of the Reproduction Group, will be more noticeable than the Non-Adaptive Group as measured by RTs.				1	-0.5	-0.5
1c	Honesty, as Study 3 representative of the Reputation Group, will be more noticeable than the Non-Adaptive Group as measured by RTs.			1		-0.5	-0.5

Note: SLT = Survival Life Threatening, SNLT = Survival Non-Life Threatening, CL = Child-Loss, H = Honesty, SI = Socially Important, SU = Socially Unimportant.

Table 12 shows that in testing the three main predictions for the current research, Predictions 1a, 1b, and 1c, the relevant adaptive contradiction RTs have each time been compared with the Non-Adaptive Contradiction RTs. In order to discount misinterpretation of the results due to family-wise error, *p* values at .01 or less have been used as the basis of interpretation.

Prediction 1a – The Survival Group compared to the Non-Adaptive Group.

In Chapter 2 - Theoretical Background it was predicted (Prediction 1a) "*that a Survival Group comprised of the two categories of survival contradictions will be more noticeable than a Non-Adaptive group comprised of the Non-Adaptive ones*". To test this prediction, a one-way repeated measures ANOVA with planned comparisons for all participants excluding outliers, using the untransformed RT data, showed that the means

of the Line 10 RTs of the Survival group, Life Threatening and Non-Life Threatening, were significantly longer than those of the Non-Adaptive Topics, Socially Important and Socially Unimportant, F (1, 174) = 12.98, $p < .001$, *partial* $\eta^2 = .069$, a moderate effect size. The planned comparison was repeated for the two sexes separately. For females, the planned comparison showed that the Survival Group did have longer mean Line 10 RTs, at 2,551 ms than the Non-Adaptive Group's, at 2,419 ms, but the planned comparison did not reach conventional statistical significance, F (1, 95) = 2.202, $p = .141$. For the male participants, the planned comparison showed that the Survival Group did have statistically significantly longer RTs at 3,111 ms than the Non-Adaptive Group at 2,728 ms, F (1, 78) = 13.637, $p < .001$, *partial* $\eta^2 = .149$, a large effect size.

The analyses were repeated with the Square Root Transformed RTs. It showed that the Square Roots of the Survival Group of topics' Line 10 RTs were statistically significantly greater than those of the Non-Adaptive Group, F (1, 174) = 11.478, $p = .001$, *partial* $\eta^2 = .062$, a moderate effect size. For female participants only, the analysis showed that the Survival Groups' transformed variables were not statistically significantly different from the Non-Adaptive Groups', F (1, 95) = 2.71, $p = .103$. For male participants, the results of the analysis were F (1, 78) = 110.4, $p = .002$, *partial* $\eta^2 = .118$, a large effect size. All three planned comparisons with the transformed data confirm the equivalent results obtained with the untransformed data. The transformations had little effect on the analyses. Prediction 1a is confirmed when the analysis is conducted using all participants' RTs, but is not confirmed according to conventional significance standards when only female participants' RTs are analysed. It is very strongly confirmed when only male participants' RTs are analysed. The results are summarized in Table 13 below.

Table 13
Summary of Prediction 1a results

Ptcpts.	Data type	*F*	*p*	*partial* η^2	Effect Size
All	Untransformed	F (1, 174) = 12.98	<. 001	.069	Moderate
Females	Untransformed	F (1, 95) = 2.202	.141	NA	NA
Males	Untransformed	F (1, 78) = 13.637	<.001	.149	Large
All	SQRT Transfmd.	F (1, 174) = 11.478	.001	.063	Moderate
Females	SQRT Transfmd.	F (1, 95) = 2.71	.103	NA	NA
Males	SQRT Transfmd.	F (1, 78) = 11.406	.002	.118	Large

Note: SQRT Transfmd = Square root transformed.

As the distributions of the untransformed RTs were moderately skewed, prediction 1a was tested using the non-parametric Wilcoxon Signed Rank Test. As this test is not available in a planned comparison form (but see Rosner, Glynn, & Lee, 2006), both Survival topics' RTs were matched with both Non-Adaptive topics. The results appear in Table 14 below.

Table 14
Wilcoxon Signed Rank Tests for Prediction 1a

Test topics	Z	Asymp. Sig (2-tailed)	r as a measure of Effect Size	
SI- SLT	-2.565	.010	-.188	Small
SI- SNLT	-3.557	<.001	-.26	Medium
SU - SLT	-0.657	.511	N/A	
SU - SNLT	-2.296	.022	-.167	Small

Note: SLT = Survival Life-Threatening, SNLT = Survival Non-Life Threatening, SI = Socially Important, SU = Socially Unimportant.

Three comparisons show statistically significant differences. The very similar means for the Line 10 RTs of the Socially Unimportant and the Survival Life Threatening, at 2,759 ms and 2,810 ms respectively (see Table 9), unsurprisingly failed to yield a significant result. Nevertheless, the results as a whole are interpreted as confirming Prediction 1a, the three comparisons with significant results prevailing over the null result for the Socially Unimportant/Survival Life Threatening comparison.

Prediction 1b - The Reproduction Group (Child-Loss) compared with the Non-Adaptive Group.

For Study 3, Child-Loss was chosen as the representative of the Reproduction Group. A one-way repeated measures ANOVA with planned comparisons using the untransformed data were conducted to test the expectation that the mean of the Line 10 Child-Loss RTs would be greater than the those of the Non-Adaptive Group. It showed that the Child-Loss RTs were statistically significantly longer than the Non-Adaptive Group, F (1, 180) = 41.219, $p < .001$, *partial* $\eta^2 = .19$, a large effect size. For female participants, the planned comparison showed a similar result, F (1, 97) = 23.525, $p < .001$, *partial* $\eta^2 = .195$, a large effect size. For male participants, the planned comparison also yielded a similar result, F (1, 82) = 17.833, $p < .001$, *partial* $\eta^2 = .179$, a large effect size.

As the prediction reached statistical significance, the test with all participant data were repeated with Square Root transformed Line 10 RTs. It showed that the transformed mean of the Child-Loss Line 10 RTs was significantly higher than the means of the transformed Non-Adaptive Group, F (1, 180) = 44.67. $p < .001$, *partial* $\eta^2 = .199$, a large effect size. As this test presents a result as clear as the analysis with untransformed data, separate tests for Males and Females were not conducted. The results are summarized in Table 15 below.

Table 15
Prediction 1b results with different sets of participants and data

Participants	Data type	*F*	*p*	*partial* η^2	Effect Size
All	Untransformed	F (1, 180) = 41.22	<.001	.195	Large
Females	Untransformed	F (1, 97) = 23.53	<.001	.19	Large
Males	Untransformed	F (1, 82) = 17.83	<.001	.179	Large
All	SQRT Transfmd.	F (1, 180) = 44.67	<.001	.199	Large

Note: SQRT Transfmd = Square root transformed.

As the planned comparison with the untransformed variables produced a significant result, the Wilcoxon Signed Rank test was used for confirmatory purposes. As planned comparisons are not available for this test, the Child-Loss topic was compared with both the Non-Adaptive ones. The results appear in Table 16 below.

Table 16

Results of Wilcoxon Signed Rank Test comparing Means of the Child-Loss and Non-Adaptive topics

Topics	Z	Asymp. Sig. (2-tailed)	*r* as a measure of Effect Size	
SI - CL	-6.175	< .001	-.314	Moderate
SU - CL	-4.864	< .001	-.247	Moderate

Note: SI = Socially Important, SU = Socially Unimportant, Cl = Child-Loss.

All tests confirm the statistically significant difference between the Child-Loss topic and the Non-Adaptive ones.

Prediction 1c – The Reputation Group (Honesty) compared with the Non-Adaptive Group.

Honesty was chosen as the representative of the Reputation Group for Study 3. A within-groups ANOVA with planned comparisons was conducted to test the expectation that the mean of the Honesty Line 10 RTs would be greater than those of the Non-Adaptive Group. The analysis showed that the Honesty Line 10 RTs were greater than the Non-Adaptive Group's but not statistically significantly greater, F (1, 179) = 1.65, p = .201, *partial* η^2 = .01. When the analysis was repeated for female participants, the result was similar with F (1. 96) = 2.06, p = .154, *partial* η^2 = .021. For males, the result was also statistically insignificant with $F < 1$. As no statistically significant results were produced, analyses with square root transformed data were not conducted. The results are summarized in Table 17 below.

Table 17

Summary of Prediction 1c results.

Participants	Data type	F	*p*	*partial* η^2	Effect Size
All	Untransformed	F (1, 179) = 1.65	.201	.010	NA
Females	Untransformed	F (1. 96) = 2.06	.154	.021	NA
Males	Untransformed	< 1	NA	NA	NA

Prediction 2a – Life Threatening compared to Non-Life Threatening.

Prediction 2a predicted that "*Life Threatening contradictions will be more noticeable than the Non-Life Threatening contradictions".* A paired samples *t*-test was conducted with the untransformed data to compare the relevant RTs. The mean of the Survival Non-Life Threatening Line 10 RTs was higher, at 2,920 ms, than the mean of the Survival Life Threatening Line 10 RTs, at 2,740 ms, $t = -1.58$ and $p = .116$. The prediction was disconfirmed. The two means are in the opposite direction from that predicted, although this result did not reach statistical significance. The *t*-tests were conducted for the two sexes separately but in neither case did they reach statistical significance, $t = .924$, $p = .357$ for females and $t = .687$, $p = .494$ for males. The failure to confirm the prediction combined with the direction of the difference between the means of the two topics is a surprising result.

Prediction 2d – The Socially Important topic compared to the Socially Unimportant.

Prediction 2d stated "*that the Socially Important contradictions would be more noticeable than the Socially Unimportant ones".* A paired samples *t*-test conducted to test this prediction produced a statistically significant result, $t = 2.001$, $p = .047$, *partial* $\eta^2 = .02$, a small effect size, with the mean of the Socially Unimportant Line 10 RTs, at 2,722 ms, higher than the mean of the Socially Important Line 10 RTs at 2,537 ms. The prediction was disconfirmed. When the paired samples *t*-test was repeated for female participants the means of the Socially Unimportant Line 10 RTs, at 2,623 ms. was significantly longer than the Socially Important Line 10 RTs, at 2,358 ms, $t = -2.122$, $p = .036$, *partial* $\eta^2 = .024$, a small effect size. For male participants, the mean of the Socially Unimportant Line 10 RTs at 2,833 ms, was again higher than the mean of the Socially Important ones, at 2,740 ms, $t = -0.687$, $p = .494$, but failed to reach statistical significance. These results indicate that the male participants were insensitive to any difference in the two categories of contradiction, relative to female sensitivity. These result are summarised in Table 18 below.

Table 18
Prediction 2d Results Summary. All times in ms.

Participants	SI Mean	SU Mean	t	p	*partial* η^2	Effect Size
All	2537	2722	2.001	.047	.020	Small
Females	2358	2637	-2.122	.036	.024	Small
Males	2740	2833	-0.687	.494	NA	NA

As the analysis of the combined participants' RTs reached statistical significance, albeit in the counter prediction direction, a paired samples *t* - test was conducted on the square root transformed data of all participants. The mean of the square roots of the Socially Unimportant Line 10 RTs, at 50.93 $ms^{1/2}$, was greater than that of the square roots of the Socially Important Line 10 RTs, at 49.21 $ms^{1/2}$, $t = 1.994$, $p = .048$, *partial* $\eta^2 = .021$, reaching conventional significance. When the tests were repeated with Female data, the mean of SqrtSU, at 49.89 ms, was greater than the mean of SqrtSI, at 47.39 ms, t = -2.103, $p = .038$ *partial* $\eta^2 = .024$. When the *t*-test was repeated with Male participant data, the mean of SqrtSU at 51.25 $ms^{1/2}$ was greater than the mean of SqrtSI at 50.09 $ms^{1/2}$, $t = -.674$, $p = .502$. Both *t*-tests using the two sexes' separate square root data replicated the results of the tests with untransformed data. These result are summarised in Table 19 below.

Table 19
Summary of Prediction 2d Results with Transformed Data. All times in $ms^{1/2}$.

Participants	SQRT SI Mean	SQRT SU Mean	*t*	*p*	*partial* η^2	Effect Size
All	49.21	51	1.994	.048	.021	Small
Females	47.39	49.89	-2.103	.038	.024	Small
Males	50.09	51.25	-0.674	.502	NA	NA

Note: SQRT SI = Square root of Socially Important, SQRT SU = Square Root of Socially Unimportant

The two sets of results from Tables 18 and 19 are condensed below into Table 20 for clarity.

Table 20
Significance of SI and SU t-test analyses

	Data Type	
Participants	Untransformed	Sqrt Transformed
All	Sig	Sig
Female	Sig	Sig
Male	Non-Sig	Non-Sig

Note: SI = Socially Important, SU = Socially Unimportant.

Because the results of the *t*-tests with the data of both sexes combined and the data of female participants reached statistical significance, a Wilcoxon Signed Rank Test using untransformed data were conducted. This failed to reach conventional statistical significance, Sig (2-tailed) $p = .096$.

The failures of the Wilcoxon test and both analyses with male sets of data to reach statistical significance counters the findings of the *t*-tests with the full and female sets, which produced significant results with small effect sizes. Nevertheless, it is surprising that the direction of the difference between the two topics is consistently contrary to that predicted. An explanation may lie in possible psychological complications, for example, an aversion to traffic related fines. More generally, there may be an aversion to socially imposed constraints on one's behavior to which well understood penalties for infringements are attached. Such an aversion could have caused some participants to have pressed the space bar quickly.

Discussion

Summary of Results.

Despite the application of an Outlier policy consisting of the elimination of all data points greater than the Mean plus two standard deviations, the data collected and analysed in this chapter was positively skewed. It was therefore analysed untransformed and transformed using the square root transformation. Prediction 1a that *a Survival Group comprised of the two categories of survival contradictions will be more noticeable than a Non-Adaptive group comprised of the Non-Adaptive ones* was confirmed when all participants' data was analysed and when male participants' data was analysed, It was not confirmed when only female data was analysed. Prediction 1b that *the Reproduction Group of contradiction would be more noticeable the Non-adaptive ones* was confirmed using the Child-loss topic as a representative of the Reproduction group of topics, when all data was analysed and when male and female data sets were analysed separately. Prediction 1c that *the Reputation group of contradictions will be more noticeable than the Non-Adaptive contradictions* was not confirmed, using Honesty as the representative topic of the Reputation Group of topics. Prediction 2a that t*he Life-Threatening contradictions will be more noticeable than the Non-Life Threatening ones* was disconfirmed. Prediction 2d that the *Socially Important contradiction would be more noticeable than the Socially Unimportant contradictions* was disconfirmed.

Major Predictions.

The hypothesis that animated this research was that CD, like other universal human psychological attributes, has evolved because it confers a selection advantage. The hypothesis was operationalised in the form of a prediction that the mind is biased, with conflicting cognitions of adaptive relevance, and therefor of greater importance, being more noticeable than Non-Adaptive ones. Adaptive contradictions would confer a selection advantage if resolved via the acquisition of new relevant information, permitting higher quality decisions to be made, while resolving Non-Adaptive contradictions would not. The reproductive rewards of higher quality decision-making constituted the selection pressure for the evolution of both the cognitive dissonance attribute itself and for a bias in sensitivities. Those who were able to make high quality

decisions, and did so, were more reproductively successful on average and over many generations than those who could or did not.

The hypothesis that the mind is biased in its sensitivity to adaptive contradictions was tested in Study 3 for the Survival Group (Prediction 1a) and confirmed when all participants' data were included in the analysis, the planned comparison reaching statistical significance with a moderate effect size. When the same prediction was tested for the two sexes separately, it was again confirmed for male participants, this time with a large effect size, but not for female ones. When the analyses were repeated with the square root transformed data, it was again confirmed when all participants' transformed data were analysed together and again when the male participants transformed data were analysed separately. When the female participants' transformed data were analysed, the analysis just failed to reach conventional statistical significance. As confirmation, further assessment was sought in four non-parametric Wilcoxon Signed Rank tests (both Survival topics compared to both the Non-Adaptive ones) using all participants' untransformed data. Three out of the four produced statistically significant results. The prediction is interpreted as having been confirmed, given that it was made for all participants' combined data. Despite female tests producing results in the predicted direction, the failure of these with both untransformed and transformed data to reach statistical significance was unanticipated. The importance that the male participants attached to survival-related contradictions, relative to females, suggests that there may be underlying factors not explicitly explored in this research influencing the results. One may be mortality salience, mentioned in connection with Terror Management Theory (TMT) in Chapter 2 -Theoretical Background. Some TMT research has shown gender moderated effects of mortality salience (Arndt, Greenberg, & Cook, 2002; Burke, Martens, & Faucher, 2010; Hirschberger, Florian, Mikulincer, Goldenberg, & Pyszczynski, 2002). In addition, evolutionary psychological considerations suggest that fear of loss of life due to predation when isolated from the extended family group and intra-male competition and violence may have played a part. Contradictory information related to survival constitutes a threat to life which could manifest itself in both sets of circumstances (Buss & Duntley, 2003; Daly & Wilson, 1988; Treves & Naughton-Treves, 1999; Treves & Palmqvist, 2007). This reasoning may explain the large effect size obtained in the male comparisons. Females' greater

emphasis on children's wellbeing, relative to that of males, may also provide a partial explanation (Burns, 2015).

The key Reproduction prediction, Prediction 1b, was tested via a comparison of a single reproduction topic, Child-Loss with the Non-Adaptive group consisting of the Socially Important and Socially Unimportant topics. For all participants combined and when separate analyses were conducted for the two sexes, the Child-Loss topic contradictions were shown to be more noticeable than the Non-Adaptive Group, confirming Prediction 1b.

The key Reputation prediction, Prediction 1c, was tested via a comparison of the Honesty topic with the Non-Adaptive Group. Although the Honesty topic's mean RT was longer than both the Non-Adaptive topics', there was no statistically significant difference, both when the analysis contained all participants' data and when separate analyses were conducted for female and male participants. This failure to confirm the prediction invited the view that a further study employing both the Reputation topics, Honesty and Reciprocity, would allow an examination of Prediction 1c with greater thoroughness and might provide a better test of the prediction.

Minor Predictions.

The predictions pertaining to relative noticeabilities of the topics within the Groups proposed that the Life Threatening Survival stories' contradictions would be more noticed than the Non-Life Threatening ones on the 'common sense' view that survival is essentially about threats to life. However, it was not confirmed as all analyses with the combined data of both sexes produced longer RTs for the Non-Life-Threatening stories, although not with a statistically significant difference. Both analyses of the separated genders' data also produced non-significant results, again with longer RTs for the Non-Life-Threatening stories. Terror Management Theory (TMT) may once again provide an explanation for these results. An expectation that the Life-Threatening contradictions would be much more noticeable than Non-Life-Threatening ones is very reasonable given the severity of the consequences of loss of life. The reverse result actually obtained, although not meeting statistical significance, can be interpreted as TMT having a substantial effect. When participants were presented with these mortality-related contradictions, rapid pressing of the space bar would have enabled them to reduce the impact of the aroused aversive emotions by quickly

diverting their attention to the next line of the story. A related explanation may be found in the threat → aversive arousal → avoidance or approach-related motivation model of Jonas, McGregor, et al (2014). The shorter RTs obtained for the Survival Life-Threatening stories may indicate an avoidance motivation reaction while the longer Survival Non-Life Threatening RTs may indicate an approach motivation. Further research with a wider set of stories may confirm these views.

The analysis of the RTs for the two Non-Adaptive topics, Socially Important and Socially Unimportant, showed a significant difference between them with the Socially Unimportant RTs being longer. This was contrary to Prediction 2d. The disconfirmation occurred again when only female participants' data were analysed but not when male data were analysed. When the analyses were repeated with transformed data a significant result was again obtained with the combined and female data. However, the analysis with the male data yielded no significant differences. One possible explanation for these results may lie in the nature of the Non-Adaptive topics. The Socially Important stories were concerned with rules of the road and associated penalties, a form of systematization perhaps more orientated to the male mind than to the female one (Baron-Cohen, 2003), while the Socially Unimportant ones were more personal, perhaps of greater interest to female participants. Female lack of interest in road rules may have resulted in rapid pressing of the space bar. This suggestion is obviously highly speculative. An explanation similar to the threat → aversive arousal → avoidance or approach-related motivation model of Jonas, McGregor, et al (2014) provided for the unexpected results of the Survival Life Threatening and Non-Life Threatening comparison can be offered. The aversive traffic offence penalties may be causing an avoidance motivation reaction with consequent shorted RTs for the Socially Important topic. Inevitably, there is a degree of cloudiness surrounding any interpretation of this result. A different result might be obtained with a different set of stories.

Study 3 examined the relative noticeabilities of the Survival and Non-Adaptive contradictions and achieved an initial exploration of the comparison of the Reproduction and the Reputation contradictions with the Non-Adaptive ones. It confirmed the hypothesis underlying the current research with respect to Survival, advances a tentative confirmation of Prediction 1b, pertaining to Reproduction, but leaves Prediction 1c, pertaining to Reputation, in need of further exploration.

Chapter 6 – Study 4: Prediction Testing (cont.)

Study 3 was designed primarily to test the Predictions 1a, 1b and 1c that adaptive contradictions would be more noticeable than Non-Adaptive ones. In the case of Prediction 1a, pertaining to Survival, this objective was achieved using both Survival related topics. In tests pertaining to the Reproduction and Reputational groups, the tests were conducted using one topic only, judged to provide the highest contrast with the Non-Adaptive Group. Study 4 was designed to compare these groups as wholes with a Non-Adaptive topic, and to assess the lower level predictions, 2b and 2c, of the relative noticeability of the topics comprising the Adaptive Groups. It also permitted an assessment the relative noticeability of all three Adaptive Groups, Predictions 3a and 3b.

Method

As for Study 3, Study 4 was conducted using the contradiction paradigm. Study 4 included the five topics of the Reproduction and Reputation Groups, Child-Loss, Family, and Fidelity, and Honesty and Reciprocity, in order to make intra-group comparisons. This meant that the Non-Adaptive Group would be represented by one topic only. In Chapter 5 - Study 3: Prediction Testing, it was reported that Prediction 2d, "*the Socially Important contradictions would be more noticeable than the Socially Unimportant ones*" was disconfirmed, the mean of the Socially Unimportant Line 10 RTs being higher than the Socially Important one. In Chapter 5 - Study 3: Prediction Testing, the Socially Unimportant topic was assessed as more likely to be unencumbered with possible psychological complications as the Socially Important topic's stories featured penalties imposed for infringements of socially important rules. Consequently, the Socially Unimportant topic was chosen for use in Study 4 as the Non-Adaptive base line topic. Not only was it regarded as less encumbered with psychological complications, it was also the topic with the higher mean of Line 10 RTs. This choice ensured that comparisons between the Reproduction and Reciprocity Groups on the one hand and the Non-Adaptive Group on the other would be conservative. The Socially Unimportant stories used in Study 4 were the same as those used in Study 3. Study 4 also permitted the different triplets of stories constituting the three newly introduced adaptive topics, Family, Fidelity and Reciprocity, to be assessed

for different intra-topic differences of Line 10 RTs. The examination of all the contradictory topics' triplets of stories appears as Appendix N - Triplet ANOVA Analyses.

Participants.

Participants were recruited for this Study in the same way as for Study 3. As far as was practical, the number of people recruited from the general public, as opposed to students from university campus sites, was very similar. In order to make the two studies as comparable as possible, the number of participants was kept at 201, 93 being female. This policy was adopted in order to reduce the likelihood of an effect size of a particular magnitude reaching conventional statistical significance in one study but not in the other.

Stimuli.

The stimuli were 12 stories, as for Study 3, but with the topics changed as explained above. To avoid order effects, the stories were varied in their presentation sequence using the same 18 version mechanism as used in Study 3. The non-contradictory stories were the same as those used in Study 3.

Material.

The software used in Study 4 was the same as that used for Study 3.

Procedure.

On commencement, each participant first read instructions (see Appendix L - Instruction) in which they were informed that the experiment had been approved by the relevant Massey University Ethics Committee and that they would be reading twelve short stories of eleven lines each. The story reading would be preceded by a short practice story and questions. In all cases, they were to read a line and press the space bar once they had finished reading that line. The program would then display the next line and so on until the end. Following each story, a page containing four simple questions

about that story would be displayed, which were to be answered by typing very short answers, many with a yes (Y) or no (N). They were informed that they would be required to provide some anonymous profile data and that providing their email addresses was voluntary. If they chose to do so, they would later receive an explanation of the objectives and results of the experiments. They were also told that they may cease to participate at any time and that their informed consent would be assumed if they chose to advance to the practice story, which largely repeated the instructions. Once participants had read these instructions, they pressed the space bar to provide the profile data. Thereafter, they proceeded to read the stories and answer the questions to completion. Afterwards, participants were given a brief description of the objectives of the research. All participants who supplied their email addresses have been sent an explanation of the objectives of the research and a summary of the results.

Results

Initial Data Analysis.

Normality Test and Descriptive Statistics.

Homogeneity of Variance was tested by application of the F_{Max} test, the ratio of the largest topic variance to the smallest not exceeding 10 (Tabachnick & Fidel, 2013). The value of F_{Max} for Study 4 RTs is 2.37. Normality checks revealed that Kurtosis values were more than twice the Standard Error (Rose, Spinks & Canhoto, 2014) indicating high Kurtosis. The RT data were positively skewed, with the Kolmogorov-Smirnov test showing significance levels under .05 for all topics. However, the Standard Error of Skewness declines to small values with large samples[6] undermining the utility of the test (Rose et al. 2014). Descriptive Statistics appear in Table 21 below.

[6] The standard error of skewness is roughly inversely proportional to the square root of the sample size as indicated by its definition, $(6n(n-1)/(n-2)(n+1)(n+3))^{1/2}$ which approximates to $n^{-1/2}$ when n is large.

Table 21

Descriptive Statistics for all Study 4 participants. All measurements are in ms.

Topic	N	Mean	SD	96% C.I.		Min	Max	Skewness		Kurtosis	
				Lower	Upper			Statistic	Std. Error	Statistic	Std. Error
SU	201	2700	1363	2522	2890	511	9478	1.37	0.172	3.053	0.341
REC	201	3619	2098	3327	3911	507	14732	2.001	0.172	6.749	0.341
H	201	2824	1576	2605	3044	622	10511	1.768	0.172	4.191	0.341
FAM	201	2735	1745	3192	3677	618	7723	1.245	0.172	1.817	0.341
CL	201	3434	1745	3192	3677	618	11221	1.251	0.172	1.961	0.341
FID	201	3159	1710	2910	3395	899	11337	1.845	0.172	4.393	0.341

Note: SU = Socially Unimportant, REC = Reciprocity, H = Honest, FAM = Family, CL = Child-Loss, FID = Fidelity.

First Language Analysis.

To ascertain whether there was sufficient difference in RTs to justify separate analyses for English First Language and Non-English First Language participants, Between-groups *t*-tests were conducted comparing the Line 10 RTs for both categories of participants for all topics. Descriptive Statistics and the results of the t-tests appear in Table 22 below.

Table 22
Study 4 Descriptive Statistics and t-test results for English and 'Other' First Language Participants. All times in ms.

Topic	Language	N	Mean	SD	Sig (2 tailed)	*t*	df
SU	English	189	2687	1374	0.581	0.552	199
	Other	12	2912	1216			
REC	English	189	3556	1948	0.089	0.976	199
	Other	12	4619	3741			
H	English	189	2801	1536	0.35	0.851	199
	Other	12	3200	2153			
FAM	English	189	2709	1341	0.396	1.089	199
	Other	12	3149	1439			
CL	English	189	3410	1676	0.438	0.778	199
	Other	12	3815	2671			
FID	English	189	3034	1504	0.046	2.238	11.314
	Other	12	5104	3182			

Note: SU = Socially Unimportant, REC = Reciprocity, H = Honesty, FAM = Family, CL = Child-Loss, FID = Fidelity.

The analysis revealed that for all topics, the Non-English First Language participants had longer RTs than the English First Language speakers with a statistically significant difference being obtained for the Fidelity topic. Examination of the Fidelity RTs revealed that for this topic some very high RTs were obtained for two participants. However, as their RTs for other topics were near average and they were from the Oceania continent, it was decided not to separate them or other non-English First Language participants for separate analysis.

Continent of Origin Analysis.

To ascertain whether participants' continent of origin acted as an independent variable, a one way between-groups ANOVA for all topics was conducted with continent of origin as the independent factor. The result showed no statistically significant differences in Line 10 RTs for any topic. Participants originated in 4 continents. Data from all participants from all continents was retained as a single body of data for subsequent analyses.

Educational Level Analysis.

As for Study 3, a one way between-groups ANOVA was conducted for all topics using Study 4 data to assess whether there were any differences in mean Line 10 RTs for the three levels of education for each topic. Statistically significantly different RTs were found for the Socially Unimportant topic only. As all analyses to test the predictions was made on the RTs after the elimination of outliers, another One Way Between-Groups ANOVA was conducted with Study 4 RTs without outliers to ascertain if the statistically significant difference between the three educational groups was retained. None was found for any topic. As a consequence, all data from all participants at all educational levels was retained for subsequent analyses.

Tests of Sex/gender.

To ascertain differences between male and female RTs, between-groups *t*-tests were conducted for all topics comparing the Line 10 RTs of male participants with those of female participants. The results appear in Table 23 below.

Table 23

Descriptive Statistics and t-test results for Study 4 Line 10 RTs for Male and Female participants for all Topics. All times in ms.

Topic	Sex	N	Mean	SD	Means diff.	*t*	Sig (2 tailed)	*Partial η2*
SU	Female	93	2517	1132	341	1.815	.07	N/A
	Male	108	2858	1523				
REC	Female	93	3438	1694	338	1.139	.256	N/A
	Male	108	3775	2389				
H	Female	93	2755	1582	129	0.58	.563	N/A
	Male	108	2884	1575				
FAM	Female	93	2584	1173	280	1.499	.125	N/A
	Male	108	2865	1475				
CL	Female	93	3154	1521	522	2.134	.034	0.023 Small
	Male	108	3676	1890				
FID	Female	93	2992	1657	309	1.279	.202	N/A

Note: SU = Socially Unimportant, REC = Reciprocity, H = Honesty, FAM = Family, CL = Child-Loss, FID = Fidelity.

For all topics, males RTs were longer than female ones but only for Child-Loss did the difference reach statistical significance. For this topic, the effect size was small. In general, the differences were smaller than in Study 3. The ratios of male and female RTs for the 6 topics are shown in Table 24 below.

Table 24

Ratio of male to female RTs for all topics. Times in ms.

	Mean RTs		
Topic	Male	Female	Ratio M:F
SU	2858	2517	1.14
REC	3775	3438	1.1
H	2884	2755	1.05
FAM	2865	2584	1.11
CL	3676	3154	1.17
FID	3301	2992	1.1

Note: SU = Socially Unimportant, REC = Reciprocity, H = Honesty, FAM = Family, CL = Child-Loss, FID = Fidelity.

As for Study 3, because the ratio of the RTs differs according to topic, separate analyses have been conducted for all participants' data and for the two sexes separately.

Outlier Policy and Application.

The outlier policy adopted for Study 4 was the same as for Study 3. Its application resulted in a loss of 4.8 % of the data. This is also below the 5% maximum proposed by Ratcliffe (1993).

Normality Tests and Descriptive Statistics of Participant Data Excluding Outliers.

Homogeneity of Variance was tested by application of the F_{Max} test; the ratio of the largest topic variance to the smallest should not exceed 10 (Tabachnick & Fidel, 2013). The value of F_{Max} for Study 4 Line 10 RTs is 2.423. Normality checks revealed that Kurtosis values were less than twice the Standard Error indicating low Kurtosis (Rose, Spinks, & Canhoto, 2014). The distributions of all topics' Line 10 RTs were positively skewed, as expected, with the Kolmogorov-Smirnov test showing significance levels under .05 for all topics. Descriptive statistics for Study 4 appear in Table 25 below.

Table 25
Descriptive Statistics for Study 4 RTs. All times in ms.

Topic	N	Min	Max	Mean	SD	Skewness		Kurtosis	
						Statistic	SE	Statistic	SE
SU	192	511	5179	2522	1079	0.582	175	-0.362	0.349
REC	191	507	7304	3280	1430	0.382	0.176	0.62	0.35
H	191	622	5839	2585	1159	0.947	0.178	0.36	0.35
FAM	192	641	5153	2550	1057	0.566	0.175	-0.35	0.349
CL	192	618	6833	3210	1411	0.722	0.175	0.003	0.349
FID	190	0.899	6279	2867	1195	0.938	0.176	0.245	0.351

Note: SU = Socially Unimportant, REC = Reciprocity, H = Honesty, FAM = Family, CL = Child-Loss, FID = Fidelity.

Elimination of the outliers has reduced Kurtosis values to less than twice the Standard Error, eliminating the high values presented in Table 21.

Figure 7 below illustrates the Means and 95 % Confidence Intervals.

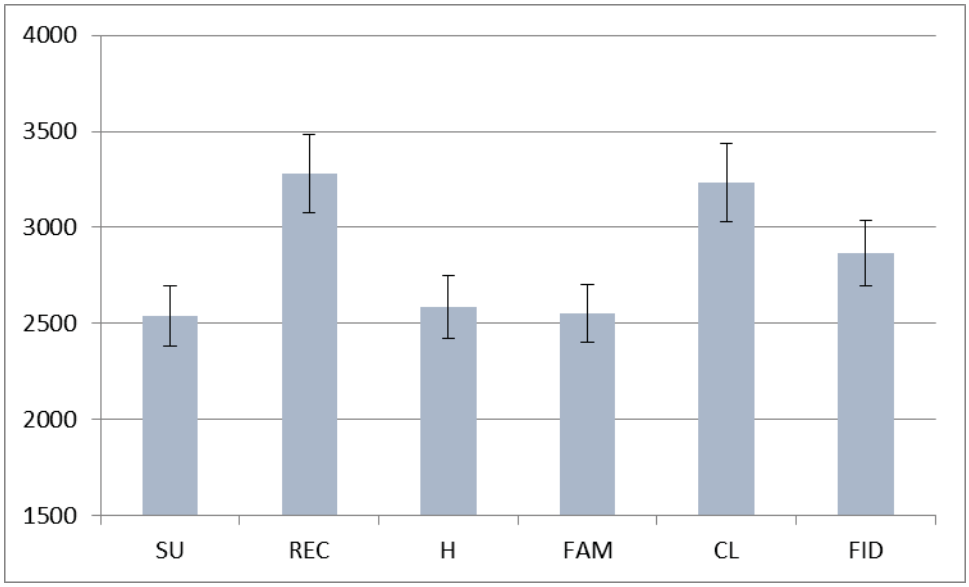


Figure 7. Histogram of Study 4 Means and 95% Confidence Intervals
Note: SU = Socially Unimportant, REC = Reciprocity, H = Honesty, FAM = Family, CL = Child-Loss, FID = Fidelity.

Study 4 RTs Skewness values were all positive and moderate according to Brown's (2015) description of the skewness range of .5 to 1. The comments made in connection with skewed data in Chapter 5 - Study 3: Prediction Testing apply to Study 4 data also so parametric tests have been conducted. Nevertheless, the RTs has been transformed to square root values in order that some analyses might be checked. Descriptive Statistics for the transformed data appear in Table 26 below.

Table 26
Descriptive Statistics for the Square Roots of Study 4 RTs.

	N	Min	Max	Mean	SD	Skewness		Kurtosis	
						Statistic	SE	Statistic	SE
SqrtSU	192	23	72	49.06	10.76	0.124	.175	-0.474	.349
SqrtREC	191	23	85	55.83	12.82	-0.092	.176	-0.563	.35
SqrtH	191	25	76	49.65	10.99	0.493	.176	-0.273	.35
SqrtFAM	192	25	72	49.41	10.46	0.144	.175	-0.587	.349
SqrtCL	192	25	83	55.47	12.32	0.234	.175	-0.3991	.349
SqrtFID	190	30	79	52.47	10.73	0.542	.176	-0.305	.351

Note: SqrtSU = Square Root of Socially Unimportant RTs, SqrtREC = Square Root of Reciprocity, SqrtH = Square Root of Honesty, SqrtFAM = Square Root of Family, SqrtCL = Square Root of Child-Loss, SqrtFID = Square Root of Fidelity.

The Skewness values have improved as a result of the transformation although the SqrtFID is just outside Brown's criterion. The skew of Reciprocity has become very slightly negative in the transformation to SqrtREC due to the higher compression of the upper range of the untransformed distribution brought about by the square root algorithm. Because the negative skewness of SqrtREC is very low, the transformed data values have nevertheless been used in ANOVAs and *t*-tests as checks on the analyses on the untransformed data where relevant. Wilcoxon Signed Rank tests have also been conducted where they are deemed helpful in the interpretation of parametric test results.

Prediction Analysis.

Planned Comparison Tables.

As in Study 3, several of the predictions made in Chapter 2 - Theoretical Background require the planned comparison of one or more of the topic Line 10 RTs with combinations of other topics' Line 10 RTs. In order to facilitate an understanding of these, Table 27 below provides the SPSS weightings used for each of the Study 4 predictions which make use of this facility.

Table 27
Weightings used in Study 4 Repeated Measures Planned Comparisons.

Prediction #	Description	Topic and Weighting					
		SU	H	REC	CL	FAM	FID
Prediction 1b	The Reproduction Group of topics are more noticeable than the Non-Adaptive Group (represented by the Socially Unimportant topic)	3	0	0	-1	-1	-1
Prediction 1c	The Reputation Group of topics are more noticeable than the Non-Adaptive Group (represented by the Socially Unimportant topic)	1	-0.5	-0.5	0	0	0
Predictions 3a and 3b	3a) the noticeability of the adaptive groups will be ranked Survival, Reproduction and Reputation and 3b) the ranking will be Reproduction, Survival, Reputation.	0	3	3	-2	-2	-2

Note: SU = Socially Unimportant, H = Honesty, REC = Reciprocity, CL = Child-Loss, FAM = Family, FID = Fidelity.

Prediction 1b Reproduction Group compared to Non-Adaptive contradictions.

In Chapter 2 - Theoretical Background, *"The prediction (Prediction 1b) is made that the Reproduction group of contradictions would be more noticeable than Non-Adaptive contradictions as measured by longer RTs"*, where the Reproduction Group consists of the Family, Child-Loss and Fidelity topics and the Non-Adaptive group is represented in Study 4 by the Socially Unimportant topic. To test this prediction, a one way repeated measures ANOVA with planned comparisons for all participants showed that the Reproduction Group Line 10 RTs were significantly longer than the Non-

Adaptive Line 10 RTs, F (1, 170) = 13.684, $p < .001$, *partial* $\eta^2 = .074$, an effect size a little larger than moderate. The planned comparison was repeated for the two sexes separately. For female participants, the Reproduction Group Line 10 RTs were statistically significantly longer than the Non-Adaptive Line 10 RTs, F (1, 86) = 6.015, $p = .016$, *partial* $\eta^2 = .066$, a moderate effect size. For male participants, the Reproduction Group Line 10 RTs were statistically significantly longer than the Non-Adaptive Line 10 RTs, F (1, 85) = 6.417, $p = .013$, *partial* $\eta^2 = .07$, moderate effect size.

As a confirmation of the above results, the one way repeated measures ANOVA with planned comparisons was conducted on the Square Root Transformed Line 10 RTs for all participants. The analysis showed that mean of the square roots of the Reproduction Group Line RTs was significantly greater than the mean of the square roots of the Non-Adaptive Line 10 RTs, F (1, 170) = 14.559, $p < .001$, *partial* $\eta^2 = .079$, an effect size a little larger than moderate. The planned comparisons were repeated for the two sexes separately. For female participants, there was a statistically significant result, F (1.84) = 7.522, $p = .007$, *partial* $\eta^2 = .082$, an effect size slightly larger than moderate. For male participants, the analysis result was similar, F (1, 85) = 7.1, $p = .009$, *partial* $\eta^2 = .077$. Analyses both with untransformed and transformed Line 10 RTs confirmed the prediction when all participants' data is analysed and when analyses are conducted for the two sexes separately. The above results are summarised in Table 28.

Table 28
Summary of the Prediction 1b Results

Participants	Data type	*F*	*p*	*partial* η^2	Effect Size
All	Untransformed	F (1, 170) = 13.684	<.001	0. 074	Moderate
Females	Untransformed	F (1, 86) = 6.015	.016	0.066	Moderate
Males	Untransformed	F (1, 85) = 6.417	.013	0.07	Moderate
All	SQRT Transfmd.	F (1, 170) = 14.559	<.001	0.079	Moderate
Females	SQRT Transfmd.	F (1.84) = 7.522	.007	0.082	Moderate
Males	SQRT Transfmd.	F (1, 85) = 7.1	.009	0.077	Moderate

Note: SQRT Transfmd. = Square Root Transformed

As Wilcoxon tests with Planned Comparisons are not available, separate Wilcoxon tests were conducted as checks on the above results, comparing the adaptive topics each with the Socially Unimportant topic. The results appear in Table 29 below.

Table 29
Results of Prediction 1b Wilcoxon Tests

Comparison	Z	Sig (2-tailed)	Effect size	
			r	Comment
FAM-SU	0.702	.483		N/A
CL-SU	5.783	<.001	.423	Medium to large
FID-SU	3.989	<.001	.296	Medium

Note: SU = Socially Unimportant, FAM = Family, CL = Child-Loss, FID = Fidelity.

Table 25 shows that the Family Line 10 RTs are marginally longer than the Socially Unimportant ones, although the Wilcoxon test shows that this is not a statistically significant difference. Nevertheless, the other tests show that Prediction 1b has been confirmed. Table 25 and the Wilcoxon tests show that this confirmation is due to the longer RTs of the Child-Loss and Fidelity topics.

Prediction 1c Reputation compared to Non-Adaptive.

In Chapter 2 Theoretical Background, "*It was predicted (Prediction 1c) that the Reputation group of contradictions will be more noticeable than Non-Adaptive contradictions"* where the Reputation Group consists of the Honesty and Reciprocity topics and the Non-Adaptive Group is represented by the Socially Unimportant topic. To test this prediction a one way repeated measures ANOVA with planned comparisons using Study 4 data for all participants showed that the Reputation Group RTs were statistically significantly longer than the Non-Adaptive RTs, F (1, 173) = 18.969, p < .001, *partial* η^2 = .099, an effect size slightly larger than moderate. The ANOVAs were repeated for each sex separately. For female participants, the analysis showed that the Reputation Group RTs were statistically significantly longer than the Non-Adaptive RTs, F (1, 84) = 12.052, p = .001, *partial* η^2 = .126, a large effect size. For male participants, the analysis also showed that the Reputation Group RTs were significantly longer than the Non-Adaptive RTs, F (1, 88) = 7.338, p = .008, *partial* η^2 = .077, an effect size slightly larger than moderate and lower than the effect size for female participants.

As a confirmation on the above results, the planned comparisons were run on the Square Root Transformed RTs for all participants. The planned comparison showed that the square roots of the Reputation Group Line 10 RTs were statistically significantly greater than the square roots of the Non-Adaptive Line 10 RTs, F (1, 173) = 18.259, $p < .001$, *partial* $\eta^2 = .095$, an effect size slightly larger than moderate. The planned comparisons were repeated for each sex separately. For female participants, the analysis showed that the square roots of the Reputation Group Line 10 RTs were statistically significantly longer than the square roots of the Non-Adaptive Line 10 RTs, F (1, 84) = 10.464, $p = .002$, *partial* $\eta^2 = .111$, a large effect size. For male participants, the analysis showed that the square roots of the Reputation Group RTs were statistically significantly longer than the square roots of the Non-Adaptive RTs, F (1, 88) = 7.871, $p = .006$, *partial* $\eta^2 = .082$, an effect size slightly larger than moderate. The results above are summarised in Table 30 below.

Table 30

Summary of Prediction 1c Results

Participants	Data type	*F*	*p*	*partial* η^2	Effect Size
All	Untransformed	*F* (1, 173) = 18.969	<.001	.099	Moderate +
Females	Untransformed	*F* (1, 84) = 12.052	.001	.126	Large
Males	Untransformed	*F* (1, 88) = 7.338	.008	.077	Moderate
All	SQRT Transfmd.	*F* (1, 173) = 18.259	<.001	.095	Moderate +
Females	SQRT Transfmd.	F (1, 84) = 10.464	.002	.111	Large
Males	SQRT Transfmd.	*F* (1, 88) = 7.871	.006	.082	Moderate

Note: SQRT Transfmd. = Square Root Transformed.

All analyses confirmed the prediction that the Reputation Group of topic contradictions are more noticeable than the Non-Adaptive ones. The interesting discovery made in the Study 4 analyses is that female participants appear to be more sensitive to reputation contradictions than males. This is particularly the case for the Reciprocity topic given that in Study 3, where the Honesty topic was compared to the Non-Adaptive Group, the comparison failed to gain statistical significance.

As a further test of the combined data tests, Wilcoxon tests showed a statistically significant result for the Reciprocity comparison but a null result for the Honesty one as shown in Table 31 below.

Table 31

Wilcoxon tests of Reputation topics compared to Socially Unimportant

Comparison	Z	Sig (2-tailed)	Effect size	
			r	Comment
REC-SU	5.92	<.001	.436	Medium to large
H-SU	0.931	.353		N/A

Although Table 25 shows that the Honesty RTs are longer than the Non-Adaptive ones, the Wilcoxon tests show that this difference is not statistically significant. Nevertheless, this difference combined with the statistically significant difference between the Reciprocity and Non-Adaptive topic demonstrate that Prediction 1c has been confirmed. There was a larger effect size in the analysis using untransformed female participants' data than was obtained with male participants' data. The differential sensitivity implies that females may be more concerned with reputation for reciprocity than are males. However, when a between-groups *t*-test was conducted to compare the means of the Line 10 RTs of the two sexes, the mean of the female Line 10 RTs, at 3,268 ms, was lower than the male equivalent, at 3,292 ms, although the difference was not statistically significant, $p = .9$. As differential reading speeds may be a partial cause of this result, conclusions about the differential sensitivity of the two sexes to Reciprocity contradictions must be confined to interpretations of the difference in effect sizes. These were large for females but moderate for males.

Prediction 2b – Child-Loss Line 10 RTs would be longer than Family Line 10 RTs.

In Chapter 2 – Theoretical Background, "*It was predicted (Prediction 2b) that the Child-Loss contradictions will be more noticeable than those pertaining to the Family topic.*" A paired samples *t* - test was conducted to test this prediction. The test found that the mean of the Child-Loss Line 10 RTs, at 3,205 ms, was statistically greater than the mean of the RTs of the Family topic, at 2,507 ms, $t = 6.155$, $p < .001$, *partial* $\eta^2 = .171$, a large effect size. The prediction was confirmed.

The analyses were repeated with each sexes' RTs separately. For females, there was a significant difference, $t = 3.634$, $p < .001$, *partial* $\eta^2 = .129$, a large effect size. For male participants, there was a significant difference, $t = 5.01$, $p < .001$, *partial* $\eta^2 = .207$, a very large effect size. Interestingly, female participants seemed to have differentiated between the two topics less than males.

Although no prediction was made in Chapter 2 - Theoretical Background pertaining to the relative size of the means of the Child-Loss and Fidelity RTs, a paired samples *t* - test was conducted to ascertain whether there was a statistically significant difference. The test revealed that the mean of the Child-Loss Line 10 RTs, at 3,163 ms, was statistically significantly greater than the mean of the Fidelity Line 10 RTs, at 2,810 ms, $t = 3.267$, $p = .001$. *partial* $\eta^2 = .055$, a moderate effect size. The difference between the mean of the Child-Loss Line 10 RTs in this test and the equivalent in the previous test is due to the different Family and Fidelity outlier exclusions. As statistically significant result had been obtained, the test was repeated separately with female and male data only. For females, the test showed that Child-Loss Line 10 RTs, at 3,039 ms was significantly greater than the female Fidelity Line 10 RTs at 2747 ms, $t = 1.988$, $p = .05$, *partial* $\eta^2 = .043$, a small to moderate effect size. For males, the Child-Loss Line 10 RTs, at 3,277ms, were significantly greater than the Fidelity Line 10 RTs at 2,869 ms, $t = 2.592$, $p = .01$, *partial* $\eta^2 = .066$, a moderate effect size. The effect sizes are slightly greater for males than for females.

Although no prediction was made in Chapter 2 - Theoretical Background as to the relative sizes of the means of the Family and Fidelity Line 10 RTs, a paired samples *t* - test was conducted to ascertain whether there was a statistically significant difference. The test revealed that the mean of the Fidelity Line 10 RTs, at 2,778 ms, was

statistically significantly greater than the mean of the Family Line 10 RTs, at 2,475 ms, t = -3.124, p = .002, *partial* η^2 = .051, a moderate effect size. As a statistically significant result had been obtained, the test was repeated separately with female and male data. For females, the test showed that the Fidelity Line 10 RTS were significantly greater at 2,736 ms than the Family ones at 2,445 ms, t = -2.026, p = .046, *partial* η^2 = .045, a small to moderate effect size. For males, the Fidelity Line 10 RTs, at 2,818 ms, were significantly greater than the Family ones, at 2,504 ms, t = -2.386, p = .019, *partial* η^2 = .057, a moderate effect size. The effect sizes are slightly greater for males than for females.

No ranking predictions for all three topics in the Reproduction Group were made in Chapter 2 - Theoretical Background. Nevertheless, the ranking of the RTs for the Reproduction Group is found to be Child-Loss, Fidelity and Family. All positions are statistically significant for all participants' combined data and for the two sexes separately.

Prediction 2c - Honesty Line 10 RTs would be longer than Reciprocity Line 10 RTs.

In Chapter 2 - Theoretical Background, *"... it was predicted (Prediction 2c) that Honesty-related contradictions would be more noticeable than the Reciprocity ones*." A paired samples *t*-test was conducted to test the prediction. The test revealed that the mean of the Honesty Line 10 RTs, at 2,533 ms, was statistically significantly lower than the mean of the Reciprocity RTs, at 3,227 ms, t = 6.047, p < .001, *partial* η^2 = .159, a large effect size. The prediction was disconfirmed.

The test was repeated for the two sexes separately. For females there was a significant difference, t = 4.610, p < .001, *partial* η^2 = .186, a large effect size. For males there was also a significant difference, t = 3.934, p < .001, *partial* η^2 = .123, also a large effect size but a smaller one than obtained with the female participants' data. Both tests disconfirmed the prediction. When between-groups *t*-tests were conducted to compare female Honesty RTs with those of males, no significant difference was detected, p = .585. For Reciprocity also, no significant differences were detected, p = .909.

Predictions 3a and 3b – The relative sizes of Survival, Reproduction and Reputation Line 10 RTs.

Prediction 3a stated that "*the ranking of noticeability for the three adaptive groups would be, first, Survival, second, Reproduction and last, Reputation*." Prediction 3b reversed the sequence of the Survival and Reproduction Groups. To assess the relative sizes of the means of the three Groups' Line 10 RTs, one ANOVA and two *t*-tests were conducted. The ANOVA was a repeated measures planned comparison between the Reproduction and the Reputation Groups' data from Study 4, using the topic weights illustrated in the last entry of Table 27. The *t*-tests compared the Survival Group RTs from Study 3 with the Reproduction and Reputation RTs from Study 4.

For the comparison of the Reproduction Group with the Reputation Group, the planned comparison ANOVA showed that there was no statistically significant difference between the two groups, F (1, 162) = 0.737, p = .392. The Descriptive Statistics for the comparison appear in Table 32 below.

Table 32
Predictions 3a and 3b Descriptive Statistics for the Adaptive Groups' topics. All times in ms.

Group	Topic	N	Mean	SD	Std Error
Reputation	REC	163	3188	1442	113
	H	163	2437	1065	83
Reproduction	FAM	163	2367	969	76
	CL	163	3187	1424	112
	FID	163	2696	1104	86

Note: REC = Reciprocity, H = Honesty, FAM = Family, CL = Child-Loss, FID = Fidelity.

The *t*-tests were used to compare one set of participants' RTs when reading one category of stories with those of another set of participants reading a different category of stories. Making such comparisons of RTs pertaining to two levels of one independent variable (Studies 3 and 4 participants) and two levels of a second one (different topics) risks the introduction of a confound if the two sets of participants' sensitivities to

contradictions are different. However, this can be tested as three topics, Socially Unimportant, Child-Loss and Honesty were common to both Studies. The availability of this data from the two studies enabled comparisons of these topics' RTs to be made with a preliminary set of *t*-tests. The expectation that there would be no statistically significant differences was met. The results of the three between-groups *t*-tests for these common topics are shown in Table 33 below.

Table 33
Studies 3 and 4 Between-groups t *- test of Line 10 RTs for the common topics*

Topic	Study 3			Study 4			*t*	Sig (2-tailed)
	N	Mean	SD	N	Mean	SD		
Socially Unimportant	196	2759	1209	193	2542	1110	1.844	.07
Child-Loss	193	3295	1435	193	3232	1441	0.426	.67
Honesty	193	2782	1273	191	2585	1159	1.584	.11

The results shown in Table 33 are interpreted as indicating that the participants of Studies 3 and 4 are not statistically significantly different in their sensitivities to contradictions. A comparison of one study's RTs for one topic with the other study's RTs for another can therefore be held to be a comparison of the two topics' noticeabilities, without there being a confounding difference between the two studies' sets of participants.

For the comparison between the Survival and Reproduction Group, the *t*-test showed that there was no statistically significant difference between the two groups, $t = 0.039$, $p = .969$. The Descriptive Statistics are shown in Table 34 below.

Table 34
Predictions 3a and 3b Descriptive Statistics for t-test of Survival and Reproduction Groups. All times in ms.

Group	N	Mean	SD	Std. Error
Survival	385	2879	1336	68
Reproduction	574	2876	1257	52

For the comparison between the Survival and Reputation Groups, the *t*-test showed that there was no statistically significant difference between the two groups, $t = 0.553$, $p = .58$. The Descriptive Statistics are shown in Table 35 below.

Table 35

Predictions 3a and 3b Descriptive Statistics for t-test of Survival and Reproduction Groups. All times in ms.

Group	N	Mean	SD	Std. Error
Survival	385	2879	1336	68
Reputation	382	2933	1346	69

The general result of the comparisons is that neither Prediction 3a nor 3b has been confirmed. None of the adaptive groups' Means of Line 10 RTs are statistically significantly different from each other's.

Discussion

Summary of Results.

Normality tests and outlier exclusions were the same as in Chapter 5. Prediction 1b *that the Reproduction group of contradiction would be more noticeable than Non-Adaptive contradiction* was confirmed, when all participants' data was analysed and when male and female participants' data was analysed separately. Prediction 1c that *the Reputation group of contradictions would be more noticeable than Non-Adaptive contradictions* was confirmed, when all participants' data was analysed and when male and female participants' data was analysed separately. Prediction 2b that *the Child-Loss contradiction would be more noticeable than those pertaining to the Family topic* was confirmed when all participants' data was analysed and when male and female participants' data was analysed separately. A comparison of the Child-Loss and Fidelity topics found that the Child-Loss contradictions were statistically significantly more noticeable than the Fidelity topic ones which in turn were statistically significantly more noticeable than the Family topic contradictions. Predictions 3a and 3b which predicted the relative noticeabilities of the three Adaptive Groups, Survival, Reproduction and Reputation, were not confirmed.

Major Predictions.

Prediction 1b predicted that reproductive contradictions would be more noticeable than Non-Adaptive ones. Prediction 1c predicted that reputation contradictions would be more noticeable than Non-Adaptive ones. In both case, these predictions were verified when all participants' data were analysed together and when separate analyses were conducted for the two sexes. They were also verified when the square root transformed data were analysed and were further verified using Wilcoxon Signed Rank Tests. There were, nevertheless, some unexpected results. The effect size of the Reputation/Non-Adaptive comparison was larger for females than for males. This suggests that the female participants may have been more sensitive to the Reputation, and particularly the Reciprocity contradictions, than male participants. In the 'gathering' communities in evolutionary times in which females possibly provided the larger proportion of the calories (Hrdy, 2009; Pennisi, 1999; Silverman, Choi, & Peters, 2007;

Wrangham, 2009) acceptance within the group, upon which they may have depended for their own and their children's wellbeing, was likely all important. Denial of access to or of information about sources of gathered food would likely be a consequence of a poor reputation for reciprocity or dishonesty. The current research was motivated by the suggested importance of reputation in ancestral times, (Gintis, Henrich, Bowles, Boyd, & Fehr, 2008; Trivers, 1971), but was not aimed at exploring differences between the sexes in depth. Both the results showing the difference and the suggested explanation for it are necessarily tentative.

Minor Predictions.

Prediction 2b that Child-Loss contradictions would be more noticeable than Family ones was verified when all participants' data were analysed together and when separate analyses were conducted with the data provided by each sex. The effect size of the analysis for males was larger than that for females, suggesting that the two sexes may be differentially sensitive to contradictions in these areas. This is interpreted as indicating that the concept of family is of more immediate relevance to females and is more closely related to the concept of child or child loss than is the case for males. For males, the non-birthing sex, there appears to have been a looser association between the two concepts, family being somewhat abstract while the loss of a child is more striking and concrete. As for the Prediction 1b result, a difference between the two sexes is regarded as tentative.

Prediction 2c that Honesty related contradictions would be more noticeable than Reciprocity ones was disconfirmed by analyses of the Study 4 data of both sexes combined and the data separated by gender. All three analyses disconfirmed the prediction with large effect sizes, the female effect size being larger than the male one, emphasising the earlier suggested female sensitivity to reciprocity related matters. The prediction was based upon the view that as honesty is fundamental to the formation of trust and cooperation with others (Rubin & Somanathan, 1998; Somanathan & Rubin, 2004), theft and lying would be more easily detected than failures to reciprocate, especially as it is possible to disguise the latter by reciprocating, but only occasionally (Trivers, 1971). The reason for the disconfirmation may lie in the nature of the material circumstances in which humanity lived over evolutionary time. Possessions were few,

perhaps almost non-existent, prior to the use of clothing, perhaps as long ago as 170 kya but possibly more recent (Toups, Kitchen, et al., 2011). As humanity evolved to its anatomically and psychologically modern state in the period prior to 200 kya (McDougal, Brown, & Fleagle, 2005), individual contributions to group wellbeing, upon which individuals' wellbeing was dependent, may have manifested themselves more in reciprocity related activities such as participation in group hunting and gathering than in the acquisition, retention and use of private possessions and non-acquisition of others'. The question merits more research.

Neither predictions 3a and 3b pertaining to the relative noticeability of the three adaptive topic groups were confirmed, none of the three adaptive groups' contradictions proving significantly more noticeable than those of the other two. Given that all three adaptive groups were found to be more noticeable than the Non-Adaptive one, this result is not regarded as particularly consequential or disappointing. Neither predictions 3a or 3b were based upon any strong theoretical basis, but were included to provide a basis for exploratory comparisons. This research was based upon particular choices of adaptive topics as representative of these groups, and within the topics, particular triplets of stories, not all of which proved to have equally noticeable contradictions. It is possible that other groups and other topics within these groups and other stories within these topics would produce a different result. However, the results of the current research suggest that whatever combination was devised, provided a wide range of adaptive contradictions were tested against Non-Adaptive ones, the adaptive ones would prove more noticeable as a whole.

Summary.

In Chapter 2 - Theoretical Background, three topic-group-related predictions, 1a, 1b and 1c, were made which were based upon the hypotheses that:
CD is a product of evolution by natural selection and natural selection has produced a bias in the human ability to notice contradictions according to topic, adaptive contradictions being more noticeable than Non-Adaptive ones.

The second hypothesis is based upon the first. The second hypothesis is directly testable while the first one is not. Studies 3 and 4 demonstrated these three topic related predictions have all been confirmed, verifying the second hypothesis. It is claimed that the verification of the second hypothesis provides evidence for the first hypothesis: CD is the product of evolution by natural selection.

Chapter 7 - General Discussion

"Nothing in biology makes sense except in the light of evolution". So pronounced the evolutionary biologist Dobzhansky (Dobzhansky, 1973, p. 125). Although he was writing of the unity of living things, the same phrase modified to "Nothing in psychology makes sense except in the light of evolution" is equally explanatory if applied to many attributes of the human psyche. It has been an underlying assumption of the current research that only an evolutionary approach can provide the reasons for their existence.

Numerous areas of research into cognitive discrepancy detection have used their own particular stimuli and have searched for and found associated effects. Cognitive Dissonance is a typical example but is only one of the wide range of psychological domains encompassed within the anxiety-to-approach model proposed by Jonas, McGregor, et al. (2014). In contrast to the current research which seeks to demonstrate that CD is the product of natural selection, the proposal of Jonas, McGregor, et al. (2014) is notably descriptive, as are some of the researches they cite. Although the term "adaptive" appears several times in the article, it does not explain the natural selection processes by which their mechanism came into existence. This left a gap for other research.

Support for the view of Jonas, McGregor, et al. (2014) that the wide range of reactions to discrepant cognitions are manifestations of the same underlying mechanism is provided in a recent paper by Randles, Inzlicht, Proulx, Tullett, and Heine (2015). They adopt the term 'fluid compensation' (Alport, 1943, p. 466) to refer to discrepancy detection in one area of cognition giving rise to an adjustment in an apparently unrelated one. Randles et al. (2015) use classical CD experimental paradigms, induced compliance and forced choice, to produce affirmation and abstraction effects that have previously been found in other areas of cognitive discrepancy research, for example, the Meaning Maintenance Model (Proulx & Inzlicht, 2012). The affirmations related to enhanced declarations of a moral attitude and belief in the existence of a supernatural entity, and the abstraction to enhanced motivation to see patterns where none existed. In four experiments, they demonstrated both the conventional CD and the predicted

affirmation and abstraction effects with large sets of participants from a broad range of cultural backgrounds. Their experiments are probably only the first of many to demonstrate the credibility of the Jonas, McGregor, et al. (2014) model of discrepant cognitions processing. In the current research, stimuli were either adaptive and non-adaptive, and within the adaptive, different effect sizes were demonstrated, suggesting that adaptive stimuli spanning a spectrum of potency are discoverable. This contrasts with the literature of TMT, MMM and the other discrepant detection research which describe experiments in which a single type of discrepant stimulus is used and evidence of a particular effect sought. The results of the current research combined with that of Randles et al. (2015) suggest that a set of stimuli spanning a spectrum of adaptive relevance should be used to detect a variety of the five A effects described by Proulx and Inzlicht (2012) and that these effects would be of measurably different sizes. Evidence that the strengths of these different effects varied with the adaptive relevance of stimuli would support an evolutionary explanation for the model's existence. In the remainder of this discussion, it is assumed that this will be the case and that the general correctness of the model will be accepted in the research communities. As a consequence, numerous questions arise some of which are addressed in the following subsection.

Generalisations from the current research

Several generalisations can be based upon the results of the current research if it can be assumed that in the exigent and dangerous circumstances of our evolutionary past, a substantial proportion of knowledge acquisition was oriented to the SRR topics. First, in the process of acquisition, new information would be tested against acquired information previously found to be trustworthy. Any discrepancy implied a threat in the sense that the new information conflicted with that which had proved reliable in the past. Second, an equally important function of the mechanism was to ensure that discrepancy detection led to action. Evolving a discrepancy detection ability on its own has little point as without consequential action, there would be no purchase for natural selection to exploit. One can therefore hypothesise that the mechanisms that detected the discrepancy, generated a motivational aversive state and initiated the consequential action all evolved simultaneously. Given the substantial volume of neural tissue needed

in humans for these connected activities, a very lengthy period of time would have been required.

Such argumentation supports the view, expressed in Appendix B - The Evolution of Linguistic Abilities, that humanity's language ability evolved over many hundreds of thousands, perhaps millions, of years. Merely conveying information between people may not have been enough to justify its evolution, although its communicative function has been claimed to provide an adequate benefit (Pinker, 1994, p. 354). Without an ability to assess new information via comparison with previously acquired information known to be reliable, we would have remained infinitely gullible, accepting anything that was part of the current inflow.

The evidence outlined in Chapter 2 -Theoretical Background suggests that the comparison between discrepant cognitions may take place in the ACC, given that a large variety of discrepancy detections are accompanied by ACC activity (Proulx, Inzlicht, & Harmon-Jones, 2012). Recent research, for example, Fedorenko and Thomson-Schill (2014) and Mesulam, Thompson, et al. (2015) strongly suggests that the storage of concepts and of other forms of knowledge is widely distributed in the brain. The CD-related activation in the ACC and the wide distribution of concept storage raise a number of possibilities. Representations of thoughts or concepts may be carried or transferred from a location of storage to the location of comparison, presumably the ACC. If this is so, in what form and how and what is the mechanism of comparison? On the other hand, comparisons could take place in those areas where the stored concepts are activated during cognitive processing. In this case, the comparison would be made between the representations of the two similar conflicting thoughts in the same area of the brain and a signal representing discrepancy transferred to the ACC in order that dissonance be experienced. Cognitive Dissonance appears to be the phenomenological correlate of activity in the ACC. In the first case, the brain/mind can be compared to a modern computer in which data is stored in multiple locations, on discs, in random access memory chips, and in cache stores, for example, while comparisons of data take place only in a processor with its limited number of registers. In the second case, all areas devoted to the storage of concepts would also be capable of making comparisons between the representations of similar but conflicting cognitions. The ACC might then function in a way analogous to a fire alarm which produces the

same output irrespective of where the fire is located as a consequence of being positioned at the centre of converging cabling. Such analogies may be dangerous, given so many differences between the architecture of the brain, von Neumann computers and security and alarm networks, but these high level similarities may not be absurd. The generation of a high amplitude N400 that accompanies detection of semantic discrepancies "likely generated by a highly distributed neural source including the temporal lobes" (Harmon-Jones & Harmon Jones, 2012, p. 353) gives some credibility to these ideas. Brouwer and Hoeks (2013) discuss plausible retrieval and integration mechanisms. This area of neuropsychology must be regarded as immature (Brouwer, 2014), and intriguing studies into the mechanisms of discrepancy detection can be expected in the next few years.

The implications of the evolution of a discrepancy detection mechanism via selection-based processes are far more extensive than one might gather from reading even the wide ranging papers of Proulx and Inlicht (2012), Jonas, McGregor, et al. (2014) and Randals et al. (2015). Arguably, crucial elements of culture and enterprise became possible only with the capacity to detect discrepancies between multiple information sources. Detection and assessment of opposing or at least competing courses of action can take place only if the advantages and disadvantages can be pitted against each other and oppositions detected and evaluated in strength. Scepticism can exist only when new information can be compared with trusted knowledge, leading to questioning and critical evidence-based thinking, the foundation of science, technology and much philosophy. Our powers to detect deception or indeed to initiate conscious deception are dependent upon our abilities to pose one thought against another. There is probably very little in today's knowledge-based world that could exist in its current form without an information discrepancy detection mechanism. Evolutionary psychologists should point out the contribution the evolution of this mechanism has made to all human life and, arguably, the inevitability of this evolution.

How old is the capacity for discrepancy detection?

The results of the current research described in Studies 3 and 4 in Chapters 5 and 6 provide strong evidence that CD (and the generalised capacity for cognitive discrepancy detection) is a product of natural selection and has evolved to promote the

passage through the generations of the information stored in the genes that enable it, via the relative reproductive success of the bodies the genes inhabit at any one time (Dawkins, 1976). Cognitive Dissonance can therefore be regarded as (a) one of a family of cognitive discrepancy detection phenomena deriving from a single mechanism and (b) having evolved over a lengthy period of development as proposed in Appendix B - The Evolution of Linguistic Abilities. Accepting that CD is one discrepancy detection phenomenon among several, the question arises: when did the underlying neurological machinery first evolve? Some evidence suggests that the anxiety-to-approach mechanism may be based upon a neural substrate present in mammals millions or even tens of millions of years ago. Although the violation of expectation literature in non-human animals appears to be sparse, evidence for surprise and expectation violation has been found in apes (Bräuer & Call, 2011), dolphins and sea lions (Singer & Hendersonm, 2015), and perhaps rats (Baumeister, 2005, p. 196). The existence of this capacity in non-human animals would indicate a very old mechanism capable of at least coding and storing current perceptual experiences and comparing them with earlier ones. Further research in non-human animal behaviour may confirm a very early date for the origin of this mechanism. It is suggested that the Jonas, McGregor, et al. (2014) mechanism, which compares discrepant experiences defined in words, may have recruited an older mechanism that stored and compared new and recent non-verbal experiences. This supposition is supported by the similar architecture of the ACC of humans and non-humans (Devinsky, Morrell, & Vogt, 1995; Shackman, Salomans, et al., 2011). If further investigation of this idea proves fruitful, it will demonstrate that during its evolution, CD and features of our linguistic ability recruited a pre-existing mechanism.

Weaknesses in the methods employed to demonstrate the effects of CD

Forced Choice Experiments.

Randles et al. (2015) challenge a variant of CD theory that concentrates upon the self (Aronson, 1991), by citing “dissonant researchers” who employed animals and young children as subjects. Egan, Santos, and Bloom (2007) used macaque monkeys and young children who are assumed not to have a well-developed view of self, and Lydall, Gilmour, and Dwyer (2010) used rats. These citations are problematic. The

forced choice methodology used by Egan et al. (2007) was challenged by Chen and Risen (2010). In Egan et al. (2010), the same team report a replication of the results of their 2007 paper but no longer claim CD as the explanatory mechanism. Lydall et al. (2010) in their effort justification experiment are explicit in their rejection of CD as the explanatory mechanism and cast doubt on all experiments that previously used the effort justification paradigm to demonstrate dissonance effects. Lydall et al. (2010) cannot be referred to as "dissonance researchers" as they are in Randles et al. (2015). Harmon-Jones and Harmon-Jones (2008) report on Brehm's (1956) experiment using the free/forced choice paradigm attacked by Chen and Risen (2010). The publication date of their paper, 2008, makes it clear that they could not have responded to the challenge represented by Chen and Risen (2010), but an experiment reported by Harmon-Jones et al. (2015) in which they fail to cite either Chen and Risen (2010) or Egan et al. (2010) leave their readers wondering whether their results can be considered to provide evidence of CD.

The previous paragraph illustrates a debate still being conducted amongst CD researchers. While the Chen and Risen (2010) view has attracted support (for example Gawronski, 2012; Johanson, Hall, Tarning, Sikstrom, & Chater, 2014), other researchers, (for example, Izuma, Matzumoto, et at. 2010; Kitayama et al. 2013) have conducted forced choice experiments using fMRI equipment which detected activity in the ACC. These latter experiments suggest that dissonance may be being felt in forced choice experiments

The debate is a consequence of the acceptance of indirect evidence that CD has been experienced. Effects such as changes of attitude or preferences, rationalisations of decisions, and changes of perceptions are noted and measured. However, with the exceptions of Elliott and Devines's (1994) experiment in which they used their Dissonance Thermometer, participants are not asked to give any indication of their state of discomfort. The measured effects are assumed to be the consequences of an undocumented state of discomfort following induced discrepancy detection but the opportunity to ascertain the emotional state of the participants is not exploited. The documentation of a possible state of discomfort was included in the current research in two ways. The third question of the four following each story was always framed in such a way as to provide an opportunity to reveal whether the contradiction between lines 3 and 10 had been noticed. Some participants did use it for this purpose and some of the answers revealed a state of discomfort with words such as 'conflicted!'

‘conflicting statements!’ ‘Very important and then not important?’ ‘i dont know contradictory evidence’ and ‘ambiguous’ being used in the answers. During data gathering for both Studies 3 and 4, a brief description of the purpose of the research was given after participants had completed their contribution. This gave them the opportunity to reveal their reactions to the contradictions that they remembered noticing. As reported in Appendix H - Contradictory Paradigm Assumption Testing, words such as ‘confusing’, ‘puzzling’, ‘weird’ and even ‘contradictory’ itself when used with the appropriate tone of voice, indicated that they had indeed been experiencing the discomfort normally associated with the term CD. Observation of the exclamations, grimaces, posture changes, and head and body movements of the participants helped to reinforce the impression that when they were reading the contradictory stories, they really were feeling the discomfort of CD. Head and body movements do communicate a state of mind (Livingstone & Palmer, 2015).

The experiments conducted in accordance with the contradiction paradigm appear to fall into two categories; (1) those that do not explicitly investigate the feelings of participants and which are now being challenged as to their ability to demonstrate CD as the explanatory mechanism, and (2) those that do provide opportunities for the expression of participants’ feelings of discomfort. This contrast suggests that all future research in which it is claimed that CD is the causative explanatory mechanism should be designed to provide some avenue for participant expressions of incidental discomfort. These expressions would provide evidence that CD is indeed the explanatory mechanism for the phenomenon being measured.

The importance of the separate female and male analyses

Although tests for the statistical significance of the variation of the differences between male and female topic RTs failed to show conventional significance, the separate analyses did provide one interesting comparison. While the difference in effect size between the two sexes’ comparisons for Predictions 1b (Reproduction) and 1c (Reputation) may be of no importance, the difference in the Survival analyses was surprisingly large. The comparison with the Non-Adaptive group for female participants did not demonstrate a significant difference while the male comparison did, with a large effect size. While a tentative explanation for this was provided in Chapter 5 - Study 3,

the result should be regarded with caution and suggests further research designed specifically to explore this surprising result.

The differential importance of the topics in the validation of Predictions 1a, 1b, and 1c

The seven adaptive topics contributed very unevenly to the success of the main predictions. While the contributions of each of the Survival topics to the comparison with the Non-Adaptive Group were reasonably close, this was not the case for the other two adaptive groups. In the case the Reproduction Group, the contribution of the Child-Loss and fidelity topics was much greater than that of the Family topic. For the Reputation Group, the Reciprocity contribution far exceeded that of the Honesty topic. Despite some topics' RTs being predicted to be greater than others, these were surprising results. A more even set of contributions had been expected. If further research in this area is ever conducted to overcome some limitations of the current research (see below), these contrasts could be further explored.

Limitations of the current research

Although the experiments conducted in the current research all support the motivating hypothesis that CD is a product of natural selection, the studies manifest a number of limitations which could be overcome in future research.

Assumptions underlying the Contradiction Paradigm.

The Contradiction Paradigm does not typically include any specific measure of the subjective feelings of the participants. Rather it measures the effects of the participants noticing the contradiction, the extended Read Time, and assumes that the participant feels the same aversive cognitive dissonance as has been measured in some other paradigms such as the Forced Compliance one (Elliot & Devine, 1994). The issue is discussed in more detail in Appendix H.

Lack of standardised contradictoriness.

There is no set of contradictory pairs of sentences that is recognised as a standardised tool, e.g., all conforming to the rules of De Marneffe, Rafferty and Manning (2008), and widely used in a number of contexts within the discrepant cognitions research community. This absence forces researchers to generate their own and submit them to individual tests of contradictoriness, such as the 6+ threshold used in this research, which are non-standard and ultimately arbitrary. It may be impossible today to conduct research in any other way, but the lack exposes this area of research to charges of immaturity and caprice.

Three stories to represent each topic.

Although the rationale for requiring only three stories to represent a topic is provided in Chapter 3 – Study 1, Item Construction, it is unfortunate that this limitation was necessary particularly since it was discovered that the RTs for the three stories for some topics showed differences that were statistically significant. Although an attempt has been made to explain these in Appendix N - Triplet ANOVA Analyses, it would be preferable if more effort was dedicated to testing a larger number of stories per topic to reach a standardised level of contradictoriness and subsequently tested for RTs in pilot sessions. This would permit later research to use a larger number of stories that are known to be more similar in both respects to represent each topic. A greater degree of standardisation could be achieved with greater reliability of results.

Limitation of number of topics to represent areas of interest.

The practical issues also imposed limitations on the number of topics that were used to test areas of interest. The Survival Group was represented by life-threatening and shelter related stories but could have been represented also, for example, by stories featuring food shortages. The Reproduction Group of topics was represented by child loss, family and fidelity, leaving out sexual attractiveness and fertility which are also relevant. The Reputation Group was represented by only honesty and reciprocity but could have included loyalty. The Non-Adaptive Group was represented by driving regulations (socially important) and paper things (socially unimportant) but a wealth of

topics other than these such as clothing conventions (socially important) or type fonts (socially unimportant) could have been included as bases for comparisons. In general, the wider the bases of comparison, the more robust the results.

The mono-cultural nature of the research.

Although some participants were of overseas origin, assessed by their having a first language other than English, the research must be regarded as mono-cultural. The percentage of non-New Zealanders was small at 5.5 % and the difference in RTs for this small fraction, although longer, rarely reached statistical significance due to the small numbers and higher standard deviations. In any case, lengthy RTs were eliminated from analyses by the outlier rules. In addition, from their physical appearance, the New Zealand participants were judged to be largely of European origin, although no attempt was made to confirm this via the personal profile information. Cross cultural considerations were not part of this research.

Since the hypothesis for this research is based upon natural selection, which has universal relevance, the results obtained in this research should also be obtained if similar research was conducted in other cultures. Confirming this prediction remains an outstanding obligation (Henrich, Heine, & Norenzayan, 2010).

Participants' backgrounds.

Participants were recruited in two categories of locations; lecture theatres and public areas of university campuses and dining areas of public shopping malls. Participants recruited in university premises are not a representative sample of the New Zealand population. A more representative sample was obtained by recruiting participants in the dining areas of two shopping malls, despite the greater difficulty. However, the general backgrounds of the participants recruited at the mall located in the Wellington central business district indicated that they were at the higher end of the educational spectrum. The other mall was in a residential suburb and may therefore have been more representative. Based upon the available literature (e.g. Deary, Strand, Smith & Fernandes, 2007; Rohde & Thompson, 2007; Source, Lennon, Lennon & Leandro, 2015), it's seems possible that there was a skew in the distribution of

participants towards the higher end of the intelligence spectrum. Although tests of educational achievement levels failed to produce statistically significant differences in RTs once outliers had been eliminated, it is hypothesised that effect sizes would have been different with a sample of participants more representative of the population as a whole. Confirmation of this hypothesis remains a further obligation.

Alternative Interpretations

The predictions of bias in humanity's sensitivity to contradictions discovered in the experiments reported in this thesis were based upon the logic of Darwinian natural selection. The experimental results and the cognitive discrepancy literature suggest that CD is one of a family of psychological attributes that evolved as a consequence of threats to survival and/or reproductive success in ancestral times. Given the mechanism suggested by Jonas et al. (2014) and the experiments reported by Randles et al. (2015), this family of phenomena probably all made a contribution to the selection pressure for the evolution of the Jonas et al. (2014) mechanism. Although this claim may be the most economical explanation for the existence of CD and the bias in its manifestation, it may not be the only one.

Gould and Vrba (1982) proposed that the term 'exaptation' could be applied to a biological phenomenon that had evolved, not because it conferred any survival or reproductive benefit itself, but as a by-product of one or more others that did. The white colour of bones and the red colour of blood are often mentioned as examples, the white of bones being a by-product of the strengthening calcium content of bones and the red being a by-product of the oxygen carrying capacity of the haem-based constituents of blood. In principle, the term might also be applicable to psychological attributes. Both CD and a bias in its manifestation could be exaptations of other phenomena. Obvious possibilities are the other forms of cognitive discrepancy detection briefly describe by Jonas et al (2014) and others in this thesis. They have the advantage of pertaining to threats to survival. In addition, there could be other antecedent attributes unrelated to these. One such possibility is a speech correction mechanism, which probably evolved as part of the evolution of our linguistic abilities. When listening to human speech, we form a neural representation of what we hear in order to respond to it. This capacity is invoked not only when we listen to others but also when we hear ourselves talking.

Occasionally we say something other than what we meant to say, sometimes even the opposite. On these occasions, there is a cognitive discrepancy between what was said and what was intended. Our cognitive discrepancy detection ability is invoked and action is taken to correct our speech to more accurately reflect our intention. Both these facets of our speech correction mechanism, discrepancy detection and subsequent action, map on to the action-based model of CD describe by Harmon-Jones and Harmon-Jones (2008). One disadvantage of this proposal is that, at first sight, it would not appear to predict a bias in sensitivity in discrepancy detection according to topic. The mechanism may well be domain general because it must make comparisons between an unlimited set of contradictory self-generated pairings and must do so successfully and quickly irrespective of both content and context. However, a detailed examination of a large corpus of recorded conversations could reveal a bias. Conceivably we correct some types of speech errors more than others and do so possibly because some errors are more consequential than others. Only considerable further research will determine whether any of these several possibilities, alone or in combination, present a challenge to the explanation for the existence of CD proposed in this thesis and the bias in its manifestation revealed in the experiments described in Chapters 5 and 6.

Strengths of the research

Threshold of 6+ contradictoriness.

The 6+ threshold required of contradictory pairs to be taken to the later stages of the research proved to be far from easy to achieve. Several pairs thought to be as contradictory as others that did meet the threshold were rejected, without any detectable cause. The difficulty in obtaining them is measured by the need to employ as many as 13 panels of 60 participants, a grand total of 780 people. Because of these difficulties, the threshold engenders confidence in the rigour of the experiments even if it is arbitrary.

Reproducibility of results.

A recent study of the reproducibility of psychological research results (Open Science Collaboration, 2015) found that only 39% of a sample of previously published

psychological research results could be subjectively judged to have been reproduced when the original experiments were repeated by teams of independent researchers. The likelihood of the original research results being successfully reproduced correlated highly with the effect sizes of the original research. The authors of the 2015 study draw attention to the higher frequency of within-subject designs used in cognitive psychological research relative to social psychology. They note that this design often has a greater power than a between-groups design to detect effects with a given number of participants.

These Open Science Collaboration (2015) findings, although dismal for psychological research as a whole, are encouraging when considered in connection with the current research. In this research, two within-subjects experiments were conducted with independent sets of participants, both having high N at 201. Both were aimed at demonstrating the same effect when defined in general terms, namely, that the mind is biased in its ability to detect contradictions, it being more sensitive to adaptive contradictions than to Non-Adaptive ones. Both demonstrated this to be the case. It can therefore be claimed that Study 3 first demonstrated the truth of the underlying hypothesis and that Study 4 replicated the result of the first. Although some of the effect sizes were small, most were moderate or large.

In sum, the current research is characterised by high participant numbers, many large effect sizes, the within-subject design and a replication of the results of the first experiment. This combination suggests that the current research passes the Open Science Collaboration (2015) test of reproducibility.

Conclusions

This research was inspired by the view that all universal human psychological attributes exist because they produced reproductive benefits during humanity's evolution. While this view has probably been widely accepted for many decades by all psychologists of a Darwinian persuasion, for some attributes, such as CD, there appeared not to be any evidence to demonstrate how the mechanisms of natural selection could have given rise to their existence. This research was designed to provide this evidence by showing that there is an enhanced sensitivity to adaptive

contradictions, which, according to the available literature, leads to action to resolve them (Harmon-Jones and Harmon-Jones, 2008) frequently taking the form of questioning aimed at obtaining better quality information (Graesser & McMahen, 1993). It is claimed that the current research has done so although CD is no more than one example of cognitive discrepancy detection. Randles et al. (2015) have produced strong evidence that CD is just one manifestation of an underlying cognitive discrepancy mechanism at work, almost certainly that of Jonas, McGregor, et al. (2014). In providing evidence for CD's evolutionary origin, the current research provides evidence for the origin by natural selection of this mechanism and all its other manifestations. In doing so, it begins to fill a gap in the evolutionary psychology literature.

Appendices

Appendix A – Discourse Comprehension

In this appendix, the discourse comprehension literature is reviewed with a view to generating rules to guide the construction of stories encompassing the tested pairs of contradiction. As far as possible, there should be no reason for participants to hesitate after having read a contradictory line 10, other than having noticed the contradiction and feeling CD. The rules for the construction of the stories, deduced from this review of the discourse comprehension literature, appear at the end of the appendix.

Modes of Research

In recent decades, research into language and discourse comprehension has taken two major directions (Louwerse & Jeuniaux, 2008); (1) viewing comprehension as modal, employing specific sensory modalities, for example, visual or auditory, and embodied or grounded in cerebral representations on the one hand and (2) symbolic or a-modal on the other. In the former case, researchers investigated whether multiple modes of psychological processing, and by implication activity in their neurological substrates, were invoked during language comprehension. For example, some early researchers, for example, Klatzky Pellegrino, McCloskey, and Doherty (1989), produced evidence that "the motor system plays an integral part in facilitating interpretation of phrases describing hand/object interactions" (p. 73). More recent researchers have been able to employ fMRI instruments while participants read texts or viewed images or objects. Burin, Acion, et al. (2014) have gathered evidence that the ventromedial prefrontal cortex is activated when inferences are being made about the socio-emotional aspects of verbally described situations. Indeed, the purpose of these researchers has been to establish that symbols such as words or images are grounded in bodily (neurological) representations of their referents and that purely symbolic manipulation in the brain is inadequate as a complete explanation of comprehension (Chow, Mar, Yisheng, Siyuan, Suraji, & Braun, 2014; Horshak, Giger, Cabral, & Pochwatko, 2014). These researchers may be said to be exploring the cerebral geography of comprehension and have convincingly established that the understanding of symbols such as lexical items in texts is necessarily accompanied by the activation of neurological representations of the same concepts.

On the other hand, symbolic or a-modal researchers have investigated external responses to more complex stimuli, such as bodies of text (McKoon & Ratcliff, 1986, 1992, 1995). They have concentrated on understanding the nature of the sequential processing that might be taking place within the brain and often take a computational approach (Kintsch, 1988; Myers & O'Brien, 1998). Their emphasis has been on the representation and processing of a body of text, rather than the effects of hearing isolated words or viewing particular objects, as has been the case for the embodiment school. Computer models simulate a presumed representation in the brain of both the entities mentioned in the text and their interrelationships as portrayed by the grammatical structures of the sentences. This 'model' can then be processed to make inferences and answer questions (Kintsch, 1988).

Because the current research has employed bodies of text in the form of short stories, it is the latter branch of research that is the more relevant. While the embodiment of textual understanding is not denied, it is the symbolic or a-modal branch of the research literature which has provided the relevant guidance. Without an investigation into the processes that appear to take place in the brain during discourse comprehension, it might be easy to incorporate unintentionally several features into the stories' texts which would act as confounds. Accordingly, the a-modal literature has been examined with a view to extracting rules of story composition aimed at minimizing such confounds. A review of this literature appears below and it is followed by a list of the rules that have been extracted.

Preliminary points

In testing sensitivity to contradictions dealing with survival, reproduction and reputation (SRR) via participants' readings of text-borne contradictions, little experimental noise is likely to be introduced if the contradictions are phrased explicitly, that is, the contradictions conform to such straightforward forms as 'verb' and 'not verb' or are stated in equivalently explicit form. If expressed in non-concrete form, for example, contradictions requiring inferencing to be understood, any difficulty or subtlety of the SRR contradictions must be no more so than that of the non-SRR ones. Dissimilarities could introduce confounds. It is easier to ensure equality of

understanding by avoiding the need for inference. It is also obvious that for all topics, the intervening texts between the contradictory pairs be similar in complexity and neutral to the contradiction, both explicitly and inferentially. These two points have contributed to the identification of the rules of story construction.

The symbolic or a-modal research on text comprehension

Much of the current research assumes and tests a tripartite model of memory based largely on the proposal of Kintsch and Van Dijk (1978). Although the terminology of different researchers varies, a common surface level is reliant upon the components of Working Memory (WM) (Baddely, 2000; Baddely & Hitch, 1974). Parts of working memory contain representations of the exact words and therefore the phrases or sentences of the text. These can be accurately reproduced by the listener if required within a short time period, typically measured in the low tens of seconds. However, they are displaced by subsequent equivalent elements as the reader progresses through the text. A deeper level, termed the textbase or propositional level, consists of the meanings of the text, encoded and held in short term memory. The text propositions are understood in the sense that they provide answers to questions about the content of the text. However, this level's content reflects the meaning of the propositions rather than the text itself. Readers may not be able to reproduce the text some minutes after having read it. The third level is termed the situation model and includes not only the meaning of the text itself but also much other information associated with the entities and propositions in the text. These typically include information primed by the entities that are explicitly mentioned and inferences which cover a wide set of relationships between them. Creation and storage of the situation model in Long Term Memory is usually regarded as an indication of successful comprehension of the text.

In the various simulations discussed below, these three levels are provided with explicit form within the computer programs that model them. Terms such as *resonance*, *activation*, *focus*, and *persistence* are used to refer to those entities or propositions that are current, as opposed to those that have been extracted from the text or databases representing personal knowledge but which are no longer playing a significant role. The recent neuroscience studies, described in Chapter 2 – Theoretical Background, strongly

suggest that this latter terminology is useful in a literal sense rather than merely as a metaphor.

During the last few decades of the 20th century, a multiplicity of models of discourse comprehension were developed. Two of them, the memory or minimalist, and the constructionist, formed the basis of a classic academic dispute (Graesser, Singer, & Trabasso, 1994). Their differences lie mostly in the extent to which humans are believed to enrich the representation of the text in memory from their own knowledge and the extent to which they make inferences 'on-line', that is, automatically. More recently, there has been a unification in the form of the Constructionist-integration mode (see below). All the models concern themselves with establishing in the reader's mind text coherence but differ about how many and what types of inferences will be drawn by readers in order to achieve this. Short discussions of text coherence and inference types are therefore provided prior to a discussion of the models.

Text Coherence

Coherence refers to the relatedness and consistency of the propositions in a discourse. For a discourse to be coherent, each proposition within it must relate semantically and syntactically to those preceding and following it (Caron 1992), permitting it to be conceived as an integrated whole. Without coherence, propositions become a disjunctive list.

Local coherence.

According to Graesser et al. (1994) two textually presented propositions are locally coherent if they:

1. are adjacent or close to each other spatially in the text. They follow each other sequentially or are separated by only a small number of other propositions. They are consequently both read within a few seconds, and are both therefore likely to be held in short term memory simultaneously,
2. are semantically connected in some manner. Several means of connection are suggested, causal connectives such as 'because' are used, both contain the same entity name, or one proposition is anaphoric with respect to the other, one

proposition containing the referent of a pronoun used in the other, or uses synonymous or general terms to refer to an entity mentioned previously.

3. require only case structure inferences if 2. does not apply (see below)
4. are consistent with each other. Neither contradicts the other nor is inconsistent with the other, the latter meaning that inferences can not be drawn which contradict one or other of the propositions.

Zwaan and Radvinsky (1998) suggest that local coherence is broken whenever any of the protagonist or protagonist's motivation, the time, cause or location of an event featured in an incoming piece of text is inconsistent with the current contents of WM. Such loss of local coherence manifests itself as an increase in reading time of the second and inconsistent proposition. Van Den Broek and Gustfson (1999) add the role of protagonists' goals in the maintenance of coherence. However, these appear to be very similar to Zwaan and Radvinsky's (1998) motivations.

Global Coherence.

This term refers to an overarching theme that can be extracted from the set of individual propositions. A text is globally coherent to the extent to which all or most of the propositions contained in the text contribute to and refrain from detracting from a main theme. This latter may not be explicitly mentioned in the text at all, and in such a case, would be identified only by inference. Concepts participating in local coherence may be interrelated to form and contribute to higher order concepts. Global coherence may not be obtainable if pre-requisite background knowledge is not available.

Types of Inference

The literature reveals that there is no broadly based acceptance of any particular taxonomy of inference types (Singer, 2007), despite the use of such terms as "unified model" of inferences (Schmalhofer, McDaniel, & Keefe, 2002). This taxonomic non-existence pertains to two aspects of inferencing:

1. the various subject matters of the inferences,
2. the degree of difficulty that the relevant surface texts imposes.

McKoon and Ratcliff (1992) define an inference as "any piece of (relevant) information that is not explicitly stated in a text" (p. 440). They accept that this definition is very broad as it encompasses a gamut extending from ascertaining the referent of a pronoun to identifying the murderer in a detective novel from a wide variety of clues, prior to the author's dénouement. This definition fails to identify categories of inferential difficulty, and its associated processing burden, and suggests the impossibility of defining categories of inferences. Despite this, they note (p. 441) that some inferences seem to be made automatically without any awareness that inferencing is taking place, and are presumed to be automatic, that is, achieved without conscious thought. They appear to be in agreement with Graesser et al. (1994) in suggesting that other inference types are of a problem-solving nature and require conscious effort. For example, Graesser et al. (1994) describe domain general inferencing tools such as propositional calculus, predicate calculus, theorem proving, mathematics and statistics as "a struggle" (p. 376) and suggest that they would not normally play any part in on-line inferencing. Although McKoon and Ratcliffe (1990) indicate that some types of inferencing take place more quickly than others, there appears to be no comprehensive research covering all types of inferencing which lists them according to difficulty. In the apparent absence of such comprehensive research, a list is provided below, sequenced in accordance with a 'common-sense' view of increasing difficulty.

1. Word Overlap: these are inferences that encode two instances of the same word as referring to the same entity.
2. Pronoun resolution: this form of anaphoric inference establishes the relation between a pronoun and its referent.
3. General Term inference: this form of anaphoric inference establishes that the general term 'the boy' in one proposition refers to a specific entity such as 'John' appearing in a nearby previous one,
4. Causal Inferences, sometimes regarded as a form of bridging inference, (see below): these are inferences that establish a causal connection between one proposition and another. One event occurs because of a preceding one. There are two types: physical and motivational, the former referring to a mechanism and spatial arrangements and the latter referring to an animate state of mind, implying intention.

5. Bridging inferences; these inferences, which "reliably accompany comprehension" (Singer, 2007, p. 347), supply linkages between two propositions to establish a semantic continuity that cannot be supplied by any of the previously mentioned inference types. In the pair of sentences "Police are hunting a man in hiding. The wife of Bob Birch disclosed illegal business practices in an interview on Sunday" (McKoon & Ratcliff, 1992, p. 444), a bridging inference, that the man in hiding is Bob Birch, has to be made to connect the two propositions. In this process, the two propositions cohere. Schmalhofer, McDaniel, and Keefe (2002) use the term Bridging inferences to refer exclusively to inferences that connect the just-read proposition to previous ones and therefore regard them as backward inferences.
6. Pragmatic Inferences are made on the basis of the reader's general knowledge and have an explicatory function. Reading that someone is pounding a nail may lead someone later to believe that a hammer was explicitly mentioned (Caron, 1992, p. 126) They are sometimes referred to as elaborative or expository inferences.
7. Predictive or Forward inferences (Singer & Ferreira, 1983) in which readers infer the consequences of events described in the text. These are not required to contribute to text coherence unless consequential events are explicitly mentioned. Schmalhofer, McDaniel, and Keefe (2002) noted that forward or predictive inferences are based upon the just-read proposition and a relevant knowledge base. Some authors, such as Graesser et al. (1994), have noted the absence of these in their research, and common observation appears to confirm that determining the future implications of an event often does not take place.
8. Global Inferences refer to a wide range of possible inferences that might establish initiating settings, characters, goals, including those of the author of the text, events and outcomes, sometimes collectively referred to as a story grammar.

Graesser et al. (1994) present a list of 13 inference types which broadly fall into 6 categories: referential inferences, case structure roles, causal relationships between events, emotional states of the protagonist, the goals of the protagonist, and the belief system or values of the protagonists or author. The term 'case structure roles' refers to the roles that a noun, pronoun or noun phrase plays within a sentence or clause, for example, the agent (subject in surface grammar), object, indirect object, instrument, recipient, beneficiary, adessive (the locative case), etc. (Fillmore, 1968). If a text

indicates that a man is stationery and fishing, the inference that he is using a fishing rod would be an example of an instrumental case structure inference.

Working Memory Capacity

The concept of Working Memory proposed by Baddeley and Hitch (1974) was modified later (Baddeley, 1995, 2000) to extend the concept to include a central executive containing various types of content but still of relatively low capacity. Ericsson and Kintsch (1995) retained the concept of working memory but considerably extended it. They proposed that during tasks for which the necessary skills have been developed, among which they include discourse comprehension, people exploit what they call Long Term Working Memory (LT-WM). This entity exploits control processes that encode important information in conventional Long Term Memory in retrievable form and employs retrieval structures which enable rapid access. Beside discourse comprehension, they cite research in such fields as chess, mental arithmetic, remembering multiple restaurant dinner orders, and medical diagnostics that support their concept. Just and Carpenter (1992) and Gathercole (2007) have proposed a more general concept of individual differences in Working Memories capacity, drawing on the view that WM includes the processing of its arguments or content as well as their storage. This notion has received wide support in the literature. The processing encompasses language comprehension, requiring syntactic and semantic computations. Just and Carpenter's (1992) computer simulation of WM is realised by a set of productions (If ... then ... statements) that represent "activations". The contents of WM and the processing performed upon it are limited by the total of activation an individual has available. Individuals differ in their activation capacities and therefore in their performance in any task requiring the resources of the WM. Within any individual, there is an inverse relationship between the activation available for processing and storage; as more is demanded for one, that available for the other diminishes. By conducting prior reading span tests (tests which ascertain the number of last words of sentences read to them as a series that they can correctly remember averaged over several tests), Just and Carpenter (1992) were able to establish two sets of participants: those with high ($\geq$4.5) and low ($\leq$2.5) spans, high and low spans being used as a proxy for WM capacity. They applied a set of moving window reading tests to both groups with a view to ascertaining whether their indicator of WM capacity manifested itself in

a number of text comprehension tests and found that it did. High WM capacity participants fared better when answering comprehension questions on read texts, understanding garden path sentences, identifying ambiguous sentences as such, making use of semantic cues, maintaining extrinsic memory loads, and reading relative clauses more quickly. They applied their capacity model to distance effects discovered by previous researchers, namely that readers' abilities to make the first four types of inferences listed above are diminished to the extent that the two lines of text connected by the inference are separated by other lines. Inferencing performance diminished with increasing separation of lines but does so more quickly for those of low span/WM capacity than for those with high values. They note that "intermediate text that does not reinstate a referent causes forgetting with more forgetting by low span than high span subjects" (p. 134). The capacity model of WM is used below in the creation of the rules for the intermediate texts for the studies described in this thesis.

The Discourse Comprehension Models

Using Baddeley's (2000) terminology, all the models are based upon the supposition that an initial transformation takes place as representations of the text move from the visual-spatial sketch pad of WM into representations of the propositions and their arguments in the episodic buffer. The models also include the priming, in psychological terminology, or activation, in the terminology of computer modeling of memory (Collins & Loftus, 1975; Quillian, 1967) to varying degrees of long term memory contents that are related to those in WM. Two aspects of priming are important:

1. the relationships existing between the priming and the primed. These can be either semantic or phonological (Meyer & Schvandenveldt, 1971), that is, the priming and primed are related by meaning or sound, and
2. the nature of the activation or retrieval, much of it taking place "without strategic control" (Sparks & Rapp, 2010, p. 375).

Given the above, there is a fluctuation in the activation of the contents of the discourse. The activation intensity of particular characters, instruments, events, propositions, etc.

and those that are primed by these, varies as one set is replaced by another as the reader progresses continuously through a text.

Inferencing Models.

Two schools of discourse modeling thought developed in the latter decades of the last century. They differed mainly in terms of how much information was kept active in the memory-based text representation as readers progressed through the text, and therefore how many and which inferences were made and kept active. They are briefly outlined below.

The Memory or Minimal Inference Model.

As the main developers of the Memory or Minimalist model, McKoon and Ratcliff (1986, 1992, 1995) assert that readers retain in STM/WM only the most recent propositions and claim that in the absence of a goal directed strategy, only two kinds of inferences are automatically drawn during reading. These are:

1. "those that establish (if necessary) locally coherent representations of the part of a text that are being processed concurrently and
2. those that rely on information that is quickly and easily available." (McKoon and Ratcliff, 1992, p. 440).

Local coherence would typically require anaphoric and some bridging inferences to be made but no more. Thus, the minimalist model asserts that readers will not normally construct inferences that fully represent the situation describe in the text unless they have particular motivations for doing so. Implied predictable events will be encoded in memory only minimally. Information that is quickly and easily available includes that which is explicit in the text and relevant well-known (to the reader) general knowledge, including some schemata.

According to McKoon and Ratcliff (1992), contradictions if present in non-adjacent texts may not be noticed. In one of their experiments, one text "was coherent at the local level but local information contradicted global information. The data showed

no effects of this contradiction" (p. 446). Other results indicate that global coherence is sought only when local coherence can not be found, and then only probabilistically and slowly, taking over half a second to deduce. McKoon and Ratcliff (1995) claim that the Minimalist hypothesis has survived 20 years of research and that "no incontestable evidence has been found to support the notion that readers engage in the construction processes [see below] that would be required to generate a large number and variety of inferences" (p. 97). Unfortunately, some of the terms used to describe the Minimalist position are ambiguous. What information can be regarded as 'well known' is a question that raises obvious difficulties. It is clearly determined by individual's personal background, which makes experimental results harder to interpret, and of course, the term 'well' is particularly difficult to define.

The Constructionist Model.

The idea that memories are constructed is many decades old. From the inventions in participants' acts of memory recall which he observed in his experimental work, Bartlett (1932) concluded that memory is "an imaginative reconstruction" (p. 213) and "a great amount of what is said to be perceived is in fact inferred" (p. 33). He used the phrase 'active settings' to designate what today are called schemata, frames or scripts, sequences of activities associated with frequent everyday functioning, such as visiting a restaurant.

Graesser et al. (1994) criticise their 'constructionist' predecessors, such as Bartlett, for not specifying and predicting precisely what inferences are made and which are not, when narrative texts are being read. In defining the constructionist model in more detail, they assert that readers engage in a search after meaning and claim that from their experimental results, the following principles emerge:

1. Readers construct a meaning representation in memory which reflects and addresses their own goals,
2. They endeavour to ensure that this meaning representation is coherent at both local and global levels,
3. They endeavour to construct their representation such that it explains the actions, events and states mentioned in the text.

On the basis of these principles, the authors predict that if readers have the necessary background knowledge, are not reading the text for 'shallow' purposes, such as proof reading, and the text is globally coherent, then in addition to the few inferences that the minimalist model proposes, readers will further make the following types of inferences 'on-line' that is, while reading the text, rather than subsequently, which they term 'off-line':

1. where necessary, the goals, beliefs, emotional states, and values of protagonists will be inferred,
2. causal antecedent events will be inferred to explain the events in the text if absent from the text,
3. a global theme or objective may be inferred from the mentioned events,
4. the author's goal(s) may be inferred.

They also note that readers do not typically make inferences that predict the future or track the spatial positioning of people or objects featured, although other authors contest this, for example, Rapp and Gerrig (2006). Graesser et al. (1994) admit that there may be a continuum between on-line and off-line inferencing with inferences being coded 'to some degree' (p 376). Of relevance to this thesis, they claim that narrative texts parallel our everyday experiences and that it is adaptive to understand fully the events occurring in our environments, making it necessary sometimes to employ inferencing mechanisms (p. 372). In particular, they suggest that readers are sensitive to changes that convey information about discrepancies from normal states, violations of normative standards, conflicts, danger 'and abnormal occurrences that have adaptive significance to the organism' (p. 379). Implicit in the above is the view that readers would be more sensitive to changes or discrepancies that:

1. could affect them rather than affect the protagonists in a story,
2. could affect the SRR of protagonists relative to such changes which could not.

Graesser et al. (1994) claim that their constructionist view of readers' mental representation is required if full comprehension of a text is achieved. Such comprehension exists when there is conformity between the author's intentions, the explicit text, and the reader's constructions, the latter including the relevant inferences. Only then is it possible for readers to create summaries or paraphrases, ask relevant

questions, and detect contradictions. They also insist that comprehenders make particular effort to explain why the activities described in the text are occurring. They search for causal explanations of events and claim that this finding is compatible with causal attributions investigated within social psychology. It is also compatible with the inferred causality explored in perceptual processing (Scholl & Tremoulet, 2000) and the hyperactive agency detection device (Barrett & Johnson, 2003) discussed in the evolutionary psychological explanations of religion (Barrett & Lanman, 2008).

The Constructionist-Integration Model.

Kintsch (1988) proposed a model which incorporated a two stage view of the creation of the representation of a text leading to the formation of a situation model. It focused on the propositions and the entities of a text progressively encountered but also real world descriptive propositions and inferences which were based upon real world knowledge. He implemented this model in the form of a computer program which he successfully tested to reproduce experimentally observed results. The first stage is a parser which transforms the focal text into a textbase which reflects the set of propositions and entities featured in the former and the linkages between them. Thus the proposition ‘Mary baked a cake’ is transformed into

1. BAKE[Agent:PERSON, Object:CAKE]

2, PERSON[Mary]

with links between them and any other textbase expressions to which they are linked by common arguments or anaphoric reference, the strength of the links depending upon the degree of separation of the propositions from each other in the original text. An associative network is then constructed by adding a small number of propositional nodes, “its closely associated neighbours, from a general knowledge database” (p. 166). These nodes have very variable relevance to the main propositions, being added only because of some semantic or homonymous relationship to those already in the network. These added nodes would be linked to those already present with link strengths determined partly from the database and partly by the number of common arguments. For example, in the example above, the nodes

3. BAKE{ins:SUN,Objext:PART-OF-COUNTRY} and
4. LIKE[MARY,EAT[MARY.CAKE]]

would also be constructed from more generic forms, such as LIKE[agent.PERSON, object.EAT[PERSON, CAKE]], and connected as inferences although semantically not expressed in the original manifest text in this example. The 'promiscuous' linkage and activation of numerous connected nodes, some irrelevant and/or inconsistent, provides scope for a subsequent Integration phase employing matrix algebra operations upon the strength of the linkages. In these operations, linkages are pruned, with many propositions and/or entities losing their connection to the network. In the example above, proposition 3. would be pruned because of the paucity of connections to others, while proposition 4. would survive, as an inference, because of the greater number and strengths of connections to other activated propositions. As a text is progressively read, the model cycles through the construction-integration phases adding new nodes to the network from new text propositions and re-applying the pruning algorithm to eliminate the less relevant ones, with some inferences surviving and the situation model becoming ever richer. Although the 'computerese' representations of nodes and connections bear little resemblance to how these are represented in the brain, the model successfully reproduces the type of the inferences that are typically made by humans and the timing delays which occur between creating the representations and the generation of inferences.

Myers and O'Brien (1998) extended the Kintsch (1988) model by placing more emphasis on the activation of distant, early propositions in the later stages of model construction in response to the need to resolve anaphoric references or causes of actions and events. They were included in a pattern matching facility, the equivalent of Kintsch's (1988) matrix algebra operations, which ensures that they resonate as a function of the degree of match to the current input. The pattern matching facility in their model is perceived to model a cerebral resonating function. Their model reflected several variables that the authors believe influence the accessibility of anaphor, that is, the reactivation of antecedent information. Because their resonance model works by pattern matching, one of the variables is the extent of feature overlap. A second one is the degree of elaboration of the antecedent information prior to the anaphoric reference.

In the interval between the two, the original proposition or entities may or may not be inferred as featured in other events. The more that they are, then the easier it is to make the connection when an entity or event is referred to anaphorically. The other side of the same coin is the presence of entities or propositions that distract. If mention is made of distracting propositions in intervening statements, they diminish the ease with which subsequent anaphoric references are understood. Likewise, the distance: the greater the distance between the two, the more difficult the retrieval of the antecedent information. A complicating factor manifests itself if the anaphoric term contains more than one word. An adjectival modifier in the anaphoric term may be sufficient to activate the entire anteceding information, but only if the distance between the two is small. If large, the entire anaphoric term is required to activate the antecedent information.

A unique feature of these C-I models is the wide range of real world knowledge that may be activated. If a reader's real world knowledge included concepts that were closely related to the antecedent information, they would resonate at the same time as the antecedent information did, particularly if they shared arguments with it. For example, if the term 'bank' appears in the surface layer, both river and money related entities would be activated in the situational model. When the anaphoric term was read, readers may link it to the resonating term rather than the information mentioned explicitly in the text. Myers and O'Brien (1998) believe their resonance model is consistent with a number of experimental studies which explore the effect of contradictory information in the second or target statement. In these cases, if the second contradictory sentence broke global coherence, even if it did not break local coherence, readers would search background information to try to re-establish global coherence. This search manifests itself as a longer reading time for an anomalous sentence than a consistent one. One experiment reported in O'Brien and Albrecht (1992) explored the degree to which readers remain sensitive to the location of the protagonist of a story. They found that readers took longer to read a second sentence pertaining to the protagonist's location that was inconsistent with an earlier one than when the second one was consistent with it. However, this was the case only when the phrasing enabled the protagonist's location to be easily determined, which the experimenters interpreted to mean that readers will adopt the perspective of a protagonists if the text permits it.

Guidance for the current experiments

Methods employed.

It is noticeable from the a-modal discourse comprehension literature that a wide range of experimental methods have been employed. To some extent, this lack of uniformity must be presumed to explain the conflicting results obtained and the different conclusions drawn by various schools of researchers. Methods include comprehension tests, timed and untimed target word identification, timed question answering, and timed priming. The texts employed constitute another source of variation, differing in their subject matter, degrees of complexity and subtlety, length, cohesion, change of theme, and other characteristics. Additionally, the majority of the experimental participants are first year American psychology undergraduate students. The frequently noted (e.g., Henrich, Heine, & Norenzayan, 2010) lowered ecological validity of research conducted with such a narrow base of participants introduces further problem of acceptance.

Anomaly detection and subsequent activity.

In the McKoon and Ratcliff (1992) experiments reported previously, the inconsistency present in globally incoherent text was not noticed irrespective of the degree of local cohesion. Timed test word identifications did not differ in globally inconsistent texts from globally consistent ones, which they claim as evidence undermining the constructivist models. On the other hand, Huitema, Dopkins, Klin, and Myers (1993) did find that global incoherence was detected when participants read a test line that was anomalous with respect to an earlier line of text. Numerous other researchers have found similarly. Hess, Foss, and Carroll (1995) found global incoherence caused longer reading times than local incoherence, and Rapp (2008) found reading times of test sentences were longer when they conflicted with well-known facts than when they did not. Clemerson (2009) also found that reading times of sentences which contradicted sentences appearing six lines earlier in passages were longer than non-contradictory ones.

Graesser and McMahen (1993) noted that previous research had established that students, and by implication people in general, frequently fail to detect contradictions in texts. In their experiments, they investigated this quantitatively but also compared this failure with reactions to other forms of incoherence. They hypothesized that contradictory information would prompt more questions than textual changes such as deletions and insertions of salient or subtle irrelevancies and posited a sequential model of anomaly detection, question articulation and social editing to explain it. They conducted five experiments all of which confirmed their predictions. The current research assumes a wide application of this finding, namely, cognitive dissonance generated by the contradictions acts as a greater spur to the generation of questions, relative to the paucity of question generation following other types of textual changes. Graesser and McMahen (1993) also discovered that physical barriers and social circumstances (having to get up from the table and go to another room to ask a question) could significantly diminish the extent of the contradiction-triggered questioning, implying that feelings such as embarrassment could swamp and/or inhibit CD feelings or actions normally triggered by it. They concluded that their hypothesized three stage sequential actions model of anomaly detection, question articulation, and social editing was verified.

Inferencing and Timing

Bridging Inferences.

Haviland and Clark (1974, cited in Caron, 1992) provide evidence for the making of bridging inferences and the relative difficulty of these by comparing reading times for the same sentence in two different contexts. Participants were presented with the two following sentence pairs

1. We got some beer out of the trunk. The beer was warm.
2. We checked the picnic supplies. The beer was warm.

The reading times for ‘The beer was warm’ were 0.2s longer in the second case relative to the first, implying that the greater complexity of the second inference, that the picnic supplies included beer, imposed a greater processing load than the first one, namely, only that the word ‘beer’ in both the sentences in the first pair referred to the same beer.

Elaborative Inferences.

Singer (2007) notes a timing difference between a bridging inference and an elaborative one, that is, one based of the reader's general knowledge. Participants were asked to judge the truth of the sentence "The tourist used a camera" after reading one of the following texts:

1. Explicit: The tourists took the picture with the camera, the scene was more beautiful than he remembered.
2. Bridging Inference: The tourist took the picture of the church. The camera was the best he had ever owned.
3. Elaborative Inference: The tourist took the picture of the church. The scene was more beautiful than he remembered.

Participants required approximately identical times to verify the test statement for the explicit and bridging inference cases but needed about 250 ms longer to verify the elaborative inference text. He concluded that the bridging inference accompanies comprehension, that is, takes place on-line supporting the constructivist view of discourse comprehension.

The difference between the two research cases outlined above seems only to be the semantic distance, that is, the extent of the difference in meaning, between the two sentences that have to be linked. The greater the semantic distance, the greater the time needed to make the inference. However, there appears not to be an objective way of assessing this distance independent of such timing experiments.

Conclusions

The mutually contradictory results of the experiments reported in the discourse comprehension literature frustrate a unified view of the psychological processes taking place during the reading of texts. The many different experimental methods employed probably contribute to this disunity of view. Nevertheless, Zwaan and Radvansky (1998) summarised the discourse comprehension research at that date by noting that, despite McKoon and Ratcliffe's many claims to the contrary, sufficient evidence had

accumulated that text readers did, to varying degrees on given occasions, create a situation model that could be described as having five dimensions: protagonists, objects, space, causation/intentionality, and time. Protagonists and objects form the kernel of an evolving situation model and are the subjects of the majority of inferences, namely anaphoric ones. Most focus is on the protagonists, with attention on objects being dependent on contextual cues, with many such inferences being case based.

Much of human behaviour is goal oriented. Zwaan and Radvansky (1998) interpret the results of research by Graesser and colleagues to mean that readers do infer high level goals from actions and use them to provide coherence clues to subsequent actions or events as they are incorporated into the integrated situation model. Readers will causally relate the propositions as each is met, with local and global coherence being tested. To the extent they are limited by the information provided and their personal motivation, readers create a mental image of arrangements of the objects mentioned, a spatial structure. Finally, time inferencing will take place, sequences of action and events being determined by the sequence of propositions in the text unless qualified by time markers such as ‘before’ or ‘prior to’.

In an apparent attempt to bring the dispute between the two schools of researchers to an end, the claim is made by van den Broek, Rapp, and Kendeou (2005) and Sparks and Rapp (2010) that the memory and constructivist based views can be reconciled, the first describing the initial activation in memory of the concepts, propositions and simple inferences of the text, and the second one leading to the activation of the more complex inferences. The latter encompass the five dimensions mentioned above. From this article and others, it appears reasonable to conclude that the degree of inferencing and the richness of the situation model constructed depend upon readers’ goals, their size of working memory however defined, general aptitude, their expertise or knowledge of the subject recounted, and the richness and coherence of the text.

Rules and relevant references

From the resources cited above, the following rules governing the generation of the stories to be used in the current research were formulated. The rules should be used for both the SRR and non-SRR texts. Contradicting pairs should:

1. be separated by the same number of lines of distracting text (Myers & O'Brien, 1998).
2. have identical degrees of elaboration of the main proposition (Myers & O'Brien, 1998).
3. have the same number of arguments (Myers & O'Brien, 1998; Just & Carpenter, 1992).
4. have the same number and types of distractions in the intervening propositions (Myers & O'Brien, 1998).
5. not include inferences that predict the future (Graesser et al, 1994)
6. not pit agentic inferences against non-agentic ones (Barrett & Johnson, 2003; Barratt & Lanman, 2008).
7. avoid unequal reinforcing of referents in intermediary lines (Just & Carpenter, 1992).
8. avoid calling upon WM resources differentially in the sense of having different degrees of syntactic or semantic difficulty (Baddeley 2000, Just & Carpenter, 1992).
9. avoid having unequal degrees of clarity regarding protagonist's location if that is the subject of a contradiction (O'Brien & Albrecht, 1992).
10. have the same dimension of (in)coherence, selected from space, sequence, causality, protagonist, and objects (Zwaan & Radvinsky, 1998).

The following additional rules specific to the current research experiments, emerged from a consideration of the stories, post first composition, when they were being assessed for adherence to the above rules,

1. The intermediate lines between line 3 and 10 should not gradually change the theme of the story towards a justification of the sense of line 10, thereby reducing the impact of the contradiction.

2. Any significant difference in valence (interpreted to mean pleasantness or unpleasantness) between lines 3 and 10 should be avoided as this would act as a primer/trigger for the detection of contradictory statements.

3. Consequently, all line 9s should ensure neutrality of line 9. It should be descriptive only, not referring to feelings.

4. Line 9 should not introduce a jarring change of subject, thereby possibly introducing a confounding cause of lengthier reading time for line 10.

Appendix B - Evolution of Linguistic Abilities

In the Theoretical Background section, arguments were presented supporting the view that the human capacities for language and cognitive dissonance developed in tandem over a lengthy period. In this appendix, the evidence for this lengthy evolution, measured in terms of many hundreds of thousands of years, and more speculatively, in the low millions, is presented. However, it must be noted that the main findings of the current research, that humanity has a greater sensitivity to contradictions of adaptive relevance, are not dependent upon the correctness of the notion of a lengthy period of language evolution. However, all characteristics of humanity's linguistic abilities are more credibly viewed as the products of natural selection if it can be demonstrated that changes necessary to equip humanity with language have taken place over a lengthy period.

In Chapter 2 – Theoretical Background, it was established that dissonance is experienced when a cognition in the form of a statement, attitude, opinion, or action (all capable of being described in words) conflicts with another such cognition. The ensuing conflict alleviation is often effected via speech, e.g., trivialising the source of one of the cognitions (Simon, Greenberg, & Brehm, 1995) or asking questions (Graesser & McMahen, 1993). In such cases, there is a dependency of the resolution of CD upon language and the relevant cognitive abilities. If language developed slowly, it follows that the relevant cognitive abilities would also have done so. However, CD has a further dependency. There is no obvious reason that the mere existence of language would have led to the development of a capacity for CD. Cognitive Dissonance required its own selection pressure. It is the major hypothesis of this thesis that that selection pressure was the contribution to survival and reproductive success that feeling CD and then acting upon it was able to make: those that felt CD in a situation where conflicting options affecting survival and reproductive success existed and whose CD sensitivity to the conflict was higher than average were more likely than others to seek a resolution that was advantageous to their survival and/or reproductive success. In addition to aiding the gradual evolution of CD, this process would also bias its expression towards circumstances of relevance to survival and reproductive success. Those whose expression of CD was not so biased would make fewer contributions to the genome of later generations than those who were so biased.

Such biased resolution over many generations would slowly cause the accumulation and fixing in the human genome of the relevant genes. This process could commence as soon as linguistic abilities began to develop but hardly before, and could and probably would take place contemporaneously with the development of language. In principle, the more sophisticated the state of language at any one time, the more frequent the opportunities for CD to lead to adaptively beneficial decision making. At the very least, it can be expected that the rate of development of the capacity for CD would initially have been very slow, but likely accelerated as language facilities developed.

The point for this thesis is that CD is largely dependent upon the refinement of the cognitive abilities necessary for language to contribute to survival and reproductive success. However, language's contribution to survival and reproductive success is in turn dependent upon the feelings which motivate speakers to use language in ways that can produce survival and reproductive benefits. CD is one such feeling and its development would likely accelerate the use of language as an aid to survival and successful reproduction. It is proposed that this mutual dependency would have led to a contemporaneous evolution and strengthening of both. It follows from this reasoning that the earlier the development of the capacity of language to represent and misrepresent reality, the more time there has been for the development of the capacity for CD and any biases in its expression.

The dating of the evolution of the human language capacity can not yet be said to be settled. While some authors have claimed a recent date, e.g., around 100kya, (Noble & Davidson, 1996), Pinker (1994) suggests that "the first traces of language could have appeared as early as *Australopithecus afarensis*" (p. 352) about 4 mya. Notwithstanding this wide range and the various meanings different authors ascribe to the word 'language', evidence is accumulating that language, at least in primitive or proto-language form, probably predated the evolution of *H. sapiens,* generally accepted as having reached its current state by about 150 kya or earlier (Klein 2002). The evidence supporting this view is presented under five headings below.

Anatomical and Physiological Evidence

In the subsection below, a set of conclusions pertaining to language relevant anatomical and physiological facets of modern humans are briefly described, each illustrating the view of the relevant researchers that human language is not recent.

The sound frequency range in modern humans maintains a high sensitivity to signals from 2kHz to 4kHz, which is important to the interpretation of spoken language. Martinez, Rosa, et al. (2004) examined the influence of skeletal structures on the acoustic filtering of the outer and middle ears in five fossil specimens dating from Middle Pleistocene which dates from ca. 750 kya to 125 kya. They concluded that these hominins had skulls that were compatible with a human-like pattern of sound transmission through the outer and middle ear at frequencies up to 5KHz. This range contrasts with that of chimpanzees and most other anthropoids that lack this sensitivity.

All primates other than humans have an air sac attached to the hyoid bulla which helps them produce booming noises. De Boer (2012) created artificial plastic vocal tracts with and without the air sacs and allocated subjects to each design, asking them to listen to the vowel sounds produced when air was blown through them. Background noise was added to ascertain the ease with which each design best enabled the distinguishing of vowels. He found that the sounds produced by the vocal tracts without air sacs were distinguishable against a higher level of background noise than those with air sacs. In addition, he found that the air sacs made it more difficult to produce and distinguish consonants. He noted that at the time of *Australopithecus afarensis*, the hyoid bulla was still present but by the time of *Homo heidelbergensis,* about 2.5 mya, the hyoid bulla and the air sac had been lost. He concluded that a gradual loss of the air sacs permitted an increasing clarity of vocal communication. The loss implies a selection pressure for its elimination. The obvious adaptive advantage of linguistic communication is suggested as the likely factor.

Bickerton (1990) suggests that inhibition of vocalisation probably developed during the Afarensis period (4-3 mya) as australopithecines moved on to the savannah. Alarm

barks and involuntary food barks would have been likely to have endangered survival. A capacity for voluntary silence necessarily means simultaneously developing a capacity for voluntary control and use of vocalisations. “It seems likely that by the time of *habilis* [voluntary vocalization] had been achieved.” (p. 143). By this time, full bi-pedalism had been developed, that is, it developed during the Afarensis period, changing the orientation of the laryngeal tract from somewhat horizontal to more vertical and simultaneously allowing the tract to descend down the neck. In the former position, swallowing and breathing can occur simultaneously. This does not occur in a fully grown human, although it does in babies. Also, the basal cranium, the “hole” in the bottom of the skull through which the peripheral/bodily part of the central nervous in the spinal column connects with the brain, migrated from the back of the skull to the bottom of the skull as the skull rotated forwards to ensure that the eyes continued to look forward as body posture became more upright. These two changes have had the effect of enlarging the laryngeal cavity permitting a large variety of sounds to be emitted. Richmond and Jungers (2008) even suggest that this development commenced as long ago as 6 mya. It would appear likely that by Erectine times, sufficient enlargement had taken place to support proto language although this was not necessarily phonemic speech typical of Indo-European or Semitic languages. As when children are just beginning to speak, continually modulated but distinguishable noises are created, like the sounds of tonal languages. Such sounds could have been sufficiently distinct to designate a small vocabulary of lexical items, forming different proto languages as *H. erectus* populated the Old World. Even chimpanzees have a vocabulary of distinct noises which appear to be interpretable by both other chimpanzees and humans (Crockford & Boesch, 2005). A more recent discovery indicates that even some monkeys extract meaning from combinations of calls (Arnold & Zuberbühler, 2008) suggesting that the linguistic Merge operation in which two syntactic objects are combined to form a new syntactic unit may have originated far earlier than Erectine times and done so in multiple species in parallel.

Lieberman, Laitman, Reidenberg, and Ganon (1992) suggest that the Supralaryngeal Vocal Tract (SVT) of *Homo sapiens*, which is responsible for the production of human vocalisations of speech, is probably an index of the neural mechanisms and cognitive abilities necessary for speech also. They note that archaeological data suggests *Homo neanderthalensis*, a cousin species contemporary

with archaic *H. sapiens* (see below) "likely employed some form of language" (p. 464) and also that *H. neanderthalensis* probably did not have an SVT identical to that of *H. sapiens*. They conclude from this that speech was present in *H. neanderthalensis* but was not as efficient as that of modern humans. Aiello (1998, cited in Jackendoff, 1999). has contested even this charge of inefficacy.

Kay, Cartmill, and Balow (1998) propose that the size of the hypoglossal canal, which carries the nerves that supply the muscles of the tongue, may provide an indication of the motor co-ordination of the tongue, a greater nerve supply reflecting the evolution of speech and language. They measured the sizes of the canal in chimpanzees, Australopithecines and *H. habilis* and find them comparable. They conclude that these species were without language. However, the size of the canals of *H. neanderthalensis* and early *H. Sapiens*, measured at Skhul (Israel), in Kabwe (Africa), and Swanscombe (UK) dating to middle Pleistocene times are comparable to modern *H. sapiens*. Because of the hypoglossal canal size of Neanderthal and human skulls dating to the middle Pleistocene are similar, they conclude that human vocal abilities may have been essentially modern by a date earlier than 300 kya.

Lee and Wolpoff (2003) graphed the gradual increase in encephalisation of early hominin species to modern *H. sapiens* during the Pleistocene (see below) and concluded that their results are "incompatible with an interpretation of punctuated equilibrium during this period" (p.186). To be more explicit, the authors' remarkable understatement makes clear their view that the gradual increase in encephalisation was required for the gradual increase in the abilities, including language, that the modern humans exhibit. The graph, Figure 8 below, suggests a gradual increase over a period of nearly 2 million years, implying gradual increases in abilities in ancestor species.

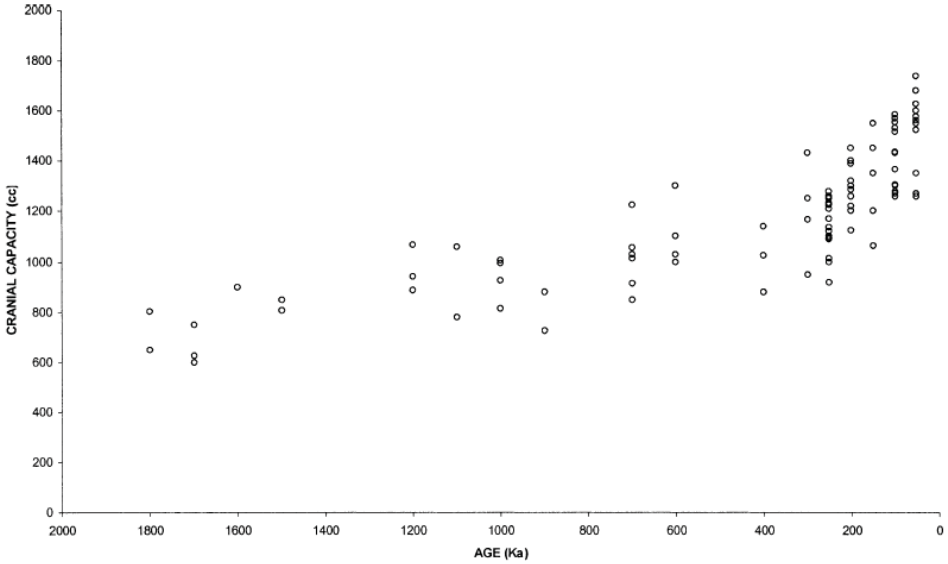


FIGURE 2. Cranial capacity (in cc) as a function of time for 94 specimens of Pleistocene *Homo*. The dates of the specimens between 1800 Ka and 300 Ka are rounded to the nearest 100 Ka, and to the nearest 50 Ka for specimens between 300 Ka and 50 Ka. The distribution has the visual appearance of curvilinearity.

Figure 8. Cranial Capacities of specimens of the Genus Homo. From Lee and Wolpoff (2003)

Early *Homo. erectus* in Africa (from about 1.7 to 1 million BP) averaged 900 cc in brain size, but later specimens from .5 My BP average 1100-1200 cc, which falls within the range of the brain size of modern humans. The earliest or archaic forms of *H. Sapiens* date to 300,000-400,000 BP and average over 1200 cc. The Neanderthal skull had a brain size of 1500 cc, which is larger than the brains of most modern humans. The average *H. sapiens sapiens* is around 1400 cc. (details from Washington State University, 2008).

In each of the articles summarised in the paragraphs above, the research compared anatomical and physiological facets of early Homo species with that of modern humans. None of these articles provides definitive proof that language was in use in the periods mentioned, hundreds of thousands and even millions of years in duration, but the comparisons drawn between earlier species and *H. sapiens* provide support for the view that language was present during these periods, although probably in primitive forms.

Genetic Evidence

Lai, Fisher, Hurst, Vargha-Khadem, and Monaco (2001) studied the 'KE' family in the United Kingdom in which several members across three generations have

verbal dyspraxia characterized by deficits in several facets of language processing, such as the ability to break up words into their constituent phonemes, for example, breaking down the word 'dogs' into its four phonemes /d/, /o/, /g/, /z/ (Snowball & Bolton, 1999, p. 67), and grammatical skills including production and comprehension of word inflections and syntactical structure. They identified in the affected members of the family differences in the FOXP2 gene which were not present in the unaffected members. They also examined an unrelated person who had similar deficiencies and found similar damage to the same gene. They concluded that the FOXP2 gene is involved in the development process that culminates in speech and language. Enard, Przeworski, et al. (2002) sequenced the FOXP2 related protein of humans, the chimpanzee, gorilla, rhesus monkey, and mouse, finding that the human protein contains two amino acid differences from the chimpanzee, gorilla and rhesus monkey version and that these three species' version differs by only one further amino acid from that of the mouse. They note the stability of the FOXP2 gene and associated protein across evolutionary time and suggest that a selective sweep, a process that eliminates less advantageous alleles, has occurred in the human version and dated it to between 120,000 and 200,000 BP. Examining the FOXP2 related protein sequence, Krause, Lalueza-Fox, et al. (2007) found that the two *H. Sapiens* differences from the non-human apes were shared with Neanderthals. These changes are present in the human haplotype (a particular length of a chromosome) that had been subject to the selective sweep. The authors push back the date of the selective sweep to earlier than 400,000BP. Coop, Bullaughy, et al. (2008) note the possibility of contamination with modern human DNA undermining the conclusions of Krause et al. (2007) but Enard (2011) notes the presence of the *H. sapiens* FoxP2 form in Denisovans which "make it also very unlikely that contamination with modern human DNA or gene flow from or to Neanderthals happened" (p. 417). This puts the date of the mutation at the upper end or beyond the range of 270,000 - 440,000 by which time the Neanderthals are believed to have split from *H. sapiens'* ancestors, according to nuclear DNA, or even beyond 500,000 years ago when the split is dated according to mtDNA (Reich, Green, et al. 2010). Gomez-Robies, De Castro, Arsuaga, Cabonelli, and Polly (2013) suggest a date of separation as long ago as 1 mya. Verne, Spiteri, et al. (2007) have demonstrated that the human FOXP2 gene is a transcription factor which influences the expression of other genes modulating synaptic plasticity, neurotransmission, neurodevelopment and axon guidance.

The discovery by Lai, Fisher, et al. (2001) that the modern human form of FOXP2 appears necessary for language, the discovery that the same form was present in Neanderthals (Krause et al., 2007) and Denisovans (Enard, 2011) and Verne et al.'s (2007) discovery that it controls neuronal developmental processes provide suggestive evidence that the selective pressure forcing the sweep may have been the early development of language.

Neurological Evidence

Aboitiz and V (1997) discuss a wide range of cortico-cortical connections linking various divisions of Wernicke and Broca's areas to each other and to some parts of the prefrontal and some infero-parietal regions that are active during speech. They suggest that such a large set of interconnections would evolve only gradually and propose therefore a Darwinian gradual evolution of speech abilities as opposed to "the Chomskian concept of a discrete language organ that has arisen by a single macro-mutation." (p. 394). Unfortunately, the authors do not date the time interval of this gradual evolution but the complexities of the cerebral connections they discuss suggests a period of several hundred thousand years, culminating in a language ability essentially identical to the modern one achieved at least one and probably several hundred thousand years ago. Schoenemann (2012) supports this view in a review of the increases in volume of those parts of the human brain, relative to other primates, active during speech. In particular, he draws attention to the increases in the prefrontal cortex relevant to semantics and syntax, temporal lobes relevant to connecting words to meanings and concepts, and Broca's and Wernicke's areas, which, although not unique to humans, appear to have been elaborated for language. He also draws attention to pathways allowing "for enhanced conscious control of speech [that] have been either added (to the larynx) or strengthened." (p. 455). Although Schoenemann (2012) does not date these changes, the increased volumes of the parts of the brain that have been enlarged to enable speech appear to rule out any saltation or indeed any period of time less than many hundreds of thousands of years. The details of the areas of expansion, and the implied time scale, appear consistent with the general cerebral expansion illustrated in figure 5 above.

Rilling, Glasser, et al. (2008) used diffusion-tensor imaging to map white matter in mammalian brains. They found that the cortical terminations of the arcuate fasciculus (AF), which connects the temporal lobes to areas of the frontal lobes, increase in complexity progressively from macaque monkeys to chimpanzees to humans. Damage to it causes speech aphasias in humans. Macaques have a voice area connected to the AF that is especially sensitive to conspecific vocalisation, although located at a different position on this fibre complex, relative to human voice processing. They suggest that the prominent temporal lobe projection on the human AF has aided the development of language in humans. Petkov, Kayser, et al. (2008), using fMRI, located a voice processing region in the anterior temporal plane for macaque monkeys with a signal sensitivity that enables the identification of conspecific individuals, a function performed by the same regions in humans. Ghazanfar (2008) concludes that this region has been strongly but gradually modified during human evolution and rejects the notion of any macro mutation for speech or the proposal that it is spandrel arising from increased brain size. He concludes that "the neural circuitry in humans evolved gradually from primitive precursors, which parallels findings from ethology which indicate a gradual emergence of vocal sophistication in the primate lineage" (p.484). The neurological evidence supports the view that the human language ability evolved over a lengthy period of many hundreds of thousands of years and was therefore likely to be doing so as far in the past as Erectine times and possibly further.

Paleo-archaeological Evidence

Karkanas, et al. (2007) report a consensus among paleoanthropologists supporting repeated use of fire in the Middle Palaeolithic period in the Qesem Cave in Israel. Their own explorations of strata dating to the end of the Lower Palaeolithic (400kya – 500kya) push back the use of fire to no later than 400 kya. "Hearths formed hubs around which other activities were carried out in the cave" (p. 208) such activities including the use of blades manifesting use-wear damage in conjunction with numerous cut marks on bones, the latter evidencing prey butchering, defleshing and marrow extraction. The authors see no distinction between behavioural patterns observed in Middle Palaeolithic populations elsewhere in the Levant and those reported in their paper. The particular activities discovered strongly suggest that language had developed

sufficiently by 400 kya for complex and dangerous activities to have been successfully coordinated within social groups (see discussion of Wadley, 2013, below).

Alperson-Afil (2008) proposes that frequent and controlled use of fire by hominins can be firmly dated to approximately 780 kya. She reports on the locations of burned and unburned flint microartifacts at the Gesher Benot Ya'aqov site in Israel. A statistical analysis of their spatial distribution from eight Acheulian archaeological horizons indicates that they occur in dense concentrations rather than being evenly distributed. She concludes that the fire making skills of the Acheulian hominins (*H. erectus*) enabled them to create fire at will in diverse environmental settings. Making the same interpretive assumptions as above, namely that fire hearths evidence complex social activities, Alperson-Afil's discoveries indicate a likelihood that language was developing in Erectine times.

Wilkins, Schoville, et al. (2012) report "multiple lines of evidence that ≈500,000 year old stone points…. functioned as spear tips" (p. 942) indicating that hafted weaponry was being manufactured at this time in Kathu Pan, South Africa. Multiple techniques were used to date the horizon in which the stone tips and a tooth of *Equus capensis* were found. Flakes of stone appear to have been removed from the bases of many of the stones. "This working could reflect intentional removals to shape the point to allow hafting." (p. 943). Wilkins et al. (2012) conducted experiments with replications of the hafted weapons, using a calibrated crossbow to simulate a thrusting spear and to ensure a consistent force. They found that most of their weaponry "had to be shot multiple times before exhibiting any visible damage". (p. 943).

Wadley (2013) discusses the discoveries of Wilkins et al. (2012) and many others with a view to defining some criteria by which to judge when complex cognition, implying language, could have been in existence. She draws on technological innovation as a better indicator than decorative skills, although Henshilwood, D'Erroco and Watts (2009) report evidence of the existence of engraved abstract depictions and associated skills "to at least 100 ka" (p. 45). Using the term "transformative technologies" to mean combining components to create an irreversible end product, Wadley (2013) highlights hafting tools, as opposed to wooden spear making, as one example. Together with planning over long distances or time and response inhibition,

she argues that technological transformations imply a capacity for analogical reasoning, cognitive fluidity, complex goal-directed action, flexibility in problem solving and task switching. She suggests heat treatments of ochre as a further exemplar of technological transformations, giving details of such treatment at the Becov 1 site in Bohemia dated to the Middle Palaeolithic, where "there is a possibility that heat treatment began long before 100kya." (p. 172). Such a date long predates the arrival of *H. sapiens* in Europe, implying advanced cognitive abilities of *H. neanderthalensis*. She expresses agreement with Ambrose (2010) that the existence of the cognitive abilities listed above strongly imply some degree of linguistic sophistication.

Experimental archaeology is making an increasing contribution to answering evolutionary questions. In a recent experiment into the instructional requirements of Acheulian technology, which dates from about 1.7 mya, Putt, Woods, and Franciscus (2014) divided participants into two groups. One group was instructed in knapping techniques non-verbally, by only observing an instructor, while the second was instructed verbally. An independent assessment of the two sets of "Acheulian" biface hand axes showed that the bifaces of the non-verbally instructed participants were of the same quality as those of the verbally instructed, suggesting that language was not necessary for Acheulian technology. However, a more recent experiment by Morgan, Uomini, et al. (2015) tested the transmission of the skills needed for Acheulian technology down a chain of practitioners, rather than the use of these skills. They employed six chains of participants, each chain headed by a trained expert and each successive member passing on knapping skills acquired from the previous one. There were five levels of instruction, ranging from nil (with each participant reverse engineering a biface) via different levels of imitation/emulation to verbal instruction and only a 5 minute teaching session. They found all levels of instruction led to decreased performance, relative to the trained experimenter, along chains of either five or ten participants long as information was lost. However, the drop in skills was much faster in the chains in which verbal instruction did not take place, relative to the one in which it did. Given the very short instruction session times, they conclude that their experiment demonstrated a strong selection pressure for the evolution of language during the Acheulian era and imply that hominins "possessed a capacity for teaching – and potentially simple proto-language – as early as 1.7 mya" (p. 6).

Lycett, von Cramon-Taubadel, and Eren (2015) go further with their exploration of the Levallois technology, present in Europe, Africa, and parts of Asia from 300 kya. They note the high levels of skills required for its manufacture and emphasise the very high level of morphological consistency across the large geographical range. They suggest that Levallois products stand out as the most plausible candidate of lithic products requiring "active instruction during the Lower-Middle Palaeolithic" (p. 14). They do not insist that this active instruction was delivered via language but they are insistent that the technology required specific mechanism of social learning, minimally both imitative and emulative copying modes of learning but likely including teaching. Combined with Morgan, Uomini, et al.'s (2015) experiment, the case for the social learning mechanism being language-based becomes at least plausible.

Homo. *neanderthalensis* and associated technologies

In a review of evidence proposed by numerous researchers to explain the demise of the Neanderthals, Villa and Roebebroek (2014) note that Neanderthals as long ago as 200 kya were using fire to synthesize pitch from bark as a form of glue for hafting tools. This required the use of fire "that involved distillation in the absence of oxygen and within a temperature interval of 340°C–400°C". The authors write that "The evidence of pitch production as early as 200 kya by European Neanderthals shows that those "elevated cognitive abilities" were not the exclusive domain of modern humans." ("Hafting Procedures, Heat Treatment and Cognition", para. 2). The recent discovery (Soressi, McPherson, et al., 2013) of lissoir (smooth) bone tool fragments from Neanderthal sites in the Dordogne, France, dated to 47,710 – 41, 130 BP and bladelets from Mousterian sites (Faivre, 2012; Faivre, Discamps, et al., 2014) prior to the final Middle Palaeolithic, demonstrate that such specialisms were present in Neanderthal Europe. The lissoir tools "predate the oldest known age for the use of similar objects in Europe by anatomically modern humans" (Soressi et al., 2013 p. 14186) suggesting that the cognitive abilities of the Neanderthals by this time were comparable with those of *H. sapiens*. The separation of the two species is dated to in excess of 500 kya by Gómez-Robles, de Castro, et al., (2013) and at between 555 kya and 765 kya by Prüfer, Racimo, et al. (2014). Assuming only that the cerebral complexities required for this tool making are the consequence of homologous descent from a common ancestor, emergence of these elevated cognitive abilities can be dated to this era.

Conclusion

Several lines of argument combine to support the view that early or proto-language was developing in Erectine times. A gradual process of cognitive development during the next several hundred thousand years via archaic *H. sapiens* 300 kya (Aitchison, 2000), evidenced by progressively more sophisticated tool making and decorative abilities, culminated in language abilities comparable to those of modern *H. sapiens* by approximately 150,000 years BP, the time when *H. Sapiens* was reaching its current state of development (Klein, 2002). Prior to the exodus of *H. sapiens* from Africa or west Asia, approximately 60,000 years ago, there was a period of around a hundred thousand years during which *H. sapiens* was resident only in these locations. Modern genetic studies suggest that during this period the gene exchange taking place was sufficient to produce the fairly uniform human genome and range of psychological attributes that exist today, including of course, the capacities for language and cognitive dissonance and any detectable bias in its expression. The mutual reinforcement of a general language ability and cognitive dissonance, described in the Theoretical background, implies that the evolution of cognitive dissonance took place over several hundred thousand years.

Appendix C - Contradictory Pairs for Assessment

Finding warm shelter on cold nights will be difficult.	1=No contradiction		3 or 4 =slightly contradictory				7 = Very contradictory
Finding warm shelter on cold nights will be easy.	1	2	3	4	5	6	7
I need to wear glasses when I read a book.	No contradiction		3 or 4 =slightly contradictory				7 = Very contradictory
I don't need glasses when I read newspapers.	1	2	3	4	5	6	7
We are confident that our food will never run out.	No contradiction		3 or 4 =slightly contradictory				7 = Very contradictory
We are sure our food will run out very quickly.	1	2	3	4	5	6	7
Yesterday morning, I went running.	1=No contradiction		3 or 4 =slightly contradictory				7 = Very contradictory
Yesterday morning, I went shopping.	1	2	3	4	5	6	7
Several times I thought I was about to die.	1=No contradiction		3 or 4 =slightly contradictory				7 = Very contradictory
I never believed my life was in danger.	1	2	3	4	5	6	7
I believe newspapers slant the news as their owners dictate.	1=No contradiction		3 or 4 =slightly contradictory				7 = Very contradictory
I'm annoyed by newspaper articles that ignore important facts.	1	2	3	4	5	6	7

Appendix C – Contradictory Pairs

That statement of John's is very meaningful	1=No contradiction		3 or 4 =slightly contradictory			7 = Very contradictory	
That statement of John's is truly meaningless.	1	2	3	4	5	6	7
I often find myself lusting after my partner.	1=No contradiction		3 or 4 =slightly contradictory			7 = Very contradictory	
I never find myself lusting after my partner.	1	2	3	4	5	6	7
We expect our new fridge to last a long time.	1=No contradiction		3 or 4 =slightly contradictory			7 = Very contradictory	
We expect our new fridge to last for a while.	1	2	3	4	5	6	7
He has the right to voice his opinion.	1=No contradiction		3 or 4 =slightly contradictory			7 = Very contradictory	
He does not have the right to voice his opinion	1	2	3	4	5	6	7
I have read the difficult book James gave me.	1=No contradiction		3 or 4 =slightly contradictory			7 = Very contradictory	
I have not read the difficult book Bill gave me.	1	2	3	4	5	6	7

Appendix C – Contradictory Pairs

I disagree with all his opinions.	1=No contradiction 3 or 4 =slightly contradictory 7 = Very contradictory
I always agree with his opinions.	1 2 3 4 5 6 7
I arrived at exactly the right time yesterday.	1=No contradiction 3 or 4 =slightly contradictory 7 = Very contradictory
I arrived at completely the wrong time yesterday.	1 2 3 4 5 6 7
I felt wretched when my partner had affairs.	1=No contradiction 3 or 4 =slightly contradictory 7 = Very contradictory
My partner's many affairs did not upset me.	1 2 3 4 5 6 7
I find word definitions wholly fascinating.	1=No contradiction 3 or 4 =slightly contradictory 7 = Very contradictory
I find word origins somewhat interesting.	1 2 3 4 5 6 7
We look forward to having several children.	1=No contradiction 3 or 4 =slightly contradictory 7 = Very contradictory
We have decided not to have any children.	1 2 3 4 5 6 7
It's never necessary to repay favours.	1=No contradiction 3 or 4 =slightly contradictory 7 = Very contradictory
To me it's essential to return favours.	1 2 3 4 5 6 7

Appendix C – Contradictory Pairs

I think housework is tedious and uninteresting. 1=No contradiction 3 or 4 =slightly contradictory 7 = Very contradictory
I enjoy living in a house kept clean and tidy. 1 2 3 4 5 6 7

I'm well known to be scrupulously honest. 1=No contradiction 3 or 4 =slightly contradictory 7 = Very contradictory
I'm generally known to be dishonest. 1 2 3 4 5 6 7

Her telephone has a pleasing pink tone. 1=No contradiction 3 or 4 =slightly contradictory 7 = Very contradictory
Her telephone has a pleasant ring tone. 1 2 3 4 5 6 7

I've had no sexual affairs since I got married. 1=No contradiction 3 or 4 =slightly contradictory 7 = Very contradictory
I've had eight sexual affairs since I got married. 1 2 3 4 5 6 7

My new letter paper is coloured a light purple. 1=No contradiction 3 or 4 =slightly contradictory 7 = Very contradictory My new
letter paper is coloured a light green. 1 2 3 4 5 6 7

My favourite cup has a crack in the side. 1=No contradiction 3 or 4 =slightly contradictory 7 = Very contradictory My
favourite cup has no crack in the side. 1 2 3 4 5 6 7

Appendix C – Contradictory Pairs

I intensely dislike driving all types of cars.	1=No contradiction	3 or 4 =slightly contradictory 7 = Very contradictory
I prefer driving automatics to manual shift cars.	1 2 3 4 5 6 7	
Happily, the opposite sex find me attractive	1=No contradiction	3 or 4 =slightly contradictory 7 = Very contradictory
Sadly, I'm unattractive to the opposite sex.	1 2 3 4 5 6 7	
My favourite food is fish and chips.	1=No contradiction	3 or 4 =slightly contradictory 7 = Very contradictor
The food I eat most often is spaghetti.	1 2 3 4 5 6 7	
The requests always arrive on Fridays only.	1=No contradiction	3 or 4 =slightly contradictory 7 = Very contradictory
The requests arrive evenly throughout the week	1 2 3 4 5 6 7	
I'm widely known to be unfaithful to my spouse.	1=No contradiction	3 or 4 =slightly contradictory 7 = Very contradictory
Everyone knows I'm both married and faithful.	1 2 3 4 5 6 7	

Appendix D – Non-Contradictory Stories

1. I think we should have a family holiday.
2. I have about three weeks holiday leave owing.
3. On the way home I will get some travel brochures.
4. We will have to choose between skiing and the sun.
5. I am sure that the children would prefer skiing.
6. We parents would prefer the sun, sand and surf.
7. Family discussions about locations are best.
8. Children always turn discussions into quarrels.
9. I should probably just announce the location.
10. I will discuss this idea with my partner first.
11. I am glad to have decided on this action.

1. How many weeks leave to you have owing?
2. What will you take home?
3. Between skiing and sun, which do you prefer?
4. Do your children quarrel?

1. I enjoy the many kinds of trees in the forest.
2. I wonder at the differences between them.
3. I did not notice their variation at first.
4. I have been learning their names from library books.
5. I try to name them as I pass during my walks.
6. I am proud to acquire my new knowledge.
7. I am also learning about their biology.
8. Sometimes I explain their growth patterns to friends.
9. They usually find it interesting too.
10. They do not like the detail as much as I do.
11. I have to remember that when I talk to them.

1. What have you been learning about from library books?
2. Is the word biology mentioned?
3. Do your friends find growth patterns interesting?
4. Do they like detail as much as you do?

1. Some years ago my bank opened a branch near home."
2. I noticed another bank did the same soon after.
3. Both of them were quite busy for a while afterwards.
4. Then all the other banks did something similar.
5. The best shopping streets were soon awash with banks.
6. I began to read there were too many bank branches.
7. Some of them then put ATMs facing the pavements.
8. I thought that this would cost them a lot of money.
9. The banks always want to increase their profits.
10. Then I noticed they started closing retail branches.
11. Now I sometimes can't find a bank when I want one.

1. When did the banks start opening branches locally?
2. Did they fill the shopping streets?
3. Did you think the cost of the ATMs was a lot or a little?
4. Can you always find a bank today when you want to?

1. I recently received a book about architecture.
2. It had designs and pictures of palaces in it.
3. The dimensions of rooms were carefully calculated.
4. The ratios were determined by the ancient Greeks.
5. There were no corridors by the sides of bedrooms.
6. The bedrooms were all interconnected by doors.
7. You had to walk through other people\'s to your own.
8. They all had four poster beds with curtains.
9. With no corridors, this was how privacy was provided.
10. The word 'corridor' appeared in the sixteenth century.
11. It's from Latin, meaning a place you can run in.

1. Who determined the desirable ratios of room dimensions?
2. Were there corridors in the palaces?
3. How was privacy provided?
4. What language does the word 'corridor' come from?

1. People of all types like some kind of music.
2. Much early music came from the human voice.
3. Musical instruments are thousands of years old.
4. Their makers had no sharp metal tools to use.
5. Even holes in bones appear randomly placed.
6. Today\’s instruments are quite sophisticated.
7. Learning to play an instrument takes patience.
8. A player must devote many hours to learning.
9. Some people practise over an hour each day.
10. That is quite difficult for a young child.
11. It is sufficient for quite rapid progress though.

1. Is music universal?
2. What were early instruments made of?
3. Is the word trombone mentioned?
4. Is playing one hour per day enough to make progress?

1. The game is simple enough for all to play.
2. You have to throw a ball at a big target.
3. The target is only four metres away.
4. There are a couple of problems though.
5. You have to stand with your back to the target.
6. You bend forward as you throw the ball.
7. Forward enough to see the target between your legs.
8. The ball must also pass between your legs.
9. It’s very difficult to adjust your arm movement.
10. Forget your dignity as you throw the ball.
11. Some people are just too inhibited to play.

1. How many metres away is the target?
2. Which way do you face to throw the ball?
3. Do you think this game is undignified?
4. Is the word inhibited mentioned?

Appendix E – Contradictory Pair Analysis

In the examples below, scores are accompanied by standard deviations in brackets ().
As an example of a small difference in wording,
“Finding warm shelter on cold nights will be difficult”
“Finding warm shelter on cold nights will be easy.”
was awarded the high score of 6.7 (.78).

Similarly, the pair
“It matters a lot if people think I’m a crook” and
“It doesn’t matter if people believe I’m a crook”
was awarded a score of 6.092 (1.43).

However,
“People can rely upon me to return favours” and
“Folk can’t rely upon me to return their favours”
was awarded a score of 5.96 (1.53).
Both the latter two pairs had directly contradicting verbs. The inclusion of the word “their” in the second sentence of the pair immediately above may have slightly reduced its score, despite the obvious opposition of “can” and “can’t”. The use of ‘People’ in the first sentence and ‘Folk’ in the second may also have detracted from the contradictoriness assessment.

The appearance similarity also applies in the comparison below.
“Swindling greatly affects our social standing”
“Swindling has no effect on our social standing”
was awarded 6.3 (1.23).

In the examples above, there appears to be a visual ease of processing effect. If the phrasing of the pair enables participants to concentrate their attention on only a few words, so as to assess the conflict between only these, the scores appear to be higher. This could be due to a perceptual fluency effect (Oppenheimer, 2008) as participants had been requested to assess the pair only for contradictoriness. Contradictions may

have been easier to identify where there was no disfluency caused by differences of word usage and sentence format.

However, this visual effect may be moderated if the similarities of appearance are in the first part of the sentences only.
"Gossip that I don't exchange fairly makes me angry"
"Gossip that I don't exchange fairly won't affect me"
was awarded 4.9 (1.96) despite the identical appearance of the first half of each sentence. The contradictory second halves of each sentence present visual differences.

There appear to be two factors which reduce scores:
the similarity of appearance occurring only in the first half of the sentences and

the contradiction appearing in words which are not similar in appearance namely "makes me angry" and "won't affect me".

The effect of dissimilarity of appearance is reinforced by the pair
"I never know when the next request will come"
"The requests always come at the end of the week"
which was awarded 5.56 (1.85). It is possible that the abstract nature of these sentences also contributed to the low score, it being well established that processing abstract sentences requires greater cognitive effort than concrete ones (Strain, Patterson, & Seidenberg, 1995). This extra effort causes disfluency (Oppenheimer, 2008).

The same dissimilarity of appearance explanation might apply to
"Several times I thought I was about to die"
"I never believed my life was in danger"
which was awarded only 5.76 (1.67) despite the unambiguous contradiction of the sense. In this instance, some effect of Terror Management Theory (see below) might also apply, diminishing the degree of attention paid to the sentences. Another factor may be that the contradiction is implicit to some degree rather than explicit, a mistake made at the time of the composition of the pair. The low score disqualified the pair from further use.

Visual similarity seems to have been less influential if the pair dealt with highly emotional and adaptive issues. For example,
“I’m widely known to be unfaithful to my spouse.”
“Everyone knows I’m both married and faithful.”
was awarded 6.18 (1.36) despite the relatively substantial visual differences. This point is supported by the following pair.
“Happily, the opposite sex finds me attractive”
“Sadly I’m unattractive to the opposite sex “
which was awarded 6.127 (1.5). The latter pair was not used in the later stages of the research due to a decision not to include sexual attractiveness as a topic.

One participant pointed out that a small number of pairs had two verbs in them, one contradictory and one not, the latter detracting from the contradictory nature of the pair. For example, the pair
“People knew that I had had several affairs”
“People didn’t know that I had had several affairs”
was awarded a score of only 5.567 (1.9).
At the time of composition, it was assumed that participants would concentrate upon the conflict between “knew” and “didn’t know”. However, some participants appear to have placed some emphasis on the phrase “I had had” in both sentences, which detracted from their assessment of the contradictoriness of the pair as a whole. This example also supports the hypothesis that the contradictoriness is more readily noticed if the discrepant wording is in the latter half of the sentence.

The following guidelines for composing pairs of contradictory sentences can be drawn from the Item Construction part of this research:

1. Ensure that the two contradictory sentences are very similar in syntax and appearance or form, particularly the second half. One method of achieving this is by employing the ‘easy’ constructions listed by De Marneffe et al. (2008).
2. If the contradictoriness is conveyed via adjectives, use only well recognised antonyms.
3. Despite the above, some dissimilarity of appearance can be tolerated if the sentences have high emotional valence.
4. Avoid the abstract.

5. Avoid placing two verbs in the sentences.
6. Avoid opposing adjectival or adverbial phrases which imply contradictoriness rather than being explicitly contradictory.

Appendix F – Tray Sort 90% Correct Placement

Kappa-type statistics

Cohen (1960) devised the Kappa Coefficient of Agreement, κ, as an assessment of the degree to which two judges, acting independently, agree when they place a set of items in mutually exclusive and exhaustive independent categories of nominal scale, according to their views of the characteristics of these items.

$$\kappa = (p_o - p_c)/ (1- p_c)$$

where p_o is the proportion of the judgements or categorisations on which they agree and p_c is the proportion of agreements that could be expected by chance alone.

A formula for an approximation to the standard error of κ is given as

$$\sigma_\kappa = \sqrt{((p_o(1-p_o)/N(1-p_c)^2)}$$

provided N is large (> 100) with

$$95\% \text{ confidence limits} = \kappa \pm 1.96\ \sigma_\kappa \text{ and}$$
$$99\% \text{ confidence limits} = \kappa \pm 2.58\ \sigma_\kappa$$

Fleiss (1971) generalised Cohen's version of Kappa for more than two judges by substituting p_0' for p_0 where p_0' is the average of agreements for each item being categorised across all judges.

Landis and Koch (1977) developed a family of Kappa-type statistics to cater for inter-observer bias and differences between sub-populations of the units being categorised. In order to allow comparisons to be made between various Kappa-type computations, they suggest that the resulting values can be assessed as illustrated in Table 36 below.

Table 36

Landis and Koch (1971) Interpretations of Kappa Statistics

Kappa-Statistic	Strength of Agreement
< 0.00	Poor
0.00 -0.20	slight
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Substantial
0.81-1.00	Almost Perfect

Application of these formulae to the tray sorting test

In order to apply the Fleiss's (1971) adaptation of Cohen's (1960) Kappa, a values for p_c must first be determined, that is, what is the probability of a random distribution of stories correctly placing all twenty-seven in the trays pertaining to their topic. This number can be calculated in a three-part computation as illustrated below

1. First computation: How many different sequences of story placements would result in the twenty-seven stories all being correctly placed in the trays pertaining to their topic theme. Let this number be X.

2. Second computation: What proportion of the total number of different distributions of stories in the trays would X represent. Let the total number of different distributions or permutations be Y.

3. Third computation: The probability of a correct distribution of the stories happening by chance, p_c, is X/Y.

Derivation of a formula for X

Let there be 4 trays, A, B, C, and D, with 3 stories each and the stories labeled as 1, 2, 3, 4 etc. to 12.

Let 1 correct distribution be

A	B	C	D
1 2 3	4 5 6	7 8 9	10 11 12

Holding the distributions in A, B and C constant, there are 6 distributions of D's stories. These combinations of stories are illustrated below

1	2	3		4	5	6		7	8	9		10	11	12
		<-					Ditto				->	10	12	11
							Ditto					11	10	12
							Ditto					11	12	10
							Ditto					12	10	11
							Ditto					12	11	10

The number of variations of D is 3! i.e., factorial the number of stories in one tray

There are the same number of variations of the stories in C, i.e. 6, and for each one (only the sequence 7 8 and 9 is shown above) there are the 6 variations in D as also shown. Keeping the sequences of trays A and B the same as shown above, there are therefore 6 * 6 different permutations which all have the correct stories allocated to trays C and D, that is, 6^2. This pattern of variation can be extended back to tray B in which case there are 6 * 6 * 6 variations of the stories in boxes B, C and D for the one permutation of A's stories as shown above, that is, 6^3. As there are also six correct sequences for tray A, the total number of correct permutations which meet the requirement that tray A contains only stories 1, 2 and 3, B contains only 4, 5, and 6, tray C contains only 7, 8 and 9, and D contains only 10, 11, and 12, is 6^4

Generalising this derivation, the formula for the number of correct possible permutations in all trays is

(factorial [the number of stories in each tray]) $^{(\text{the number of trays})}$.

Applying this formula to three stories in 9 trays, the number of correct permutations is

$$3!^{9} = 6^{9} \approx 1 * 10^{7}$$

X is approximately 10,000,000

Y, the total number of permutations is $27! \approx 1 * 10^{28}$

As $p_c = X / Y$, chance allocation of the stories to the correct trays is

$$1^{9} / 1^{28}$$

$$p_c \approx 1 * 10^{-19}$$

The likelihood of a correct distribution by chance is negligibly small.

From the observation of the actual results of the tray sorting, $p_0' = .96$. From this and from the value of p_c calculated above, the κ statistic for the tray sorting is

$$(.96 – 1* 10^{-19})/(1.0 - 1* 10^{-19}) \approx .96.$$

Applying Fleiss' formulas for standard error and confidence limits (which are very conservative when applied to this experiment as N = 27 rather than 100) the Standard error = .037, the 95% Confidence limits are .886 to 1 and the 99% confidence limits are .863 to 1. It is against the values of $\kappa \approx .96$, the confidence limits calculated above and the assessments suggested by Landis and Koch (1977) that the decision to adopt a threshold of 90% placement correctness must be judged. Even at the lower 99% confidence limit, at which $\kappa \approx .863$, the strength of agreement is still "almost perfect". The 90% threshold was both demanding and exceeded.

Appendix G – Tray Sort

Appendix H - Contradictory Paradigm Assumption Testing

The fundamental assumptions of the contradiction experimental paradigm are:
(1) the contradictory lines will have longer Response Times (RTs) than the non-contradictory ones,
(2) the longer RTs can be interpreted as the participants having detected the contradiction and
(3) the participants are feeling cognitive dissonance during the extra response time.
The opportunity has been taken to ascertain whether these assumptions can be verified using the data collected from both Studies.

Comparison of Contradictory and Non-Contradictory Line RTs

The contradictory experimental paradigm literature demonstrates that this approach has been used in several ways. In all cases, participants read a multi-line text containing anomalies, that is, one line is inconsistent with a previous one. In early experiments, once the participants had read a passage, the experimenter posed questions to ascertain whether the anomaly has been noticed (Otero & Kintsch, 1992). In later experiments, the times participants took to read each line of text were recorded (RTs). Comparisons were subsequently made between the RTs for a line which is inconsistent with an earlier one and those that are not. Statistically significantly different times between these two sets of RTs are interpreted as evidence in support of the hypothesis being tested by the experiment (Huitema et al. 1993). The first assumption above is taken for granted without independent tests having been used to demonstrate its validity. In other cases, experimenters use the contrasts between RTs of two target lines in separate texts, both of which contradict earlier lines but do so to varying and different degrees. Egidi and Gerrig (2006) used separate texts containing different degrees of consistency between goals stated in an early line and actions described in later ones. Measured RTs of target lines are compared to ascertain participants' reactions to mismatches. In the two types of experiment which employ RT measurement, it is assumed that RTs for non-anomalous lines are shorter than those containing any form of anomaly and provide a baseline for comparison. The current research provided an opportunity to test this assumption.

Appendix H - Contradictory Paradigm Assumption Testing

In a manner analogous to Egidi and Gerrig (2006), the current research employed target Line 10 RTs in separate stories to compare adaptive contradictions with Non-Adaptive ones. However, for all the contradictory stories, lines 1, 2, 3 – 9 and 11 are non-contradictory. Measuring the difference between the RTs of these lines and those of the line 10s of these stories provides the opportunity to test the first assumption. There were problems associated with doing this for unanticipated reasons. Participants were quite frequently observed to commence reading a new story by pressing the space bar to display Line 1 and then spending some time readjusting their body positions before reading it. Occasionally this readjustment would carry over to Line 2. In this way, participants introduced extended RTs with no relevance to the experiment. On other occasions, participants would sometimes commence reading a story by accidently pressing the space bar twice very quickly in succession giving themselves insufficient time to read line 1 and then spend a long time reading line 2 as if to compensate for their error. These considerations suggested that lines 1 and 2 RTs should not be included in any comparisons. In addition, in contradiction paradigm experiments, there can be a carry-over of the hesitation that follows the noticing of an anomaly to a subsequent line also (Huitema, 1993, p. 1056). This consideration suggests that line 11 RT should also be excluded. As all lines have been composed to have between 11 and 13 syllables and to subtend approximately the same angle to participants' eyes, no line is visibly different from any of the others. Every line in the 3 – 9 range is non-contradictory in the sense that none of these contradict a previous one and none is distinguishable from any other in any meaningful manner. Any of them could therefore be used in a comparison with line 10 RTs. However, there is some small merit in not using a single line. Inspection of the RTs reveals that all the data is noisy to some greater or lesser extent. A small improvement can be made over comparing the contradictory line RTs with just one of the non-contradictory ones by comparing it with the mean of the non-contradictory ones. The comparison is then more all encompassing and some of the noise in the non-contradictory line RTs is averaged out. The quality of the comparison is marginally improved. The first assumption has therefore been assessed by means of paired samples *t*-tests between the means of the Line 10 RTs and the means of the Lines 3-9 RTs for all topics. The results of the comparisons are shown in Table 37 below.

Table 37
Study 3 Paired Samples t -test Contradictory Line 10s vs Non-contradictory lines. All times in ms.

Lines	Mean	Diff.	SD	Std. Error Mean	*t*	*df*	Sig. (2-tailed)
SLT Line 10s	3096	574	1938	137	4.53	200	<.001
SLT Mean Lines 3-9	2522		938	66			
SNLT Line 10s	3269	621	2098	148	4.675	200	<.001
SNLT Mean Lines 3-9	2648		926	65			
CL Line 10s	3613	1199	2250	159	8.377	200	<.001
CL Means Lines 3-9	2414		841	59			
H Line 10s	3062	347	1961	138	2.842	200	.005
H Mean Lines 3-9	2715		935	66			
SI Line 10s	2821	229	1681	119	2.106	200	.004
SI Mean Lines 3-9	2591		936	66			
SU Line 10s	2893	265	1564	110	2.872	200	.005
SU Means Lines 3-9	2628		1013	71			

Note: For all t-tests, N= 201. SLT = Survival Life Threatening, SNLT = Survival Non-Life Threatening, CL = Child- Loss, H = Honesty, S = Socially Important, SU = Socially Unimportant.

The analysis demonstrated the validity of the assumption. Contradictory Lines Mean RTs were between 8.8% (Socially Important) and 49.7% (Child-Loss) longer than the Non-Contradictory Lines. The low difference for the Socially Important topic was a surprise. As a confirmation two further tests were made. The Means of Lines 3-11 for all six of the Non-Contradictory stories were compared with the Line 10 RT of the Contradictory stories with Outliers excluded and included. As neither Lines 10 nor 11 of the non-contradictory stories contradicted an earlier one, they were included in the calculation of the means. The results of the paired samples *t*-test comparisons appear below in Tables 38 and 39 below.

Table 38

Paired Samples t-test - RTs of Line 10 without Outliers with means of line 3-11 of Non-Contradictory Stories. All times in ms.

		Mean	N	Diff	SD	Std. Error Mean
Pair 1	SLT10_wo_OL	2810	192	234	1337	97
	Non_Contr_Story_Means	2575	192		736	53
Pair 2	SNLT10_wo_OL	2949	193	380	1335	96
	Non_Contr_Story_Means	2569	193		710	51
Pair 3	CL10_wo_OL	3295	193	732	1435	103
	Non_Contr_Story_Means	2563	193		716	52
Pair 4	H10_wo_OL	2782	193	211	1273	92
	Non_Contr_Story_Means	2571	193		726	52
Pair 5	SI10_wo_OL	2575	192	4	1157	84
	Non_Contr_Story_Means	2571	192		709	51
Pair 6	SU10_wo_OL	2759	196	168	1209	86
	Non_Contr_Story_Means	2591	196		733	52

Note: SNLT10_wo_OL = Survival Non-Life Threatening Line 10 RTs excluding Outliers. CL10_wo_OL= Child-Loss Line 10 RTs excluding Outliers, H10_wo_OL = Honesty Line 10 RTs excluding Outliers, SI10_wo_OL= Socially Important Line 10 RTs excluding Outliers, SU10_wo_OL= Socially Unimportant Line 10 RTs excluding Outliers, Non_Contr_Story_Means = Means of lines 3-11 for all six non-contradictory stories.

Although Outliers have been excluded from the comparisons, the Means of the RTs for Lines 3-11 of the non-contradictory stories are all less than the contradictory Line 10 RTs although only marginally so for the Socially Important topic. Conventional statistically significance is achieved for five of the topics, the Socially Important one being the exception.

Table 39

Paired Samples t-test - RTs of L 10 including Outliers with means of lines 3-11 of Non-contradictory Stories. All times in ms.

		Mean	N	Diff	SD	Std. Error Mean	*t*	*df*	Sig (2-tailed)
Pair 1	SLT_ALL	3096	201	492	1938	137	4.15	200	<.001
	Non_Contr_Story_Means	2604	201		742	52			
Pair 2	SNLT_ALL	3269	201	665	2098	148	5.03	200	<.001
	Non_Contr_Story_Means	2604	201		742	52			
Pair 3	CL_ALL	3613	201	1009	2250	159	6.98	200	<.001
	Non_Contr_Story_Means	2604	201		742	52			
Pair 4	H_ALL	3062	201	458	1961	138	3.8	200	<.001
	Non_Contr_Story_Means	2604	201		742	52			
Pair 5	SI_ALL	2821	201	217	1681	119	2.02	200	.044
	Non_Contr_Story_Means	2604	201		742	52			
Pair 6	SU_ALL	2893	201	289	1564	110	3.15	200	.002
	Non_Contr_Story_Means	2604	201		742	52			

Note: SLT10_ALL = Survival Life Threatening Line 10 RTs including Outliers.
SNLT10_ALL = Survival Non-Life Threatening Line 10 RTs including Outliers. CL10_ALL= Child-Loss Line 10 RTs including Outliers, H10_ALL = Honesty Line 10 RTs including Outliers, SI10_ALL= Socially Important Line 10 RTs including Outliers, SU10_ALL= Socially Unimportant Line 10 RTs including Outliers, Non_Contr_Story_Means = Means of lines 3-11 for all six non-contradictory stories.

With the Outliers included in the comparison, the means of the Non-contradictory story RTs are all well less than the Line 10 contradictory story RTs, reaching conventional statistical significance for all topics. All three tests have validated the first assumption of the contradiction paradigm. As Study 3 had shown consistent differences, the same set of differences for the six topics included in Study 4 are shown by listing them rather than conducting t-tests. The results are shown in Table 40 below.

Table 40

Comparison of Contradictory and Non-Contradictory RTs. All times in ms.

Topic	Contradictory Line 10 RTs	Mean of Lines 3-9	Difference
SU	2700	2673	27
REC	3619	2757	862
H	2824	2780	44
FAM	2735	2703	32
CL	3434	2512	922
FID	3157	2624	533

Note: Su = Socially Unimportant, REC = Reciprocity, H = Honesty, FAM = Family, CL = Child-Loss, FID = Fidelity.

The mean of the contradiction Line 10 RTs for all Study 4 (Line 10s) is 3,078 ms. The mean of all Lines 3 - 9 for all Study 4 topics is 2,675 ms. Lines 1 and 2 RTs were not included in the comparison for the same reason given for Study 3. As in Study 3, the contradictory lines on average had longer RTs than the non-contradictory lines, although for three topics, Socially Unimportant, Honesty and Family, the differences were very small. In general, the Study 4 findings support those of study 3.

Question 3 answer as an indication of having noticed the contradiction.

The second assumption upon which contradiction paradigm experiments are based is that lengthier RTs indicate that participants have noticed the anomaly or contradiction. Both Studies 3 and 4 were designed to generate some evidence in support of the validity of this assumption. As soon as participants finished reading each story, they were presented with a screen displaying 4 questions about the story. They were invited to provide short answers, many of the questions having been designed to be answered with a yes (Y) or no (N). Without leading the participant in any way, question 3 was phrased so as to provide an opportunity to state whether a contradiction had been noticed. All participants' answers to this question were subsequently coded to reflect

whether their answers indicated that the contradictions had been noticed. An independent samples *t*-test of the RTs coded in this way was conducted to ascertain if those participants who indicated they had noticed the contradiction had longer line 10 RTs than those who provided no indication. The results appear in Table 41 below.

Table 41

Study 3: Independent Samples t-test of RTs of participants who indicated in Q 3 that they noticed the contradiction compared to those who gave no indication. All times in ms.

Topic	Noticed		No indication		*t*	*df*	Sig (2-tailed)
	N	Mean	N	Mean			
Survival Life Threatening	19	3672	182	3036	-1.09	20.2	.288
Survival Non-Life Threatening	15	4082	186	3203	-0.96	14.7	.355
Child-Loss	29	4337	172	3491	-1.89	199	.061
Honesty	13	3333	188	3043	-0.52	199	.606
Socially Important	27	3402	174	2731	-1.58	30.69	.126
Socially Unimportant	19	2651	182	2918	0.71	199	.48

The number of participants who provided an indication in their answers was generally low reducing the likelihood of the analysis reaching statistical significance. For 5 out of the 6 topics, the means of the Line 10 RTs of participants who provided indications were longer than those of participants who did not, although none of them reached conventional statistical significance, the case of Child-Loss just falling short. In the case of the Socially Unimportant topic, those who did notice the contradiction had Line 10 RTs shorter than those who provided no indication, although the difference between the means was the lowest of all topics. In general, the results confirm the view that participants who stated that they noticed the contradictions had longer Line 10 RTs, interpreted as a manifestation of their CD, although weakly. It should be noted that the question seeking an indication of the contradiction having been noticed was positioned as the third one rather than, say, the first, in order to reduce the likelihood of participants noticing that there was a question aimed at exposing the existence of contradictions in the stories immediately after they had finished reading them. This precaution may also have had the effect of blunting some participants' inclination to note the contradictions due to the distraction of the intervening two questions.

Exactly the same coding and analysis procedure was conducted for Study 4 as was conducted for Study 3. The results appear in Table 42 below.

Table 42

Question 3 - Results of t–tests of Line 10 RTs of participants who noted a contradiction compared to those who did not.

Topic	Participants	Noted Contradiction?	N	Mean	SD	Diff	*t*	*df*	Sig (2-tailed
Honesty	All	No	186	2783	1522	559	-1.324	199	.187
		Yes	15	3342	2134				
	Without Outliers	No	177	2566	1182	264	-0.819	189	.414
		Yes	14	2830	817				
Reciprocity	All	No	176	3672	2126	-686	1.573	199	.119
		Yes	25	2986	1368				
	Without Outliers	No	168	3358	1501	-372	1.167	191	.245
		Yes	25	2986	1368				
Child-Loss	All	No	182	3392	1726	452	-1.074	199	.284
		Yes	19	3843	1919				
	Without Outliers	No	175	3191	1407	448	-1.256	191	.211
		Yes	18	3639	1748				

Table 42 (cont.)

Topic	Participants	Noted Contradiction?	N	Mean	SD	Diff	*t*	*df*	Sig (2-tailed
Fidelity	All	No	162	3065	1668	478	-1.573	199	.117
		Yes	39	3543	1849				
	Without Outliers	No	156	2850	1236	200	-0.876	189	.382
		Yes	35	3051	1164				
Family	All	No	179	2786	1373	-469	1.54	199	.124
		Yes	22	2318	1052				
	Without Outliers	No	171	2607	1110	-290	1.158	191	.248
		Yes	22	2318	1052				
Socially Unimportant	All	No	175	2630	1275	547	-1.921	199	.056
		Yes	26	3176	1812				
	Without Outliers	No	168	2485	1071	298	-1.267	190	.207
		Yes	24	2783	1119				

The *t*-tests provided equivocal results. Four topics, Honesty, Child-Loss, Fidelity, and Socially Unimportant indicated that those participants who noted the contradictions had longer Line 10 RTs than those who did not, while Reciprocity and Family topics indicated the opposite. This pattern was the same irrespective of whether the outlier Line 10 RTs were included in the tests or were omitted. When a *t* - test was conducted for all topics and all participants, (6 X 201 = 1206 RTs) there was no statistically significant difference between the mean of the Line 10 RTs of the participants who noted the contradiction, at 3,235 ms, and the mean of the RTs of those who did not, at 3,048 ms, with $p = .191$. As for Study 3, the results do suggest that those participants who noted the contradiction in their answers to question 3 did have longer Line 10 RTs, but only weakly so.

Testing that Longer RTs are evidence for Participants' CD.

The third assumption upon which contradiction paradigm experiments are based is that during participants' longer RTs for contradictory lines, they are feeling cognitive dissonance. Although no formal tests were conducted to check this assumption, the explanation of the experiment's objectives given to participants on completion of their contributions provided an opportunity for them to express their reactions to reading the contradictory lines. Comments that they felt the contradictory lines to be "confusing", "weird", "puzzling", "contradictory" in a tone of indignation or similar were commonly made, suggesting that the assumption is well based. It was also noticeable that after participants had been reading a contradictory story for about half a minute and could be assumed to be reaching line 10, they would commonly furrow their brows, scowl, lean towards the screen, tilt their heads, occasionally let out a yelp, shift their seating position, or otherwise indicate a disturbed state of mind (Livingstone & Palmer, 2015). While the comments and body language do not provide formal and analyzable evidence for an aroused state of CD, they strongly suggest one.

The same types of comments were made during Study 4 about the contradictory lines as were made during Study 3, "weird", etc. Similarly, the facial and body reactions were sometimes witnessed when participants were assessed as having reached to the contradictory lines. Careful observations were conducted to see whether these reactions occurred when participants were reading the non-contradictory stories. They did not.

These contrasts are interpreted as reinforcing the view developed during Study 3 that participants did feel CD after having read a contradictory line.

Appendix I – Topic Sequences

In order to prevent order effects when participants read the stories on the screen, the position of each contradictory topic in the sequence of stories presented to participants was varied so that each topic appeared equally frequently in each of the six positions. Six topic sequences were therefore necessary. In Table 43 below, which illustrates this concept, all sequences commence with a non-contradictory story, with contradictory topics in positions 2, 4, 6, etc., Contradictory topics are in bold.

Table 43
Sequences of topic presentation

Position in reading sequence	Topic Position Variation 1	Topic Position Variation 2	Topic Position Variation 3	Topic Position Variation 4	Topic Position Variation 5	Topic Position Variation 6
1	Non-Contradictory	Non-Contradictory	Non-Contradictory	Non-Contradictory	Non-Contradictory	Non-Contradictory
2	**Surv. Non-Life Thr**	**Honesty**	**Child-Loss**	**Socially Unimp**	**Survival Life Thr**	**Soc. Imp**
3	Non-Contradictory	Non-Contradictory	Non-Contradictory	Non-Contradictory	Non-Contradictory	Non-Contradictory
4	**Honesty**	**Child-Loss**	**Socially Unimp**	**Survival Life Thr**	**Soc. Imp**	**Surv. Non-Life Thr**
5	Non-Contradictory	Non-Contradictory	Non-Contradictory	Non-Contradictory	Non-Contradictory	Non-Contradictory
6	**Soc. Imp**	**Surv. Non-Life Thr**	**Honesty**	**Child-Loss**	**Socially Unimp**	**Survival Life Thr**
7	Non-Contradictory	Non-Contradictory	Non-Contradictory	Non-Contradictory	Non-Contradictory	Non-Contradictory
8	**Child-Loss**	**Socially Unimp**	**Survival Life Thr**	**Soc. Imp**	**Surv. Non-Life Thr**	**Honesty**
9	Non-Contradictory	Non-Contradictory	Non-Contradictory	Non-Contradictory	Non-Contradictory	Non-Contradictory
10	**Survival Life Thr**	**Soc. Imp**	**Surv. Non-Life Thr**	**Honesty**	**Child-Loss**	**Socially Unimp**
11	Non-Contradictory	Non-Contradictory	Non-Contradictory	Non-Contradictory	Non-Contradictory	Non-Contradictory
12	**Socially Unimp**	**Survival Life Thr**	**Soc. Imp**	**Surv. Non-Life Thr**	**Honesty**	**Child-Loss**

Note: Surv. Life-Thr = Survival Life Threatening, Surv. Non-Life Thr = Survival Non-Life Threatening, Soc. Imp = Socially Important, Soc. Unimp = Socially Unimportant

As three stories were created for each topic, three versions or story sequences were required to fully accommodate the topics appearing for each column in Table 43. As an example, Table 44 illustrates the expansion of Variation 1 from Table 43 to accommodate the inclusion of the three stories used for each topic. The stories for each topic have been arbitrarily labelled 1, 2 and 3.

Table 44

Expansion of Topic Position Variation 1 in Table 43 enabling all three stories of each topic to be presented.

Position in sequence	Topic Position Variation 1a	Topic Position Variation 1b	Topic Position Variation 1c
1	Non-Con. 3	Non-Con. 3	Non-Con. 3
2	**Surv. Non Life-Thr 3**	**Surv. Non Life-Thr 2**	**Surv. Non Life-Thr 1**
3	Non-Con. 4	Non-Con. 4	Non-Con. 4
4	**Honesty 2**	**Honesty 1**	**Honesty 3**
5	Non-Con. 5	Non-Con. 5	Non-Con. 5
6	**Soc. Imp 3**	**Soc. Imp 2**	**Soc. Imp 1**
7	Non-Con. 6	Non-Con. 6	Non-Con. 6
8	**Child-Loss 2**	**Child-Loss 3**	**Child-Loss 1**
9	Non-Con. 1	Non-Con. 1	Non-Con. 1
10	**Surv. Life-Thr 1**	**Surv. Life-Thr 2**	**Surv. Life-Thr 3**
11	Non-Con. 2	Non-Con. 2	Non-Con. 2
12	**Socially Unimportant 2**	**Socially Unimportant 3**	**Socially Unimportant 1**

Note: Non-Con. 1 = Non-Contradictory story 1. etc.

As three stories were to be used for all six contradictory topics, all six Variation Columns in Table 43 were expanded in the manner shown in Table 44. A total of eighteen story sequences was therefore generated. The particular sequence of stories presented to participants was selected programmatically by cycling through the eighteen stored sequences.

Appendix J - Eighteen Versions

Table 45 below shows the colour coded key to the cycle of eighteen sequences of the stories presented to participants in Study 3, with each topic having its own topic colour. The abbreviations and their meanings indicate the particular story in each sequence position shown in Table 46.

Table 45

Colour codes as key to story sequences of the eighteen versions

Colour and Abbreviation	Meaning
Non-C1' to Non-C6'	Non-Contradictory stories 1 to 6
SNon-L.X	Survival Non-life Threatening Stories 1, 2 or 3
Rep-H.X	Reputation for Honesty stories 1, 2 or 3
RC.X	Reproduction - Child-Loss stories 1, 2 or 3
Soc-Un.X	Socially Unimportant stories 1, 2 or3
SL.X	Survival Life Threatening stories 1, 2 or 3
Soc-Im.X	Socially Important stories 1, 2 or 3

The first participant read the twelve stories in the sequence shown in the left hand column of Table 46 starting with the Non-contradictory story 3. The next story presented to that participant was the Survival Non-life Threatening story 3, and so on down the column. The next arriving participant was presented with the stories in the next column to the right, this time the Non-life Threatening story being story 2 for this topic and so on down the column. The next participant was presented with the sequence of stories in the next column to the right and so on until all eighteen sequences had been used. The next participant read the same stories in the same sequences as the first, and so on, cycling through the eighteen sequences as many times as necessary. The same procedure was adopted for Study 4.

Table 46

Story sequences for each of the eighteen versions

Story Sequence	Versions																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	NON-C3'	NON-C3'	NON-C3'	NON-C1'	NON-C1'	NON-C1'	NON-C4'	NON-C4'	NON-C4'	NON-C6'	NON-C6'	NON-C6'	NON-C5'	NON-C5'	NON-C5'	NON-C2'	NON-C2'	NON-C2'
2	SNon-L.3	SNon-L.2	SNon-L1	Rep-H.1	Rep-H.2	Rep-H.3	RC.3	RC.2	RC.1	Soc-Un.1	Soc-Un.3	Soc-Un.2	SL.2	SL.1	SL.3	Soc-Im.2	Soc-Im.3	Soc-Im.1
3	NON-C4'	NON-C4'	NON-C4'	NON-C2'	NON-C2'	NON-C2'	NON-C5'	NON-C5'	NON-C5'	NON-C1'	NON-C1'	NON-C1'	NON-C6'	NON-C6'	NON-C6'	NON-C3'	NON-C3'	NON-C3'
4	Rep-H.2	Rep-H.1	Rep-H.3	RC.1	RC.3	RC.2	Soc-Un.1	Soc-Un.2	Soc-Un.3	SL.2	SL.1	SL.3	Soc-Im.1	Soc-Im.2	Soc-Im.3	SNon-L.2	SNon-L.1	SNon-L.3
5	NON-C2'	NON-C2'	NON-C2'	NON-C6'	NON-C6'	NON-C6'	NON-C3'	NON-C3'	NON-C3'	NON-C5'	NON-C5'	NON-C5'	NON-C4'	NON-C4'	NON-C4'	NON-C1'	NON-C1'	NON-C1'
6	Soc-Im.3	Soc-Im.2	Soc-Im.1	SNon-L.2	SNon-L.1	SNon-L.3	Rep-H.1	Rep-H.3	Rep-H.2	RC.3	RC.1	RC.2	Soc-Un.3	Soc-Un.2	Soc-Un.1	SL.1	SL.3	SL.2
7	NON-C5'	NON-C5'	NON-C5'	NON-C3'	NON-C3'	NON-C3'	NON-C6'	NON-C6'	NON-C6'	NON-C2'	NON-C2'	NON-C2'	NON-C1'	NON-C1'	NON-C1'	NON-C4'	NON-C4'	NON-C4'
8	RC 2	RC.3	RC.1	Soc-Un.2	Soc-Un.3	Soc-Un.1	SL.2	SL.1	SL.3	Soc-Im.2	Soc-Im.3	Soc-Im.1	SNon-L.2	SNon-L.3	SNon-L.1	Rep-H.3	Rep-H.2	Rep-H.1
9	NON-C1'	NON-C1'	NON-C1'	NON-C5'	NON-C5'	NON-C5'	NON-C2'	NON-C2'	NON-C2'	NON-C4'	NON-C4'	NON-C4'	NON-C3'	NON-C3'	NON-C3'	NON-C6'	NON-C6'	NON-C6'
10	SL.1	SL.2	SL.3	Soc-Im.3	Soc-Im.2	Soc-Im.1	SNon-L.3	SNon-L.1	SNon-L.2	Rep-H.1	Rep-H.2	Rep-H.3	RC.1	RC.2	RC.3	Soc-Un.1	Soc-Un.2	Soc-Un.3
11	NON-C6'	NON-C6'	NON-C6'	NON-C4'	NON-C4'	NON-C4'	NON-C1'	NON-C1'	NON-C1'	NON-C3'	NON-C3'	NON-C3'	NON-C2'	NON-C2'	NON-C2'	NON-C5'	NON-C5'	NON-C5'
12	Soc-Un.2	Soc-Un.3	Soc-Un.1	SL.3	SL.1	SL.2	Soc-Im.2	Soc-Im.3	Soc-Im.1	SNon-L.3	SNon-L.2	SNon-L.1	Rep-H.3	Rep-H.1	Rep-H.2	RC.3	RC.1	RC.2

Appendix K – Contradictory Stories

Survival - Life threatening

1. The rain forest seemed very dangerous to me.
2. When people think of danger they think of lives.
3. Not even once was my life ever in danger.
4. At night, the wild life noises were very disturbing
5. I worried about drinking the river water.
6. Sometimes I did not want to eat very much.
7. Dysentery can cause a rapid weight loss.
8. I have a few sores and blisters on my legs.
9. I was surprised by the colours in the rain forest
10. Many times, my life was in very great danger
11. I am convinced that I will never explore again.

1. What can cause weight loss?
2. Did you hear animal noises?
3. Was your life in danger?
4. Will you explore again?

1. These notes may record my last coherent words.
2. My rain forest exploration has been disastrous.
3. I live in perpetual fear for my life,
4. I worry a bit about drinking river water.
5. I now know what the word hunger really means.
6. My dysentery is causing a rapid weight loss.
7. I used to have sores and blisters on my legs and arms.
8. I am taking several types of pills to cure them.
9. I was surprised by the colours in the rain forest.
10. I am unconcerned about losing my life.
11. I don’t suppose that I will ever explore again.

1. What is the most worrying?
2. What is the most damaging to your health?
3. Are you worried about surviving?
4. Will you explore in the future?

1. I took a trek through Amazonia recently.
2. Many believe this part of the world in dangerous.
3. I was frightened for my life throughout the journey.
4. There were enormous numbers of strange insects.
5. Some carry various types of diseases.
6. I got swollen sores all over my body.
7. I do not know what the consequences might be.
8. I found the variety of plants interesting.
9. I was surprised by the wide range of colours.
10. Never during the journey did I fear for my life.
11. I do not think I will ever go back there again.

1. Was the trek in the Amazon recent?
2. Did you reach the source of the Amazon?
3. Did you fear loss of life?
4. Will you return to Amazonia?

Survival Non-Life Threatening

1. We think our camping holiday will be challenging.
2. There will be occasional frosts on the hills.
3. Finding warm shelter on cold nights will be difficult.
4. We're told the weather can be very changeable.
5. We are apprehensive about storms and gales.
6. One tent can be carried easily by one person,
7. Two people can sleep in each tent in comfort.
8. There are several forests in the area.
9. We are not expecting anyone to leave early.
10. Finding warm shelter on cold nights will be easy.
11. Some people think we will be too vulnerable

1. Are you expecting frosts?
2. Are the tents heavy or light?
3. Will warm shelter be available on cold nights?
4. How many people sleep in each tent?

1. Out in the wilds, the weather is always a danger.
2. We try to keep all our clothes completely dry.
3. Our survival is dependent upon good shelter.
4. The winds blow noisily when we're in our tents.
5. We should be well protected by them.
6. In case some get wet we have three spare sleeping bags.
7. Our group consists of six experienced people.
8. Hypothermia comes quickly when people get wet.
9. Four days from now we will return home together.
10. We'll survive perfectly well without shelter.
11. It is very important to avoid getting cold.

1. How many people in the group?
2. Do hypothermia and the wet go together?
3. How important in shelter to survival?
4. How many days till the return home?

1. Health, warmth and shelter are closely related.
2. I'm always concerned about my family's health.
3. Our health will suffer without warm clothing in winter.
4. Warm clothing is another form of shelter.
5. We need to plan for the cold if we expect it.
6. Our house has been very well insulated.
7. Some people do not dress warmly enough.
8. Their clothing styles are dictated by fashions.
9. Bright colours are the most important thing now,
10. We don't need warm clothing to stay healthy in winter.
11. Some people do not pay attention to this issue.

1. Should we plan for cold weather?
2. Do bright colours matter?
3. Does your family's health require warm clothing in winter?
4. Are people equally concerned about clothing?

Reproduction - Fidelity

1. We have been together now for seven years.
2. I thought we were well suited to each other.
3. I felt wretched when my partner had affairs.
4. We continued to live together afterwards.
5. We stayed in touch with all our mutual friends.
6. We successfully continued with our careers.
7. We each have our own interests to pursue.
8. Our house recently needed some roof repairs.
9. We bought some new furniture a few days ago.
10. My partner's many affairs did not upset me.
11. I expect we will survive as a couple.

1. How long have you been together as a couple?
2. Do you have separate careers?
3. Did you mind about the partner's affair?
4. What does the writer expect for the future?

1. Marriage is a unique institution.
2. So much time to be spent in ones person's company.
3. I've no sexual affairs since I got married.
4. I believe some of our friends have been unfaithful.
5. We moved house a few months ago.
6. Tasks like that carried out together are important.
7. People seem to differ in all sorts of ways.
8. Some people don't like talking about themselves.
9. I don't mind talking about my own love life.
10. I've had eight sexual affairs since I got married.
11. I often discuss loyalty with my friends.

1. Is marriage unique as an institution?
2. Have you friends you have been unfaithful?
3. Have you had any sexual affairs outside marriage?
4. Do you think of yourself as discreet?

1. Faithfulness seems to be in decline today.
2. We know lots of couples in open relationships.
3. My partner's affairs cause me no jealously.
4. Some of our friends are very introspective.
5. They discuss their own temperaments a lot.
6. They know they have no inclination to play around.
7. They say they behave accordingly in their lives.
8. My partner's parents visited yesterday.
9. Both of them stay healthy by walking a lot.
10. My partner's affairs make me intensely jealous.
11. Perhaps we will discuss this subject one day.

1. Are you in a relationship?
2. Are people in open relationships mentioned?
3. Do you feel jealousy?
4. How many visitors did you have yesterday?

Reproduction-Child-Loss

1. It was a short and most unexpected illness.
2. People were kind and tried to console us both.
3. I grieved a long time after the death of our child.
4. Our daily life took a new turn afterwards.
5. There is always something interesting happening.
6. We went on a climbing holiday recently.
7. I returned home with a sprained ankle.
8. There were even some children in our party.
9. Children don't seem to enjoy the scenery much.
10. I didn't grieve at all after the death of our child.
11. Life is now pretty much back to normal.

1. How long was your child's illness?
2. Was it expected?
3. Did you grieve?
4. What is life like now?

1. We always planned to have a small family
2. Our child died when only eleven months old.
3. I will never recover from my child's death.
4. We talk together about trying again.
5. Of course, there are many other things in life.
6. We have made several new friends recently.
7. They introduced us to new activities.
8. We went to a children's holiday camp last month.
9. Some of the children were really still infants.
10. I will quickly recover from my child's death.
11. We do not know ourselves until we are tested.

1. How many children are mentioned?
2. Do you have hobbies?
3. Do you think recovery from loss is near?
4. Are you introspective?

1. We started our own family soon after our wedding.
2. Not long after, we had to make birth arrangements.
3. Our first child died suddenly while still a baby.
4. We have had an average sized family.
5. The children provided us with much satisfaction.
6. They are all adults now and have vacated the nest.
7. Our house seems so empty now that they have gone.
8. We remember raising children as hard work.
9. They are raising their own children now.
10. We've not suffered the loss of any of our children.
11. Family life is a mixture of ups and downs.

1. How many children are mentioned?
2. How many are adult?
3. Have you suffered any child loss?
4. Have you any children at home?

Reproduction Family

1. We are very happy to have got engaged.
2. We have already bought a very nice house.
3. We look forward to having several children.
4. We have a lot of interests in common.
5. Both us were born into quite small families.
6. Only one other child was born on either side.
7. So we have just one sibling each, both brothers.
8. We each have one uncle and one aunt.
9. Both the aunts are now pregnant at the same time.
10. We have decided not to have any children.
11. We look forward to our new life together.

1. What has recently been bought?
2. As a couple, do you have interests in common?
3. Will you have children?
4. How many siblings do you each have?

1. We have given a lot of thought to family life.
2. We both have several cousins with lots of kids.
3. We are firm in not wanting any children.
4. We also have several unmarried cousins.
5. They vary in their attitude to having children.
6. We have an annual family gathering.
7. The location varies from one year to the next.
8. We hired a seven bedroom farm house last year.
9. Adults slept at one end and children at the other.
10. We are both excited about having children.
11. The family continues to grow in size each year.

1. Are unmarried cousins mentioned?
2. How often does the family gather?
3. Do you want any children?
4. Is the size of the family shrinking?

1. Family life is often the most important thing.
2. For some people children come before everything.
3. Our children will be my life's greatest purpose.
4. For other people careers come first and last.
5. Most people start their adult life with a job.
6. Children usually play a part some years later.
7. Our two children were born just over four years ago.
8. Children provide a purpose to life if you choose.
9. Life's purpose is a tricky issue to fathom out.
10. Our children will provide no purpose in my life.
11. For some people, children are life's fullest reward.

1. How many children have you?
2. How many years ago were they born?
3. Do your children provide a purpose to your life?
4. Do you agree that adult life usually starts with a job?

Reputation-Reciprocity

1. There is a great deal of mutuality in life.
2. We frequently do things for other people.
3. People remember whether you return favours.
4. People concern themselves with what we do.
5. Our actions are always important to them.
6. They form opinions of us accordingly.
7. We form opinions of them similarly.
8. Our world consists of other peoples' opinions.
9. We like to keep in other people's good books.
10. People don't remember when you return favours.
11. I wish people would think of each other a bit more.

1. Do we always think of others in our actions?
2. Are we concerned with other people's opinions of us?
3. Do people remember if you return favours?
4. Do we return favours as often as we should?

1. My neighbours often do little jobs for me.
2. I try to provide assistance to them also.
3. Returning favours ensures one a good name.
4. We need the help we receive from each other.
5. It's what will keep us in each others good books.
6. Some other people almost totally ignore you.
7. What you think of them doesn't seem to matter.
8. We worry about what other people think of us.
9. Other people talk about us behind our backs.
10. Returning favours won't ensure one a good name.
11. Nevertheless, peoples' attitudes are important.

1. Do you have kind neighbours?
2. What makes life easier?
3. Does returning favours get you a good name?
4. Was the word "backs" mentioned?

1. Kindness seems to be part of human nature.
2. It always seems to be much appreciated.
3. I always return kindness from others with kindness.
4. What other people think is important to me.
5. We keep it in mind when others have been kind to us.
6. We know they'll do the same when we're kind to them.
7. My mother quite often gives to charities
8. Most of them are kind and honest but some are not.
9. She hopes they'll return the favour some time.
10. I don't return kindness from others with kindness.
11. I wonder what other people think of me.

1. Do you agree that some people are kinder than others?
2. Are we all unkind sometimes?
3. Do you reciprocate kindness with kindness?
4. Was your father mentioned?

Reputation- Honesty

1. People gossip about people who cheat others.
2. Some types of cheating attract instant scandal.
3. Cheating greatly affects our social standing.
4. Our social circle is constantly changing.
5. We try to keep in mind what people have done.
6. Most people regard some crimes as worse than others.
7. The police take violent crimes more seriously.
8. Criminal court records are never destroyed.
9. They are always available to be consulted.
10. Cheating has no effect on our social standing.
11. People take note of what we do and don't do.

1. Is the word hypocrisy mentioned?
2. Is the word gossip mentioned?
3. Does cheating affect our social standing?
4. Are you forgiving?

1. People pay attention to the views of others.
2. I have to remember that in my business.
3. I really mind if people say I cheat on them.
4. Some people ignore other people's opinions.
5. Others can't afford to be so dismissive.
6. People discuss other people all the time.
7. It's how we learn if someone is trustworthy.
8. My business works with very narrow margins.
9. Many of my competitors are closing down.
10. I don't mind if people say I cheat on them.
11. My family think highly of my business skills.

1. Are you in business?
2. Do people discuss each other?
3. Do you find cheating a problem?
4. Do your family admire you?

1. I got my job in the store straight after school.
2. Money and a good name must go together.
3. It matters a lot if people think I'm a crook.
4. I was promoted twice at my current branch.
5. I believe I am quite well thought of in this store.
6. Promotion depends on management's opinion.
7. What they think of you is all important.
8. It's also important to sell goods in high demand
9. You can detect customers' views from their faces.
10. It doesn't matter if people think I'm a crook.
11. Ones good name is only one aspect of life.

1. When did you start work?
2. What do you count?
3. Do people mind about crooked dealings?
4. Are you in retailing?

Non-Adaptive Socially Important

1. Parking is becoming more difficult every day.
2. New rules are being introduced every day.
3. I got my second parking ticket yesterday.
4. More people and more traffic create more problems.
5. It is becoming difficult to find a place to park.
6. Now I try to avoid coming in to town.
7. Car use is increasingly regulated.
8. One friend told me that she has even sold her car.
9. This way, she should avoid getting a parking ticket.
10. I've never had a parking ticket issue to me.
11. Perhaps I should learn to ride a bicycle.

1. Does the writer try to avoid coming to town?
2. What sex was the friend who sold a car?
3. Have you had any parking tickets?
4. What other means of transport is mentioned?

1. Safety regulations are becoming stricter.
2. Speed limits are being reduced every day.
3. I got issued a speeding ticket yesterday.
4. There were very few pedestrians around.
5. The road had a lot of trees in leaf on both sides
6. I believe accident statistics are declining.
7. Lower speeds should mean less severe accidents.
8. People should therefore suffer fewer injuries.
9. Speeding ticket numbers might decline.
10. I have never been issued with a speeding ticket.
11. I suppose greater restrictions lead to a public good.

1. Are traffic accident figures declining?
2. Are driving speeds declining?
3. Have you ever been issued with a speeding ticket?
4. Do you agree that lower speeds should mean less severe accidents?

1. The standard of people's driving is important.
2. Some people are cautious even on empty roads.
3. It's OK driving through red lights if there's no traffic.
4. Traffic lights are there to prevent collisions.
5. They impose a social discipline on us all.
6. With more traffic came more traffic lights.
7. Now they are a normal part of everyday life.
8. The first one was installed in the nineteenth century.
9. At that time there was very little traffic.
10. It's OK driving through red lights if there's no traffic
11. Nothing is ever likely to eliminate them.

1. Was the first traffic light installed over a hundred years ago?
2. Do you agree that they are a normal part of today's life?
3. Is it OK to drive through a red light in some circumstances?
4. Will we have traffic lights forever.

Non-Adaptive Socially Unimportant

1. Joan decided to visit seven shops today.
2. She uses four paper carrier bags when shopping.
3. The striped paper bag is full to the top.
4. The others are just plain with different colours.
5. Four carrier bags are usually sufficient.
6. Her purchases do not completely fill all the bags.
7. She carries them all in a wheeled shopping basket.
8. She protects the bags from rain with a plastic sheet.
9. She notices the paper is getting creased and worn.
10. The striped paper bag is completely empty.
11. Once home she folds each bag carefully for next time.

1. How many shops did Joan visit?
2. How many bags did she have?
3. Were any of the bags full?
4. Did it rain?

1. I first made papier-mâché when I was eight.
2. All you need is water, glue and some paper.
3. I love making papier-mâché from newspaper.
4. I also use old plain paper if I have any.
5. The paper must be able to absorb water.
6. I put glue between the layers of soggy paper.
7. I don\'t like the glue getting on my fingers.
8. Using a brush will normally prevent this.
9. Like the water, the glue must be well absorbed.
10. I hate making papier-mâché from newspaper.
11. One can make intricate shapes with perseverance.

1. What age were you when you first used papier-mâché?
2. Do you use plain paper?
3. What is your attitude to papier-mâché made of newspaper?
4. What do you use to apply the glue?

1. Sebastian bought Rowan a birthday present.
2. It was Rowan's birthday in a fortnight.
3. Sebastian had lots of wrapping paper.
4. There would be twenty people at the birthday party.
5. He wanted Rowan to recall who gave him his present.
6. He foresaw presents buried under wrapping paper.
7. How would Rowan remember who gave him what?
8. After searching, he bought Rowan a book.
9. He wondered if Rowan would appreciate his choice.
10. Sebastian had no wrapping paper at all.
11. The book's title was "Sebastian's present".

1. How many days to go till Rowan's birthday?
2. How many people will be at the birthday party?
3. How much wrapping paper has Sebastian got?
4. What type of object did Sebastian buy?

Appendix L - Participants Instructions

Thank you for participating in this research. It has been approved by Massey University Ethics Committee as fully complying with its Code of Ethical Conduct. We will interpret your progressing beyond this screen as indicating your informed consent to participate fully in this experiment. However, you may abandon your participation at any time by closing your browser or the tab in which you see this page.

In order to obtain the best results from your participation, we request that you close all applications running on your computer other than your browser.

This experiment consists of your reading a set of 12 stories of 11 lines each presented one line at a time on a screen. It will be preceded by a practice story of 6 lines.

All stories are followed by a page containing 4 short questions.

Whenever you are presented with a screen displaying one line of a story, please read it and press the space bar for the next line.

Whenever you are presented with a screen with 4 questions, please provide short typed replies in the boxes to the right. Then press the Enter key.

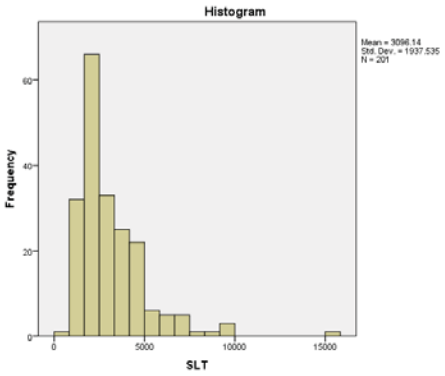
If you have left your email address in the previous screen, at the conclusion of the research we will be pleased to send you an account of what was achieved. You can also leave your email address after the last question.

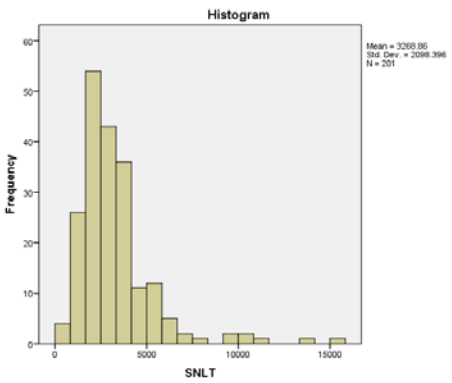
Thank you again. When you are ready, please press the space bar to proceed to the practice story.

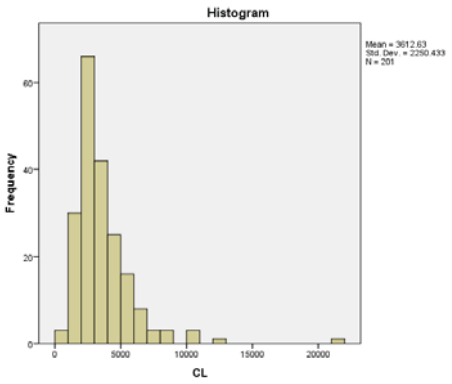
Appendix M – Histograms Illustrating the Effect of Outlier Exclusion

Study 3

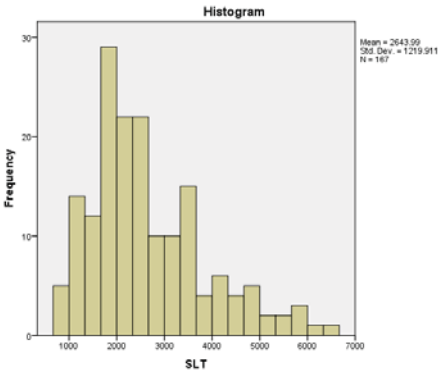
Histograms of Line 10 RTs with all 201 Participants

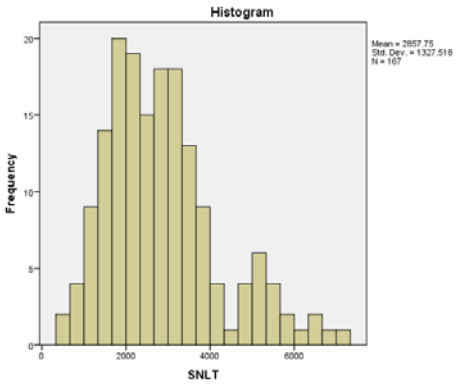


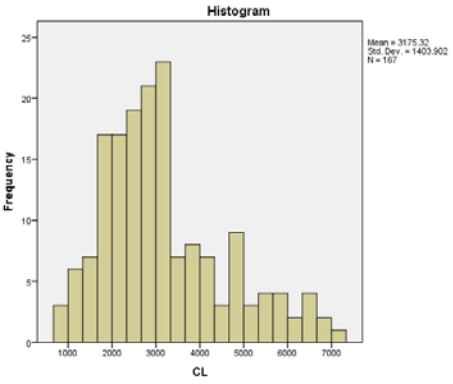




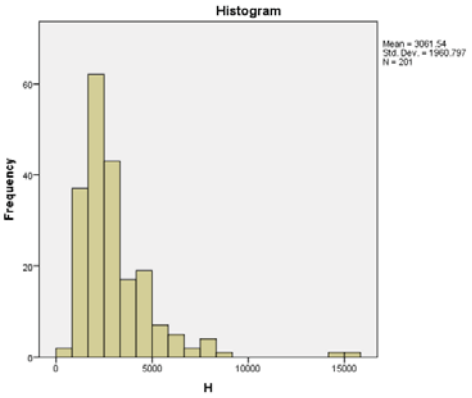
Histograms of Line 10 RTs with 2 X SD Outliers excluded

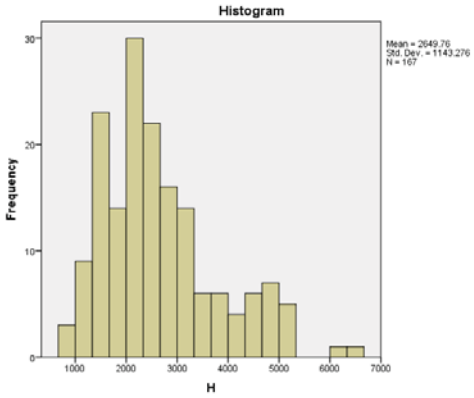


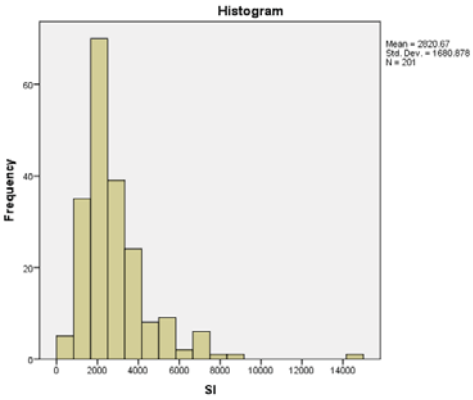


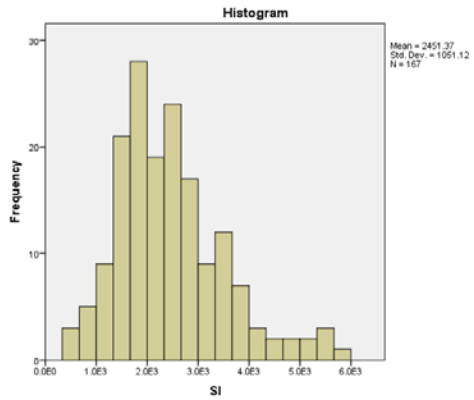


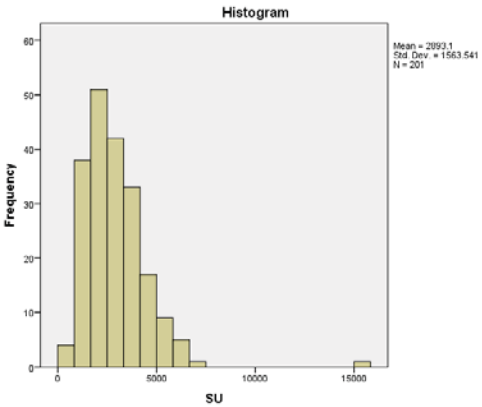
Appendix M -

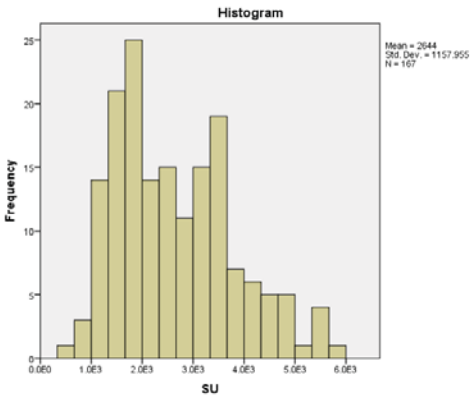

Histogram
Mean = 3061.54
Std. Dev. = 1960.797
N = 201
Frequency
H


Histogram
Mean = 2649.76
Std. Dev. = 1143.276
N = 167
Frequency
H

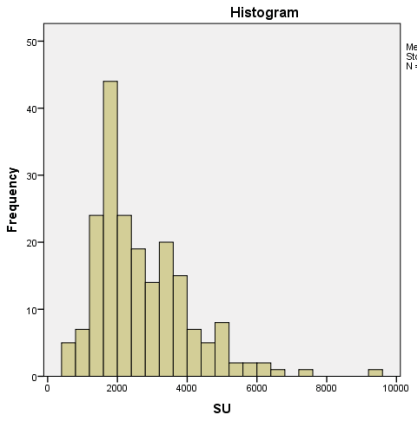

Histogram
Mean = 2820.67
Std. Dev. = 1680.878
N = 201
Frequency
SI

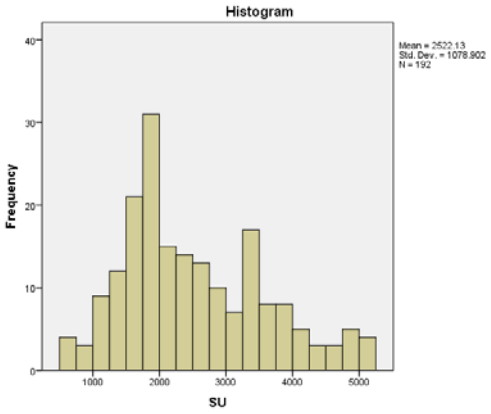

Histogram
Mean = 2451.37
Std. Dev. = 1051.12
N = 167
Frequency
SI

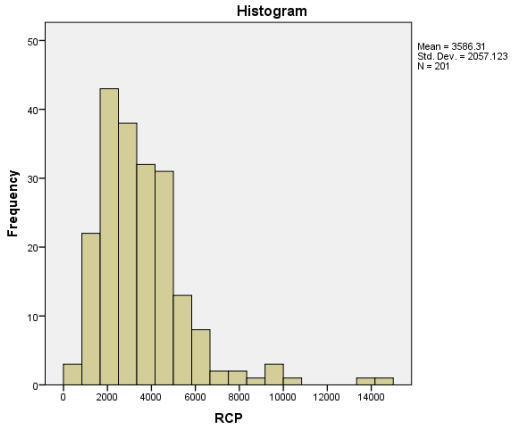

Histogram
Mean = 2093.1
Std. Dev. = 1563.541
N = 201
Frequency
SU

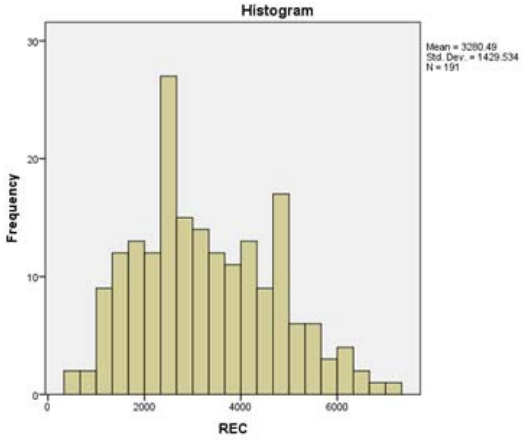

Histogram
Mean = 2644
Std. Dev. = 1157.955
N = 167
Frequency
SU

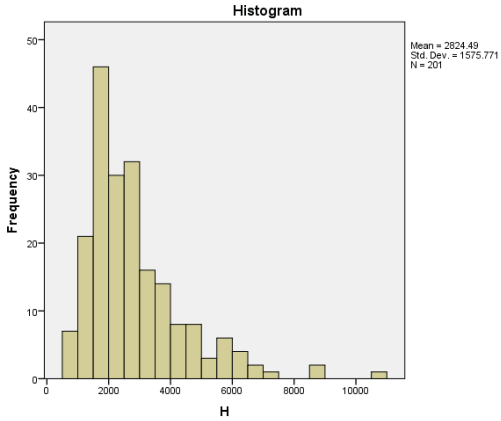
Study 4

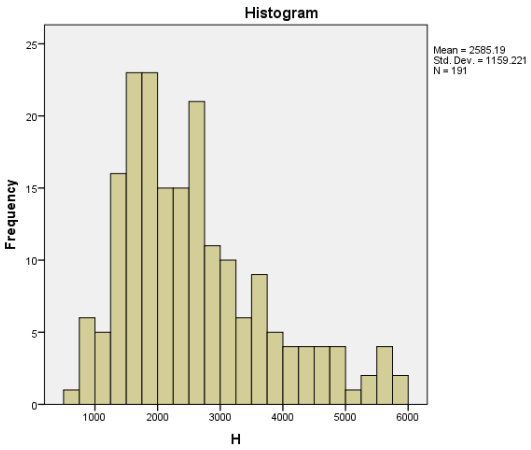

Histogram
Mean = 2700.36
Std. Dev. = 1363.345
N = 201
Frequency
SU


Histogram
Mean = 2522.13
Std. Dev. = 1078.902
N = 192
Frequency
SU


Histogram
Mean = 3586.31
Std. Dev. = 2057.123
N = 201
Frequency
RCP


Histogram
Mean = 3280.49
Std. Dev. = 1429.534
N = 191
Frequency
REC


Histogram
Mean = 2824.49
Std. Dev. = 1575.771
N = 201
Frequency
H


Histogram
Mean = 2585.19
Std. Dev. = 1159.221
N = 191
Frequency
H

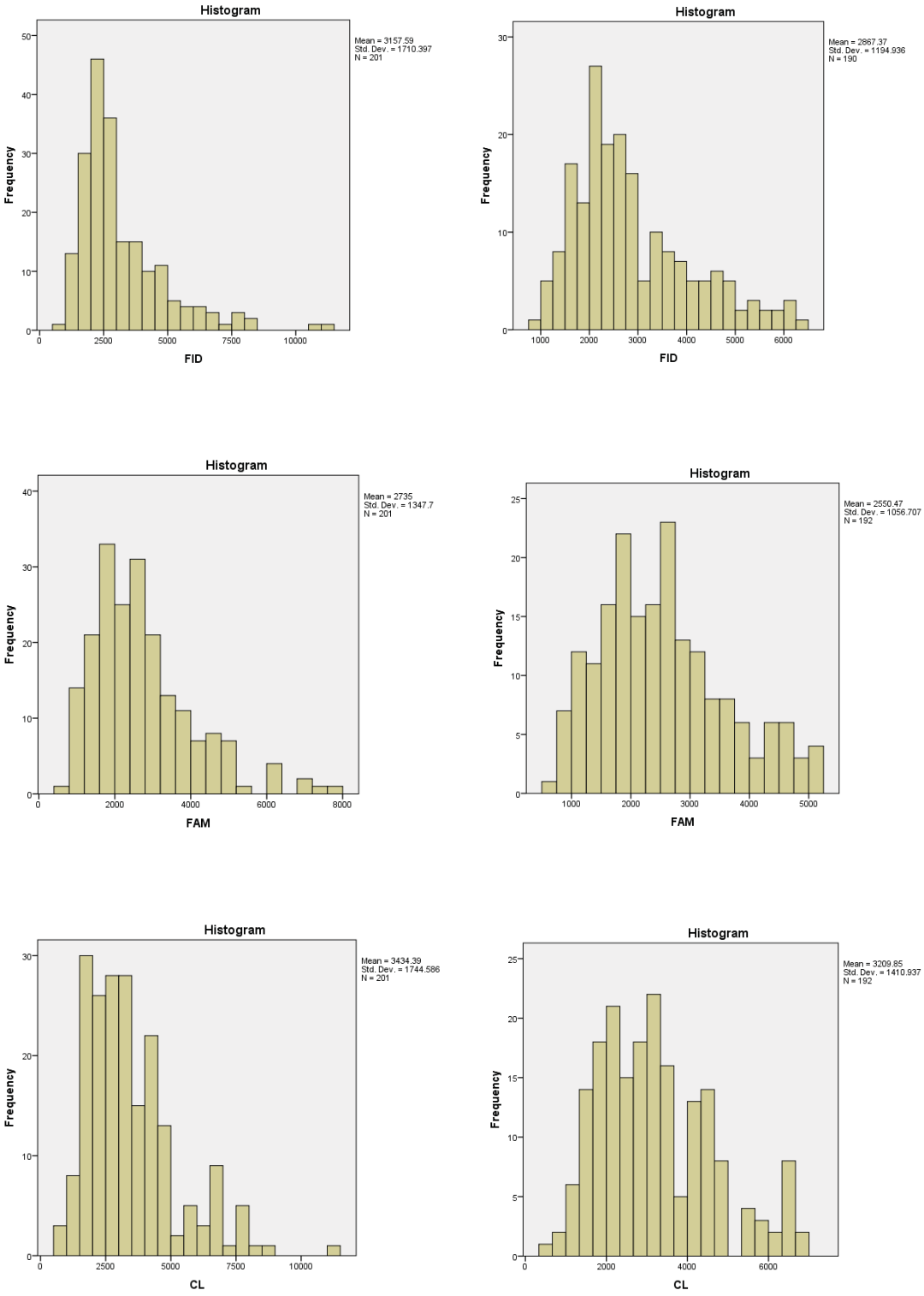


Figure 9 Histograms Illustrating the Effect of Outlier Exclusion

Appendix N – Triplet ANOVA Analyses

In Table 47 below, the means are shown for each topics' triplets of contradictory line 10 RTs. Socially Unimportant, Child-Loss and Honesty topics were used in Studies 3 and 4. This permitted a comparison to be made across the studies, the expectation being that RTs for the stories would be similar in each case. In all three cases, the highest, middle and lowest RT means pertained to the same stories in both studies. Table 36 also shows the contradictory scores obtained for each contradictory pair in Study 1. For those triplets which showed statistically significant differences between the means of the three stories' RTs, separate Mean RTs are listed for the two sexes. Two way ANOVAs between Sex and Story Number failed to show a significant interaction effect for any of these triplets.

Table 47

Line 10 RTs for each story comprising a topic. All times in ms.

Study	Topic	Contr. Pair	Text	Contr. Score	N	Mean RTs		
						All PTs without Outliers	Fem (N)	Males (N)
3	Surv Life Thr.	1	Not even once was my life ever in danger. Many times, my life was in very great danger.	6.383	65	2931		
		2	I live in perpetual fear for my life. I am unconcerned about losing my life.	6.29	67	2577		
		3	I was frightened for my life throughout the journey. Never during the journey did I fear for my life.	6.787	70	2947		
3	Surv Non-Life Thr.	1	Finding warm shelter on cold nights will be difficult. Finding warm shelter on cold nights will be easy.	6.5	61	2845		
		2	Our survival is dependent upon good shelter. Well survive perfectly well without shelter	6.145	67	2751		

		3	Our health will suffer without warm clothing in winter. We don't need warm clothing to stay healthy in winter.	6.161	65	3250		
3	Child-Loss	1	I grieved a long time after the death of our child. I didn't grieve at all after the death of our child.	6.637	65	3354	3137 (38)	3659 (27)
		2	I will never recover from my child's death. I will quickly recover from my child’s death.	6.58	66	2694	2604 (35)	2796 (31)
		3	Our first child died suddenly while still a baby. We've not suffered the loss of any of our children	6.367	62	3862	3654 (31)	4070 (31)
4	Child-Loss	1	I grieved a long time after the death of our child. I didn't grieve at all after the death of our child.	6.637	67	3179	3056 (34)	3326 (33)
		2	I will never recover from my child's death. I will quickly recover from my child’s death.	6.58	65	2836	2427 (27)	3127 (38)
		3	Our first child died suddenly while still a baby. We've not suffered the loss of any of our children	6.367	61	3713	3621 (30)	3810 (31)

3	Honesty	1	Cheating greatly affect our social standing Cheating has no effect on our social standing	6.2	67	2849
		2	I really mind in people say I cheat on them. I don't mind if people say I cheat on them.	6.246	63	2813
		3	It matters a lot if people think I am a crook. It doesn't matter if people think I'm a crook.	6.092	63	2681
4	Honesty	1	Cheating greatly affects our social standing. Cheating has no effect on our social standing.	6.2	62	2714
		2	I really mind in people say I cheat on them. I don’t mind if people say I cheat on them.	6.246	65	2697
		3	It matters a lot if people think I am a crook. It doesn't matter if people think I'm a crook.	6.092	64	2347
3	Soc. Imp.	1	I got my second speeding ticket yesterday. I've never had a parking ticket issued to me.	6.295	64	2516
		2	I got issued a speeding ticket yesterday. I have never been issued with a speeding ticket.	6.557	67	2710

		3	It's OK driving through red lights if there's no traffic People should never drive through a red traffic light.	6.612	61	2487
3	Soc. Unimp	1	The striped paper bag is full to the top. The striped paper bag is completely empty.	6.6	65	2936
		2	I love making papier-mâché from newspaper. I hate making papier-mâché from newspaper.	6.6	67	2873
		3	Sebastian had lots of wrapping paper. Sebastian has no wrapping paper at all.	6.75	54	2459
4	Soc. UnImp	1	The striped paper bag is full to the top. The striped paper bag is completely empty.	6.6	64	2894
		2	I love making papier-mâché from newspaper. I hate making papier-mâché from newspaper.	6.6	63	2537
		3	Sebastian had lots of wrapping paper. Sebastian has no wrapping paper at all.	6.75	66	2398

4	Reciprocity	1	Returning favours ensures one a good name. Returning favours won't ensure one a good name.	6.517	62	3085	3143 (25)	3046 (37)
		2	People remember whether you return favours. People don't remember whether you return favours.	6.292	67	3076	3221 (35)	2917 (32)
		3	I always return kindness from others with kindness. I don't return kindness from others with kindness.	6.283	62	3697	3427 (30)	3950 (32)
4	Family	1	We look forward to have several children. We have decided not to have any children.	6.125	64	2451	2337 (34)	2580 (30)
		2	We are firm in not wanting any children. We are both excited about having children.	6.081	66	2297	2302 (30)	2293 (36)
		3	Our children will be my life's greatest purpose. Our children will provide no purpose in my life.	6.689	62	2923	2888 (27)	2948 (35)
4	Fidelity	1	I felt wretched when my partner had affairs. My partner's many affairs did not upset me.	6.492	66	2930		
		2	I have had no sexual affairs since I got married. I've had eight sexual affairs since I got married.	6.249	62	2903		

3	My partner's affairs cause me no jealousy. My partner's affairs make me intensely jealous.	6.156	62	2766

Note: In the columns for the Female and Male Means of RTs, the number of participants are in brackets (). Surv Life Thr. = Survival Life Threatening, Surv Non-Life Thr. = Survival Non-Life Threatening, Soc. Unimp = Socially Unimportant

In both studies 3 and 4, there was a statistically significant difference between the Line 10 RTs of the stories 2 and 3 of the Child-Loss triplet, story 2 having the shortest Line 10 RTs and story 3 having the longest in both studies. Story 2 had the following contradictory pair:

I will never recover from my child's death.
I will quickly recover from my child's death.

Story 3 had the following contradictory pair:
Our first child died suddenly while still a baby.
We've not suffered the loss of any of our children

There was a difference between the two studies. In Study 3, one-way ANOVAs with story number as the independent variable indicated that the RT differences reached statistical significance for both sexes. In Study 4, the male result failed to reach significance, F (2, 99) = 1.805, p = .170, implying a slightly lower male sensitivity to these contradictions.

Examination of the two contradictions indicates that story 2 has a contradiction in the style of those proposed by De Marneffe et al. (2008) with 'never recover' opposed by 'quickly recover'. Story 3's contradiction on the other hand is present only by implication, leading to an expectation that the mean of story 3' RTs would be shorter rather than longer due to some participants not having understood the implication, albeit a fairly obvious one. This unanticipated result is compounded by the potential participant surprise when the narrator states in story 2 that s/he will quickly recover from the death of the child, an unusual reaction. The solution may lie in an analogy with TMT. The distress of reading the term 'child's death' once more might lead some participants to quickly press the space bar.

For the Family stories, the statistically significant difference between the RTs of two stories was also between stories 2 with the shorter RTs and 3 with the longer ones. Story 2 is shown below.

We are firm in not wanting any children.
We are both excited about having children.

Story 3 is shown below

Our children will be my life's greatest purpose.
Our children will provide no purpose in my life.

Separate one way ANOVAs for the two sexes RTs shows both sets reaching statistical significance, the female result being F (2, 88) = 3.333, p = .04 and the male one being F (2, 98) = 3.191, p = .045. However, examination of the Multiple Comparisons of the two ANOVAs shows that that the story 2/3 difference for the female participants did not reach conventional statistical significance p = .061, whereas the male one did, p = .035. Notwithstanding these differences, the explanation for the greater sensitivity to the story 3 contradiction may lie in the contrast between 'greatest purpose' and 'no purpose', a contradiction in the style of de Marneffe et al. (2008).

The Reciprocity stories showed below.

Returning favours ensures one a good name.
Returning favours won't ensure one a good name.

People remember whether you return favours.
People don't remember whether you return favours.

I always return kindness from others with kindness.
I don't return kindness from others with kindness.

The one way ANOVA result for all participants, F (2, 188) = 4.027, p = .019, showed that story 3 RTS was statistically longer than those for the other two stories, the Multiple Comparisons table showing p = .043 for the story 1:story 3 contrast and p = .035 for the story 2:story 3 contrast. However, when the ANOVA was repeated for the two sexes separately, the statistical significance was revealed to be due to male distinctions only, the female ANOVA result being F (2, 87) = 0.295, p = .745, while the male one was F (2,98) = 5.64, p = .006. The male story distinctions showed p = .02 for the 1:story 3 contrast and p = .009 for the story 2:story 3 contrast. While female participants found all three stories' contradictions of similar magnitude, as indicated in Table 47, male participants found story 3's particularly noticeable. As all three contradictions match De Marneffes et al.'s (2008) suggestions well, another explanation

is required. One may lie in the use of the words ‘favours’ and ‘kindness’. Females might regard them as rather similar. Males might regard kindnesses as particularly powerful, evoking greater attention, owing to a greater degree of empathy required for the recipient as a precondition for their provision, males in general exhibiting lower levels of empathy (Baron-Cohen, 2003). This explanation is obviously tentative but the issue might justify further exploration.

REFERENCES

Aboitiz, F., & V, R. G. (1997). The evolutionary origin of the language areas in the human brain. A neuroanatomical perspective. *Brain Research Reviews, 25*, 3, 381-96.

Aitchison, J. (2000). *The seeds of speech*. Cambridge; Cambridge University Press.

Allport, G. W. (1943). The ego in contemporary psychology. *Psychological Review, 50*, 5, 451-478.

Alperson-Afil, N. (2008). Continual fire-making by hominins at Gesher Benot Ya'aqov, Israel. *Quaternary Science Reviews, 27*, 1733-1739.

Ambrose, S.H., (2010). Coevolution of composite-tool technology, constructive memory, and language implications for the evolution of modern human behavior. *Current Anthropology* 51 (Supplement 1), S135–47.

Arndt, J., Greenberg, J., & Cook, A. (2002). Mortality salience and the spreading activation of world-view relevant constructs: Exploring the cognitive architecture of terror management. *Journal of Experimental Psychology: General, 131*, 3, 307-324.

Arnold, K., & Zuberbühler, K. (2008). Meaningful call combination in a non-human primate. *Current Biology, 18*, 5, 871-888.

Aronson, E. (1969). The theory of cognitive dissonance: A current perspective. *Advances in Experimental Social Psychology, 4*, 1-34.

Aronson, E. (1991). *The social animal*. (6th Ed.) New York: W. H. Freeman.

Aronson, E. (1999). Dissonance, Hypocrisy, and the Self-Concept. In E. Harmon-Jones and J. Mills (Eds.) *Cognitive Dissonance: Progress on a pivotal theory in social psychology.* Washington DC: American Psychological Association

Aronson, E., & Mills, J. (1959). The Effect of Severity of Initiation in Liking for a Group. *The Journal of Abnormal and Social Psychology. 59*, 177-81.

Aunger, R., & Curtis, V. (2013). The Anatomy of Motivation: An Evolutionary-Ecological Approach. *Biological Theory, 8,* 1, 49-63.

Averil, J. R. (1980). On the poverty of positive emotions. In K. R. Blankstein, P. Pliner, and J. Polivy, (Eds.) *Assessment and Modification of Emotional Behavior*. New York: Plenum Press.

Averil, J. R. (1975). *Semantic Atlas of Emotional Concepts*. Washington, DC: APA

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research. 3,* 2, 12-28.

Baddeley, A. D. (2000). The episodic Buffer: A new component of working memory? *Trends in Cognitive Science*. *4*, 417-23.

Baddeley, A. D. (1995). The Psychology of Memory. In A. D. Baddeley, M. D Kopelman, & B. A. Wilson, (Eds.) *Handbook of Memory Disorders.* Hoboken, NJ: John Wiley and Sons.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G.H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47–89). New York: Academic Press.

Barkow, J. H., Cosmides, L., & Tooby.J. (1992). *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. Oxford, UK: OUP

Barnett, V. & Lewis, T. (1995). *Outliers in Statistical Data* (3rd edition). New York: Wiley.

Baron-Cohen, S. (2003). *Essential Difference: Men, Women and the Extreme Male Brain*. London, Allen Lane

Barrett, H. C., & Kurzban, R. (2006). Modularity in cognition: Framing the debate. *Psychological Review*, 113, 3, 628-647. http://dx.doi.org/10.1037/0033-295X.113.3.628

Barrett, J. L., & Johnson, A. H. (2003). The Role of Control in Attributing Agency to Inanimate Objects. *Journal of Cognition and Culture*, *3*, 3, 208- 217.

Barrett, J. L., & Lanman J. A. (2008). The Science of Religious Beliefs. *Religion, 3*, 3, 109-24.

Bartels, A., & Zeki, S. (2000) The neural basis of romantic love. *NeuroReport, 11*, 3829-34.

Bartlett, F. C. (1932). *Remembering: A Study in Experimental and Social Psychology*. Cambridge, UK, Cambridge UP.

Baumeister, R. F, (2005). *The Cultural Animal: Human Nature, Meaning and Social Life*. Oxford: OUP.

Ben-Gal, I. (2010). Outlier Detection. In Oded Maimon & Lior Rokach (Eds.). *Data Mining and Knowledge Discovery Handbook*. (2nd Ed.) New York: Springer.

Berman, I. R., & Bird, C. (1933). Sex Differences in speed of reading. *Journal of Applied Psychology*; *17*, 3, p. 221.

Bickerton, D. (1990). *Language and Species.* Chicago*:* University of Chicago Press.

Blanca, M. J. Arnau, J., López-Monteil, D., Bono, R., & Bendayan, R. (2013). Skewness and Kurtosis in Real Data Samples. *European Journal of Research Methods for the Behavioral and Social Sciences, 9,* 2, 78-84.

Boas, F. (1916). Eugenics, *The Scientific Monthly, 3*, 5, 471-8.

References

Botvinick, M., Nystrom, L. E., Fissel, K., Carter, C. S. & Cohen, J. D. (1999). Conflict monitoring versus selection for action in anterior cingulate cortex. *Nature. 402*, 11, 179-8.

Bowles, S., & Gintis, H. (2003). The Origins of Human Cooperation. In Peter Hammerstein (Ed.) *Genetic and Cultural Evolution of Cooperation*. Massachusetts. MIT Press

Braasch, J. L., Rouet, J-F., Vibert, N., & Britt, M. A. (2012). Readers' use of source information in text comprehension. *Memory and Cognition, 40,* 3, 450 – 465.

Bräuer, J., & Call, J. (2011). The Magic Cup: Great Apes and Domestic Dogs (Canis familiaris) Individuate Objects According to Their Properties. *Journal of Comparative Psychology*. *125*, 3, 353-361. DOI: 10.1037/a0023009

Brehm, J. W. (1956). Postdecision changes in the desirability of alternatives. *Journal of Abnormal and Social Psychology, 52,* 384-9.

Brehm, J. W., & Cohen, A. R. (1962). *Explorations in cognitive dissonance*. New York; Wiley.

Brouwer, H. (2014). The Electrophysiology of Language Comprehension: A Neurocomputational Model. PhD Thesis submitted to Universtät des Saarlandes, Downloaded from http://www.researchgate.net/publication/262730490

Brouwer, H., & Hoeks, J. C. (2013). A time and place for language comprehension: mapping the N400 and the P600 to a minimal cortical network. *Frontiers of Human Neuroscience, 7*, 758. DOI: 10.3389/fnhum.2013.00758.

Brown, S. (2015). *Measure of Shape: Skewness and Kurtosis*. Retrieved from http://brownmath.com/stat/shape.htm on 27/02/2016.

Brown, G., & Yule, G. (1983). Topic and representation of discourse content. In *Discourse Analysis*. Cambridge, UK: CUP. Retrieved from http://dx.doi.org/10.1017/CBO9780511805226

Bulmer, M. G. (1979). *Principles of Statistics*. Retrieved from http://bobweigel.net/csi763/images/Bulmer_Principles_of_Statistics_1979_all.pdf

Burke, B. L., Martens, A., & Faucher, E. H. (2010). Two decades of Terror Management Theory: A Meta-Analysis of Mortality Salience Research. *Personality and Social Psychology Review, 14*, 2, 155-195.

Burns, J. (2015, January 7). Children's mental health is parent's greatest concern. Retrieved from http://www.bbc.com/news/education-30701591.

Buss, D. (1989). Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. Behavioral and Brain Sciences, 12, 1-49. doi: 10.1017/S0140525X00023992

Buss, D. M. (Ed.) (2005). *Handbook of Evolutionary Psychology*. New Jersey: Wiley.

Buss, D. M., & Duntley, J. D. (2003). Homicide: An evolutionary perspective and implications of public policy. In N. Dess, (Ed.) *Violence and Public Policy*. Greenwood Publishing Group.

Buss, D. M., Haselton, M. G., Shackelford, T. K., Bleske, A. L., & Wakefield.J. C. (1998). Adaptations, Exaptations, and Spandrels. *American Psychologist*. *53*, 5, 533-548.

Buss, D. & Schmitt, D. (2011). Evolutionary psychology and feminism. Sex Roles, 64(9-10), 768-787. doi: 10.1007/s11199-011-9987-3

Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between Attitudes and Evaluative Space: A critical review. *Psychological Bulletin, 115*, 3, 401- 423.

References

Cardinal, R. N., & Aitken, M. R. F. (2006). *ANOVA for the Behavioural Sciences Researcher.* Lawrence Erlbaum Associates. NJ: Mahwah.

Caron, J. (1992). *An introduction to Psycholinguistics*. Toronto, University of Toronto Press.

Carroll. J. (2004). *Literary Darwinism: Evolution, Human Nature and Literature*. NY: Routledge

Carter, C.S., & Van Veet, V. (2007). Anterior cingulate cortex and conflict detection: An update of theory and data. *Cognitive, Affective, & Behavioral Neuroscience, 7*, 4, 367-379.

Chen, M. K. (2008). Rationalization and cognitive dissonance: Do choices affect or reflect preferences? (Cowles Foundation Discussion Paper No. 1669). Retrieved from *http://cowles.econ.yale.edu/P/cd/d16b/d1669.pdf*

Chen, M. K, & Risen, J. L. (2010). How Choice Affects and Reflects Preferences: Revisiting the Free-Choice Paradigm. *Journal of Personality and Social Psychology, 99*, 4, 573-594.

Chow, H. M., Mar, R. A., Yisheng, X., Siyuan L., Suraji W., Braun, A. R. (2014). Embodied Comprehension of Stories: Interactions between Language Regions and Modality-specific Neural Systems. *Journal of Cognitive Neuroscience., 26.* 2, 279-295.

Cialdini, R. B., Trost, M. R., & Newsom, J. T. (1995). Preference for consistency: The development of a valid measure and the discovery of surprising behavioral implications. *Journal of Personality and Social Psychology, 69*, 318–328.

Clemerson, P. (2009). *Evolutionary Psychologically Predicted biases in the manifestation of Cognitive Dissonance: An Exploration*. Unpublished MA thesis, Massey University, New Zealand.

Clune, J., Mouret, J-B., & Lipson, H. (2013). The evolutionary origins of modularity. *Proceedings of the Royal Society, 280,* 1755-61. DOI:10.1098/rspb.2012.2863.

Cohen, A. R. (1962). An experiment in small rewards for discrepant compliance and attitude change. In J. W. Brem and A. R. Cohen (Eds.) *Exploration in cognitive dissonance.* New York: Wiley.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20,* 37-46.

Collins, A, M., & Loftus, E. F. (1975). A Spreading-Activation Theory of Semantic Processing. *Psychological Review, 82,* 6, 407-28.

Coop, G., Bullaughey, K., Luca, F., Przeworski, M. (2008). The Timing of Selection at the Human *FOXP2. Gene Molecular Biology and Evolution. 25*, 7, 1257-1259.

Coppin, G., Delplanque, S., Cayeux, I., Porcherot, C., & Sander, D. (2010). I'm No Longer Torn After Choice: How Explicit Choices Implicitly Shape Preferences of Odors. *Psychological Science, 21,* 4, 489-493.

Corr, P. J. (2011). Anxiety: Splitting the phenomenological atom. *Personality and Individual Differences 50,* 7, 889-897. DOI: 10.1016/j.paid.2010.09.013.

Cosmides, L., & Tooby, J. (1997). *Evolutionary psychology: A primer.* Retrieved from http://www.psych.ucsb.edu/research/cep/primer.htm.

Cosmides, L., & Tooby, J. (1992). Cognitive Adaptation for Social Exchange. In Barkow, J. H., Cosmides, L., & Tooby.J. (1992). *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. Oxford, UK: OUP.

Cotton, J. L. (1985). Cognitive dissonance in selective exposure. In D. Zillman, & J. Bryant (Eds.), *Selective exposure to communication* (pp. 11-33). Hillsdale, NJ: Lawrence Erlbaum Associates.

Crockford, C., & Boesch, C. (2005). Call combinations in wild chimpanzees. *Behaviour 142*, 397-421.

Croyle, R. T., & Cooper, J. (1983). Dissonance arousal: physiologic evidence. *Journal of Personality and Social Psychology, 45*, 4, 782-91.

Dallal, G. E. (2012). *Nonparametric Statistics*. Retrieved from http://www.jerrydallal.com/lhsp/npar.htm.

Daly, M., & Wilson, M (1988). *Homicide*. New York, Aldine.

Dawkins, R. (1976). *The Selfish Gene*. Oxford: OUP.

De Boer, B. (2012). Loss of air sacs improved hominin speech abilities. *Journal of Human Evolution*, *62*, 1, 1-6 DOI: 10.1016/j.jhevol.2011.07.007

De Marneffe, M-C., Rafferty, A. N., & Manning, C. D. (2008). Finding Contradiction in Text. *Proceedings of Association for Computational Linguistics, Human Language Technologies*, (ACL-08: HLT), 1039–1047.

De Vries, J., Byrne, M., & Kehoe, E. (2015). Cognitive Dissonance induction in everyday life: An fMRI study. *Social Neuroscience, 10,* 3, 268-281.

Devine, P. G., Tauer, J. M., Barron, K. E., Elliott, A. J., & Vance, K. M. (1999). Moving beyond attitude change in a study of dissonance related processes. In E. Harmon-Jones, & Judson Mills (Eds.) *Cognitive dissonance: progress on a pivotal theory in social psychology*. Washington; American Psychological Association.

Devinsky, O., Morrell, M. J., & Vogt, B. A. (1995). Contributions of anterior cingulate cortex to behaviour. *Brain, 118 (pt 1),* 279-306.

Dobzhansky, T. (1973), "Nothing in Biology Makes Sense Except in the Light of Evolution". *American Biology Teacher 35*, 3, 125–129.

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence 35*, 1, 13 – 21.

Dutton, D. (2009). *The Art Instinct.* New York: Bloomsbury Press.

Durkheim, E. (1895/1962). *The rules of the sociological method*. Glencoe, Il: Free Press.

Egan, L. C., Bloom, P. & Santos, L. R. (2010). Choice-induced preferences in the absence of choice: Evidence from a blind two choice paradigm with young children and capuchin monkeys. *Journal of Experimental Social Psychology. 46*, 204-7.

Egan, L. C., Santos, L. R., & Bloom, P. (2007). The origins of cognitive dissonance: Evidence from children and monkeys. *Psychological Science, 18*, 11, 978 – 83.

Egidi, G., & Gerrig, R. (2006). Readers' Experiences of Characters' Goals and Actions. *Journal of Experimental Psychology, Learning, Memory, and Cognition, 32*, 6, 1322-1329.

Erlich, D., Guttman, I., Schonbach, P. & Mills, J. (1957). Post-decision exposure to relevant information. *Journal of Abnormal and Social Psychology*, *57*, 98-102.

Elkin, R. A., & Leippe, M. R. (1986). Physiological arousal, dissonance, and attitude change: evidence for a dissonance-arousal link and a "don't remind me" effect. *Journal of Personality and Social Psychology*, *51*, 1, 55-65.

Elliot, A. J., & Devine, P. G. (1994). On the motivational nature of cognitive dissonance: Dissonance as psychological discomfort. *Journal of Personality and Social Psychology, 67*, 382-394.

Ekman, P. (1993). Facial Expression and Emotion. *American Psychologist*, 48,4, 384-392.

Ekman, P. (1992). An Argument for Basic Emotions. *Cognition and Emotion, 6*, 3-4, 169-200.

Enard, W. (2011). FOXP2 and the role of the cortico-basal ganglia circuits in speech and language evolution. *Current Opinion in Neurobiology, 21*, 415-424.

Enard, W., Przeworski, M., Fisher, S. E., Lai, C. S. L., Wiebe, V., Kitano, T., Monaco, A.P., & Paabo, S. (2002). Molecular evolution of the FOXP2, a gene involved in speech and language. *Nature, 418*, 869-72.

Ericsson, K. A., & Kintsch, W. (1995). Long Term Working Memory. *Psychological Review, 102*, 211-42.

Erlich, D., Guttman, I., Schonbach, P., & Mills, J. (1957). Post decision exposure to relevant information. *Journal of Abnormal and Social Psychology, 54*, 98-102.

Ermer, E., Guerin, S.A., Cosmides, L., Tooby, J., and Miller, M. B. (2006). Theory of mind broad and narrow: Reasoning about social exchange engages ToM areas, precautionary reasoning does not. *Social Neuroscience*. *1*, (3-4), 196-219.

Faivre, J.-P (2012)**.** A Material Anecdote but Technical Reality. *Lithic Technology*, *37*, 1, 5-24.

Faivre, J.-P., Discamps, E., Gravina, B., Turq, A., Guadelli, J-L., Lenoir, M. (2014). The contribution of lithic production systems to the interpretation of Mousterian industrial variability in south-western France: The example of Combe-Grenal (Dordogne, France), *Quaternary International (2014)*, Retrieved from http://dx.doi.org/10.1016/j.quaint.2014.05.048.

Fedorenko, E. (2014). The role of domain-general cognitive control in language comprehension, *Frontiers in Psychology, 5*, article 335.

Fedorenko, E., Behr, M. K., & Kanwisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *PNAS, 108*, 39, 16428–16433.

Fedorenko, E., & Thompson-Schill, E. (2014). Reworking the language network *Trends in Cognitive Sciences, 18*, 3, 120-6.

Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Chicago, New York: Row, Peterson and Company.

Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology. 58*, 203-10.

Festinger. L., Rieken, H. W., & Schachter, S. (1956). *When Prophecy Fails: A Social and Psychological Study of a Modern Group that Predicted the Destruction of the World.* NY: Harper Torchbooks.

Fillmore, C. J. (1968). The Case for Case. In E. Bach and R.T. Harms (Eds.) *Universal in Linguistics*. NY: Holt, Rinehart and Winston.

Finger, S. (1994). *Origins of Neuroscience: A History of Explorations into Brain Function*. Oxford. OUP.

Fisher, H. E., Aron, A., Mashek, D., Li, H., & Brown, L. L. (2002). Defining the brain systems of lust, romantic attraction and attachment. *Archives of Sexual Behavior*. *31*, 5, 413-19.

Fleiss, J. L. (1971). Measuring Nominal Scale Agreement among many raters. *Psychological Bulletin, 76*, 5, 378-382.

Fodor, J. A., (1983). *The modularity of mind*. Massachusetts; Bradford Books/MIT Press.

Fried, C. B., & Aronson, E. (1995). Hypocrisy, Misattribution, and Dissonance Reduction. *Personality and Social Psychology Bulletin*, *21*, 9, 925-933.

Friederici, A. D. (2012). The cortical language circuit: from auditory perception to sentence comprehension. *Trends in Cognitive Sciences 16*, 5, 262-268.

Fritsche, I., Jonas, E., Ablasser, C., Beyer, M., Kuban, J., Manger, A. M., & Shultz, M. (2013). The power of we: Evidence for group-based control. *Journal of Experimental Social Psychology, 49*, 1, 19–32. http://dx.doi.org/10.1016/j.jesp.2012.07.014.

Gathercole, S. E. (2007). Working Memory and Language. In Gaskel, M.G. (Ed.) *The Oxford Handbook of Psycholinguistics*. Oxford, OUP.

Gawronski, B. (2012). Back to the furute of dissonance theory: Cognitive consistency as a core motive. *Social Cognition, 30*, 6. 652 – 668.

Geary, R. C. (1947). Testing for Normality. *Biometrika, 34*, 3/4, 209-242.

Geertz, C. (1973). *The interpretation of cultures*. NY: New York: Basic Books.

Gehring, W. J., & Fencsik, D. (1999*). Slamming on the brakes: An electrophysiological study of error response inhibition*. Poster presented at the Annual Meeting of Cognitive Neuroscience Society, Washington, DC.

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. & Donchin, E. (1993) A neural system for error detection and compensation. *Psychological Science 4*, 6, 385-390.

Gemba. H., Sasaki, K., & Brooks. V.B. (1986). “Error" potentials in limbic cortex (anterior cingulate area 24) of monkeys during motor learning. *Neuroscience Letters. 70*, 223-227.

Ghazanfar, A. A. (2008). Language evolution: Neural differences that make a difference. *Nature Neuroscience 11*, 382 – 384.

Glass, G. V., Peckham, P. D., & Sanders, J. R. (1972). Consequences of Failure to Meet Assumptions Underlying the Fixed Effects Analyses of Variance and Covariance. *Review of Educational Research, 42*, 3, 237-288.

Gintis, H., Henrich, J., Bowles, S., Boyd, R, & Fehr, E. (2008). Strong Reciprocity and the Roots of Human Morality. *Social Justice Research, 21*, 2, 241-253.

Gonzaga, G. C., Turner, R. A., Keltner, D., Campos, B., & Altemus, M. (2006). Romantic love and sexual desire in close relationships. *Emotion*, 6, 2, 163-179. http://dx.doi.org/10.1037/1528-3542.6.2.163

Gómez-Robles, A., de Castro, J. M. B., Arsuaga, J-L., Carbonell, E. & Polly, P. D. (2013). No known hominim species matches the expected dental morphology of the last common ancestor of Neanderthals and modern humans. *PNAS 110* (45) 18196-18201. DOI:10.1073/pnas.1302653110

Gonzaga, G. C., Keltner, D., Turner, R. A., Campos, B., & Altemus, M. (2006). Romantic Love and Sexual Desire in Close Relationships. *Emotion, 6,* 2, 163-79.

Gosling, P., Denizeau, M., & Oberlé, D. (2006). Denial of responsibility: A new mode of dissonance reduction. *Journal of Personality and Social Psychology, 90*, 5, 722-733. http://dx.doi.org/10.1037/0022-3514.90.5.722

Gould, S. J., & Vrba, E. S. (1982). Exaptation – A Missing Term in the Science of Form. *Paleobiology*, 8, 1, 4-15.

Graesser, S. E. & McMahen, C. L. (1993). Anomalous Information Triggers Questions When Adults Solve Quantitative Problems and Comprehend Stories. *Journal of Educational Psychology*. *85*, 1, 136-51.

Graesser, A. C., Singer, M. & Trabasso, T. (1994). Constructing Inferences During Narrative Text Comprehension. *Psychological Review*, *101*, 3, 371-393.

Grange, P., Bohland, J. W., Okaty, B., Sugino, K., Hemant, B., Nelson, S., Ng, L., Hawrylycz, M., & Partha, P. M. (2014). Cell-type-specific transcriptomes and the Allen Atlas (II): discussion of the linear model of brain-wide densities of cell types. Cornell University Library *arXiv:1402.2820v3*

Green, R. E., Krause, J., Briggs, A. W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Heng, L., Weiwei Z., Fritz, M., Hsi-Yang., Hansen, N. F., Durand, E. Y., Malaspinas, A-S., Jensen, J. D., Marques-Bonet, T., Alkan, C., Prüfer, K., Meyer, M., Burbano, H. A., & Good, J. M. (2010). A draft sequence of the Neanderthal Genome. *Science 328*, 5079, 710 -22.

Greenberg, J., Solomon, S., & Pyszczynski, T. (1997). Terror management theory of self-esteem and cultural worldviews: Empirical assessments and conceptual refinements. In M. P. Zanna, (Ed.) *Advances in experimental social psychology, 29,* 61–139.

Guadagno, R. E., & Cialdini, R. B. (2010). Preference for consistency and social influence: A review of current research findings. *Social Influence, 5*, 3, 152-163.

Hardych, J. A. & Braden, M. (1962). Prophecy Fails Again: A Report of a Failure to Replicate. *Journal of Abnormal and Social Psychology*, *65*, 2, 136-141.

Harmon-Jones, E. (2004). Contributions from research on anger and cognitive dissonance to understanding the motivational functions of asymmetrical frontal brain activity *Biological Psychology*, *67*, 51–76.

Harmon-Jones, E. (2000). A cognitive dissonance theory perspective on the role of emotion in the maintenance and change of beliefs and attitudes. In N. H. Frijda, A. R. Manstead & S. Bem, (Eds), *Emotions and Beliefs*. Paris: The Press Syndicate of the Cambridge.

Harmon-Jones, E., Amodio, D. M., & Harmon-Jones, C. (2009), Action-based model of Dissonance: A Review, Integration and Expansion of Conceptions of Cognitive Conflict. *Advances in Experimental and Social Psychology*, *41*, 119-165.

Harmon-Jones, E., Brehm. J. W., Greenberg, J., Simon, L., & Nelson, D. E. (1996). Evidence that the production of aversive consequence is not necessary to create Cognitive Dissonance. *Journal of Personality and Social Psychology 70*, 1, 5-16.

Harmon-Jones, E., & Harmon Jones, C. (2012). Feeling Better or Doing Better? On the Functions of Inconsistency Reduction (and Other Matters). *Psychological Inquiry, 23,* 350–353.

Harmon-Jones, E., & Harmon-Jones, C. (2008). Action-based Model of Dissonance: A Review of Behavioral, Anterior Cingulate and Prefrontal Cortical Mechanisms. *Social and Personality Psychological Compass. 2/3*, 1518-1538.

Harmon-Jones, E., & Harmon-Jones, C (2002). Testing the Action-Based Model of Cognitive Dissonance: The Effect of Action Orientation on Postdecisional Attitudes. *Personality and Social Psychology Bulletin, 28*, 6, 711-23.

Harmon-Jones, E., Harmon-Jones, C., & Levy, N. (2015). An Action-Based Model of Cognitive Dissonance Processes. *Current Directions in Psychological Science. 24*, 3, 184-9.

Harmon-Jones, E., Harmon-Jones, C., Serra, R., & Gable. P. A. (2011). The Effect of Commitment on Relative Left Frontal Cortical Activity: Tests of the Action-Based Model of Dissonance. *Personality and Social Psychology Bulletin, 37*, 3, 395–408.

Harmon-Jones, E., Lueck, L., Fearn, M., & Harmon-Jones, C. (2006). The effect of personal relevance and approach-related action expectation on relative left frontal cortical activity. *Psychological Science, 17*, 434-440.

Harmon-Jones. E., & Mills, J. (1999). An Introduction to Cognitive Dissonance Theory and an Overview of Current Perspectives in the Theory. In Harmon-Jones, E., & Mills, J. (Eds). *Cognitive Dissonance: progress on a pivotal theory on social psychology*. Washington: American Psychological Association.

Harmon-Jones, C., Schmeichel, B., Inzlicht, M., & Harmon-Jones, E. (2011). Trait approach motivation related to dissonance reduction. *Social Psychological and Personality Science, 2*, 21-28.

Hawkins, D. M. (1980). *Identification of Outliers*. London: Chapman and Hall.

Haselton, M. G., Buss, D. M., Oubaid, V., & Angleitner, A. (2005). Sex, lies, and strategic interference: the psychology of deception between the sexes. Personality and Social Psychology Bulletin, 31(1), 3-23. doi: 10.1177/0146167204271303

Henrich, J., Heine, S. J., Norenzayan, A. (2010). The weirdest people in the world? *Behavioral Brain Sciences, 33,* 61–135. DOI: 10.1017/s0140525x0999152x.

Henshilwood, C. S., D'Errico, F., & Watts, I. (2009). Engraved ochres from the Middle Stone Age levels at Blombos Cave, South Africa. *Journal of Human Evolution, 57,* 27-47.

Hess, D. J., Foss, D. J., & Carroll, P. (1995). Effects of Global and Local Context on Lexical Processing During Language Comprehension. *Journal of Experimental Psychology, 124,* 1, 62-82.

Hirschberger, G., Florian, V., Mikulincer, M., Goldenberg, J. L., & Pyszczynski, T. (2002). Gender differences in willingness to engage is risky behaviour: A Terror Management perspective. *Death Studies, 26,* 117-141.

Holbrook, C., Sousa, P., & Hahn-Holbrook, J. (2011). Unconscious Vigilance: Worldview Defence Without Adaptations for Terror, Coalition, or Uncertainty Management. *Journal of Personality and Social Psychology, 101,* 3, 451–466.

Horshak, O. V., Giger, J-C., Cabral, M., & Pochwatko, G. (2014). From demonstration to theory in embodied language comprehension: A review. *Cognitive System Research 29-30,* 66-85.

Hoshino-Browne, E., Zanna, A. S., Spencer, S. J., Zanna, M. P., Kitayama, S., & Lackenbauer, S. (2005). On the cultural guises of cognitive dissonance: The case of Easterners and Westerners. *Journal of Personality and Social Psychology, 89,* 294–310.

Hoyt, C. S. (2003). Visual function in the brain-damaged child. *Eye, 17*, 371-386. DOI: 10.1038/sj.eye.6700364.

Hrdy, S. B. (2009). *Mothers and Others*: *The Evolutionary Origins of Mutual Understanding.* Cambridge, Massachusetts: Harvard University Press.

Huitema, J. S., Dopkins, S., Klin, C. M., & Myers, J. L. (1993). Connecting goals and actions during reading. *Journal of Experimental Psychology, Learning, Memory, and Cognition, 19,* 5, 1053-1060.

Izuma, K., Matsomoto, M., Murayama, K., samejima, K., sadato, N., & Matsumoto, K. (2010). Neural Correlated of cognitive dissonance and choice induced preference choice. *PNAS, 107*, 51, 22014 - 22019

Jackendoff, R. (1999). Possible stages in the evolution of the language capacity. *Trends in Cognitive Sciences 3*, 7, 272-279.

Johanson, P., Hall, L., Tarning, B., Sikström, S. & Chater, N. (2010). Choice Blindness and Preference Choice. *Journal of Behavioral Decision Making*, 27, 3, 281 – 289.

Jonas, E., McGregor, I., Klackl, J., Agroskin, D., Fritsche, I., Holbrook, C. Nash, K., Proulx, T., & Quirin, M. (2014). Threat and Defense: From Anxiety to Approach. In J. M. Olson & M. P. Zanna (Eds.) *Advances in Experimental Social Psychology. 49*, 219-86.

Jones, E. G. (1999). Golgi, Cajal and the Neuron Doctrine. *Journal of The History of the Neurosciences*. *8*, 2, 170-8.

Just, M. A., & Carpenter, P. A. (1992). A Capacity Theory of Comprehension: Individual Differences in Working Memory. *Psychological Review, 99,* 1, 122-49.

Just, M.A., Cherkassky, V.L., Aryal, S., & Mitchell, T.M. (2010). A Neurosemantic Theory of Concrete Noun Representation Based on the Underlying Brain Codes. *PLOSone, January 13, 2010,* DOI: 10.1371/journal.pone.0008622

Kanwisher, N. (2010). Functional specificity in the human brain: A window into the functional architecture of the mind. *Proceedings of the National Academy of Sciences, 107,* 25, 11163-11170.

Karkanas, P., Shahack-Gross, R., Ayalon, A., Mar-Matthews, M., Barkai, R., Frumkin, A., Gopher, W., & Stiner, M. C. (2007). Evidence for habitual use of fire at the end of the lower palaeolithic: Site-formation processes at Qesem Cave, Israel. *Journal of Human Evolution, 53,* 197-212.

Kassim, S., Leo, R. A., Meissner, C. A., Richman, K. D., Colwell. L. H., Leach, A-M., & Fon, A. L. (2007). Police Interviewing and Interrogation: A Self-Report Survey of Police Practices and Beliefs. *Law and Human Behavior, 31,* 381-400.

Kay, R. F., Cartmill, M. & Balow, M. (1998). The hypoglossal canal and the origin of human vocal behavior. *Proceedings of the National Academy of Sciences,* 95, 5417-5419.

Kay, A. C., Moscovitch, D., & Laurin, K. (2010). Randomness, attributions of arousal, and belief in God. *Psychological Science, 21,* 216–218.

Kay, A. C., Whitson, J. A., Gaucher, D., & Galinsky, A. D. (2009). Compensatory control: Achieving order through the mind, our institutions, and the heavens. *Current Directions in Psychological Science, 18,* 5, 264–268.

Keck, D., & Kammerer, Y. & Starauschek. (2015). Reading science texts online: Does source information influence the identification of contradictions within texts? *Computers & Education, 82,* 442-449.

Keltner, D. & Buswell, B. (1996). Evidence for the Distinctness of Embarrassment, Shame and Guilt: A Study of Recalled Antecedents and Facial Expression of Emotion. *Cognition and Emotion*. 10, 2, 155-71.

Kiehl, K. A., Liddle, P. F., & Hopfinger, J. B. (2000). Error processing and the rostral anterior cingulate: An event-related fMRI study. *Psychophysiology, 37,* 216–223.

Kintsch, W. (1998) *Comprehension: A paradigm for cognition*. Cambridge, UK. Cambridge University Press.

Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85,* 363-394.

Kitayama, S., Chua, H. F., Tompson, S., & Han, S. (2013). Neural mechanisms of dissonance: An fMRI investigation of choice justification. *NeuroImage 69,* 206–212.

Klatzky, R. L., Pellegrino, J. W., McCloskey, D. P., & Doherty, S. (1989). Can you squeeze a tomato? The role of motor representations in semantic sensibility judgements. *Journal of Memory and Language*, *28,* 56–77.

Klein, R. G. (2002). *The dawn of human culture*. New York; John Wiley and Sons.

Korgaonkar, M. S., Fornito, A., Williams, L. M., and Grieve, S. M. (2014). Abnormal Structural Networks Characterize Major Depressive Disorder: A Connectome Analysis. *Biological Psychiatry, 76,* 7, 567-574.

Krause, J., Lalueza-Fox, C., Orlando, K., Enard, W., Green, R. E., Burbano, H. A., Hublin, J-J. Hanni, C., Fortea, J., de la Rasilla, M., Bertranpetit, J., Rosas, A., & Paabo, S. (2007). The derived FOXP2 variant of modern humans was shared with neanderthals. *Current Biology, 17,* 1908-1912.

Krill, A. L., Platek, S. M., Goetz, A. T., & Shackelford, T. K. (2007). Where Evolutionary Psychology meets Cognitive Neuroscience: A précis to Evolutionary Cognitive Neuroscience. *Evolutionary Psychology, 5,* 1 232-256.

Kwon, K-A., Shipley, R. J., Edirisinghe, M., Ezra, D. G., Rose, G., Serena M. Best, S. M., & Cameron, R. E. (2013). High-speed camera characterization of voluntary eye blinking kinematics**.** *Interface*. 12the June. London: The Royal Society.

Lai, C. S. L., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F., & Monaco, A. P. (2001). A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*. *413*, 519.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 1, 159-174.

Lassonde, K. A., Surla, C., Buchanan, J. A., & O'Brien, E. J. (2012). Using the contradiction paradigm to assess ageism. *Journal of Aging Studies*, *26*, 2, 174-181.

Lazarus, R.S. (1991). *Emotion and Adaptation*. Oxford. OUP

LeDoux, J. E. (2000). Emotion Circuits in the Brain. *Annual Review of Neuroscience*. 23, 155–184.

Lee, S-H, & Wolpoff, M. H. (2003). The pattern of evolution in Pleistocene human brain size. *Paleobiology*, *29*, 2, 186-196.

Lieberman, P., Laitman, J. T., Reidenberg, J.S., & Ganon, P. J., (1992). The anatomy, physiology, acoustics and perception of speech: Essential elements in analysis of the evolution of human speech. *Journal of Human Evolution*, *23*, 447-67.

Leippe, M. R., & Eisenstadt, D. (1999). A self-accountability model of dissonance reduction. In Harmon-Jones, E and Mills, J. (Eds.) *Cognitive Dissonance: Progress on a pivotal theory in social psychology*. Washington: APA.

Livingstone, S. R., & Palmer, C. (2015, October 26). Head Movements Encode Emotions During Speech and Song. *Emotion*. Advance online publication. http://dx.doi.org/10.1037/emo0000106

Lix, L. M., Keselman, J. C., Keselman, H. J. (1996). Consequences of Assumption Violations Revisited: A Quantitative Review of Alternatives to the One-Way Analysis

of Variance F Test. *Review of Educational research, 66,* 4, 579-619.

Losch, M.E., & Cacioppo, J. T. (1980). Cognitive dissonance may enhance sympathetic tonus, but attitudes are changed to reduce negative affect rather than arousal. *Journal of Experimental Social Psychology, 26*, 4, 289-304.

Louwerse, M., & Jeuniaux, P. (2008). Language comprehension is both embodied and symbolic. In de Vega, M., Glenburg, A., & Graesser, A. (Eds.) *Symbols and Embodiment: Debates on meaning and cognition.* Oxford, OUP.

Lycett, S. L., von Cramon-Taubadel, N., & Eren, M. I. (2015). Levallois: Potential implications for learning and cultural transmission capacities. *Lithic Technology*. Published online. Downloaded 12/12/20 15. DOI: http://dx.doi.org/10.1179/2051618515Y.0000000012

Lydall, E. S., Gilmour, G., and Dwyer, D. M. (2010). Rats place greater value on rewards produced by high effort: An animal analogue of the "effort justification" effect. *Journal of Experimental Social Psychology. 46*, 6, 1134-1137.

McCready, W. (1996). Applying Sampling Procedures. In Frederick T. L. Leong and James T Austin (Eds.). *The Psychology Research Handbook: A Guide for Graduate Students and Research Assistants*. Thousand Oaks, California: Sage Publications Inc.

McDonald, J. H. (2014). *Handbook of Biological Statistics*. Baltimore, MD: Sparky House Publishing. Downloaded from http://www.biostathandbook.com/transformation.html.

McDougall, I., Brown, F. H., & Fleagle, J. G. (2005). Stratigraphic placement and age of modern humans from Kibish, Ethiopia. *Nature, 433,* 733-736. DOI:10.1038/nature03258

McGillicudy-De Lisi, A., & De Lisi, R. (Eds.). (2002). *Biology, Society and Behaviour. The Development of Sex Differences in Cognition.* Westport, CT: Ablex Publishing.

McGrath, A. (2011). Changing or Defending Our Behaviour: The Role of Attitude Importance and Choice in the Arousal and Reduction of Cognitive Dissonance. PhD Thesis retrieved from https://curve.carleton.ca/system/files/etd/568d1859-6934-416f-b966-6500bdb2259e/etd_pdf/3b6e08596b70d7abcb090a906e44693a/mcgrath-changingordefendingourbehaviourtheroleofattitude.pdf.

McGregor, I., Nash, K., & Inzlicht, M. (2009). Threat, high self-esteem, and reactive approach-motivation: Electroencephalograph evidence. *Journal of Experimental Social Psychology, 45,* 1003–1007.

McGregor, I., Nash, K., Mann, N., & Phills, C. (2010). Anxious Uncertainty and Reactive Approach Motivation (RAM). *Journal of Personality and Social Psychology, 99,* 1, 133–147.

McKoon, G., & Ratcliff. R. (1986). Inferences about Predictable Events. *Journal of Experimental Psychology: Learning, Memory and Cognition, 12*, 1, 82-91.

McKoon, G. & Ratcliff, R. (1990). Textual inferences: Models and measures. In D.A. Balota, G.B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension Processes in Reading*. Hillsdale, NJ: Erlbaum.

McKoon, G., & Ratcliff, R. (1992). Inference During Reading. *Psychological Review, 99*, 3, 440-66.

McKoon, G & Ratcliff, R. (1995). The Minimalist Hypothesis: Directions for Research. In Charles A. Weaver III, Suzanne Mannes, and Charles Fletcher (Eds.) Essays in Honor of Walter Kintsch. Hillsdale, NJ: Lawrence Erlbaum Associates.

Marth, G., Schuler, G., Yeh, R., Davenport, R., Agarwala, R., Church, D., Wheelan, S., Baker, J., Ward, M., Kholodov, M., Phan, L., Czabarka, E., Murvai, J., Cutler, D., Wooding, S., Rogers, A., Chakravarti, A., Harpending, H. C., Kwok, P. Y., & Sherry, S. T. (2003). Sequence variations in the public human genome data reflect a bottlenecked population history. *Proceedings of the National Academy of Sciences, 100,* 1, 376-381.

Martínez, I., Rosa, M., Arsuaga, J.-L., Jarabo, P., Quam, R., Lorenzo, C., Gracia, A., Carretero, J.-M., Bermúdez de Castro, J.-M., Carbonell, E. (2004). Auditory capacities in Middle Pleistocene humans from the Sierra de Atapuerca in Spain. *Proceedings of the National Academy of Sciences, 101,* 27, 9976-9981.

Mayo, N., Corley, M., & Keller, F. (2005). *WebExp2 experimenter's manual.* Retrieved April, 2010, from http://www.webexp.info/

Mesulam, M-M., Thompson, C. K., Weintraub, S., & Rogalski, E. J. (2015). The Wernicke conundrum and the anatomy of language comprehension in primary progressive aphasia. *Brain*. DOI: http://dx.doi.org/10.1093/brain/awv154. First published online: 25, June, 2015.

Meyer, D.E., & Schvandenveldt, R. W. (1971). Facilitation in Recognizing Pairs of Words: Evidence of a Dependence between Retrieval Operations. *Journal of Experimental Psychology*, *90,* 2, 227-234.

Micceri, T. (1989). The Unicorn, The Normal Curve and Other Improbable Creatures. *Psychological Bulletin, 105*, 1, 156-166.

Miller, K. (2011). Cognitive Interviewing. In Madans, J., Miller, K., Maitland, A., & Willis, G. (Eds.) *Question Evaluation Methods: Contributing to the Science of Data Quality*. Hoboken, NJ: Wiley.

Milner, B., Squire, L. R., & Kandel, E. R. (1998). Cognitive Neuroscience and the Study of Memory. *Neuron, 20,* 3, 445-468.

Mišic, B. & Sporns, O. (2106). From regions to connections and networks: new bridges between brain and behaviour. *Current Opinion in Neurobiology, 40,* 1-7.

Mithen, S. (1996). *The prehistory of the mind*. London: Phoenix.

Mobbs, D., Hagan, C. C., Dalgleish, T., Silston, B., & Prévost, C. (2015). The ecology of human fear: Survival optimization and the nervous system. *Frontiers in Neuroscience,* 9, 2, Article number 55. DOI: 10.3389/fnins.2015.00055

Mohammad, S., Dorr, B., Hirst, G., & Turney, P. D. (2013). Computing Lexical contrast. *Computational Linguistics, 39*, 3, 19-53.

Mordecai, B. (2012). *Mathematical logic for computer science*. London: Springer.

Morgan, T. J. H., Uomini, N. T., Rendall, L.E., Chouinard-Thuly, L. Street, S. E. Lewis, H. M., Cross, C. P., Evan, C., Kearney, R. de la Torre, I., Whiten, A., & Laland, K. N. (2015). Experimental evidence for the co-evolution of hominin tool-making teaching and language. *Nature Communications, 6*, 6026. DOI: 10.1038/ncomms7029

Moore, J. E. A. (1940). Further study of sex differences in speed of reading, *Peabody Journal of education, 17*, 6, 359-62.

Mosteller, F., & Parunak, A. (1985). Identifying extreme cells in a sizable contingency table: Probabilistic and exploratory approaches. In D. Hoaglin, F. Mosteller & John W. Tukey (eds) *Exploring data tables, trends and shapes*. New York: Wiley.

Murdoch, C. P. (1932). The Science of Culture. *The American Anthropologist. 34*, 2, 200-15.

Murphy, S.A. (2008). The loss of a child: Sudden death and extended illness perspectives. In M. S. Stroebe, R. O Hansson, H. Schut, & W. Stroebe (Eds.). *Handbook of bereavement research and practice: Advances in theory and intervention* pp. 375-395. Washington: APA.

Myers, J. L., & O'Brien, E. J. (1998). Accessing the Discourse during Reading. *Discourse Processes, 26*, 2, 131-57.

Nagler, R. H. (2014). Adverse outcomes associated with media exposure to contradictory nutrition messages. *Journal of Health Communication*, 19, 1, 24-40. DOI: 10.1080/10810730.2013.798384

Nairne, J. S., & Pandeirada, J. N. S. (2008). Adaptive memory: Remembering with a stone-age brain. *Current Direction in Psychological Science*, *17*, 4, 239-243.

Nairne, J. S. & Panaeirada, J. N. S. (2010). Adaptive memory: Ancestral priorities and the mnemonic value of survival processing. *Cognitive Psychology, 61*, 1-22.

Nairne, J. S., Thomson, S. R., & Pandeirada, J. N. S. (2007). Adaptive memory: Survival processing enhances retention. *Journal of Experimental Psychology: Learning, Memory and Cognition*. *33*, 263-273.

Nesse, R. M. (1990). Evolutionary Explanations of Emotions. *Human Nature*. *1*, 3, 261-69.

Noble, W. & Davidson, I. (1996). *Human Evolution, Language and Mind: A Psychological and Archaeological Inquiry*. Cambridge: Cambridge University Press.

Oberfeld, D., & Franke, T. (2013). Evaluating the robustness of repeated measures analyses: The case of small simple sizes and nonnormal data. *Behavior Research Methods, 45,* 792-812. http:/dx.doi.org/10.3758/s13428-012-0281-2.

O'Brien, E. J. & Albrecht, J. E. (1992). Comprehension Strategies in the Development of a Mental Model. *Journal of Experimental Psychology: Learning, Memory and Cognition*. *18*, 4, 777-84.

Open Science Collaboration. (2015). *Science, 349,* aac4716 DOI: 10.1126/science.aac4716, downloaded 29/8/2015.

Oppenheimer, D. M. (2008). The secret life of fluency. *Trends in Cognitive Sciences, 12*, 6, 237 241.

Osborne, J. (2002). Notes on the use of data transformations. *Practical Assessment, research and Evaluation.* Retrieved from http://PAREonline.net/getvn.asp?v=8&n=6

Otero, J., & Graesser, A. C. (2001). PREG: Elements of a Model of Question Asking. *Cognition and Instruction, 19,* 2, 143-175.

Otero, J., & Kintsch, W. (1992). Failures to Detect Contradictions in a Text *Psychological Science, 3,* 4, 229- 235.

Pallak, M. S., & Pittman, T. S. (1972). General motivational effects of dissonance arousal. *Journal of Personality and Social Psychology*. *21*, 349-358.

Park, C. L. (2010). Making sense of the meaning literature: an integrative review of meaning making and its effects on adjustment to stressful life events. *Psychological Bulletin*, *136*, 257–301.

Pennisi, E. (1999). Did cooked tubers spur the evolution of big brains? *Science, 283,* 5410, 26, 2004-5.

Petkov, C. I., Kayser, C., Steudel, T., Whittingstall, K., Augath, M., & Logothetis, N. K. (2008). A voice region in the monkey brain. *Nature Neuroscience 11*, 367 – 374.

Piaget, J. (1937/1954). *The construction of reality in the child*. NewYork, NY: Basic Books.

Pinker, S. (1994). *The Language Instinct*. London: Penguin Books.

Plant, R. R., & Turner, G. (2009). Millisecond precision psychological research in a world of commodity computers. New hardware, new problems? *Behavior Research Methods, 41,* 3, 598-614.

Plutchick, R, (2003). *Emotions and Life*. American Psychological Association.

Proulx, T., & Inzlicht, M. (2012). The Five "A"s of Meaning Maintenance: Finding Meaning in the Theories of Sense-Making. *Psychological Inquiry, 23,* 317–335.

Proulx, T., Inzlicht, M., & Harmon-Jones, E. (2012). Understanding all inconsistency compensation as a palliative response to violated expectations. *Trends in Cognitive Sciences, 16*, 5, 285-91.

Prüfer, K., Racimo, F., Patterson, N., Jay, F., Sankararaman, S., Sawyer, S., Heinze, A., Renaud, G., Sudmant, P. H., de Filippo, C., Li, H., Mallick, S., Dannemann, M., Fu, Q., Kircher, M., Kuhlwilm, M., Lachmann, M., Meyer, M., Ongyerth, M., Siebauer, M., Theunert, C., Tandon, A., Moorjani, P., Pickrell, J., Mullikin, J. C., Vohr, S. H., Green, R. E., Hellmann, I., Johnson, P. L., Blanche, H., Cann, H., Kitzman, J.O., Shendure, J., Eichler, E. E., Lein, E. S., Bakken, T. E., Golovanova, L. V., Doronichev, V. B., Shunkov, M. V., Derevianko, A. P., Viola, B., Slatkin, M., Reich, D., Kelso, J., Pääbo, S. (2014). The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature 505*, 7481, 43-49.

Putnam, A. L., Wahlheim, C. N., & Jacoby, L. L. (2014). Memory for flip-flopping: detection and recollection of political contradictions. *Memory and Cognition, 42*, 7, 1198-1210.

Putt, S. S., Woods, A. D., & Franciscus, R. G. (2014). The role of verbal interaction during experimental bifacial stone tool manufacture. *Lithic Technology, 39*, 2, 96-112.

Pyszczynski, T., Greenberg, J., & Solomon, S. (1999). A dual-process model of defence against conscious and unconscious death-related thought: An extension of terror management theory. *Psychological Review, 106,* 835-845.

Quillian, M. R. (1967). Word Concepts: A Theory and Simulation of some basic Semantic Capabilities. *Behavioral Science, 12*, 410-430.

Randles, D., Inzlicht, M., Proulx, T., Tullett, A., & Heine, S. J. (2015). Is dissonance reduction a special case of fluid compensation? Evidence that dissonant cognitions

cause compensatory affirmation and abstraction. *Journal of Personality and Social Psychology*, *108*, 5, 697-710. http://dx.doi.org/10.1037/a0038933

Rapp, D. N. (2008). How do readers handle incorrect information during reading? *Memory and Cognition. 36*, 3, 688-701.

Rapp, D. N., & Gerrig, R. J. (2006). Predilections for narrative outcomes: The impact of story contexts and reader preferences. *Journal of Memory & Language, 54*, 54-67.

Rapp, D. N., Hinze, S, R., Slaten D. G., & Horton, W. S. (2014) Amazing Stories: Acquiring and Avoiding Inaccurate Information from Fiction. *Discourse Processes, 51,* 1-2, 50-74.

Rapp, D. N., & Kendeou, P. (2007). Revising what readers know: Updating text representations during narrative comprehensions. *Memory and Cognition, 35*, 8, 2019-2032.

Ratcliffe, R. (1993). Methods for Dealing with Reaction Time Outliers. *Psychological Bulletin*, *114*, 3, 510-532.

Reich, D., Green, R. E., Kircher, M., Krause, J., Patterson, N., Durand, E. Y., Bence, V., Briggs, A. W., Stenzel, U., Johnson, P. L. F., Maricic, T., Good, J. M., Marques-Bonet, T., Alkan, C., Fu, Q., Mallick, S., Li, H., Meyer, M., Eichler, E. E., Stoneking, M., Richards, M., Talamo, S., Shunkov, M., Derevianko, A. P., Hublin, J-J., Kelso, J., Slatkin, M., & Pääbo, S. (2010). Genetic history of an archaic hominin group from Denisova Cave in Siberia., *Nature, 418*, 7237, 1053-60. DOI: 10.1038/nature09710

Reich, D., Patterson, N., Kircher, M., Delfin, F., Nandineni, M. R., Pugach, I., Ko, A. M., Ko, Y-C., Jinam, T. A., Phipps**,** M. E., Saitou, N., Wollstein, A., Kayser, M., Pääbo, S. & Stoneking, M. (2010). Denisova Admixture and the First Modern Human Dispersals into Southeast Asia and Oceania. *American Journal of Human Genetics. 89*, 4, 516–528. DOI: 10.1016/j.ajhg.2011.09.005.

Richmond, B.G., & Jungers, W. I. (2008). *Orrorin tugenensis* Femoral Morphology and the Evolution of Hominin Bipedalism. *Science, 319*, 5870, 1662-1665.

Rilling, J. K., Glasser, M. F., Preuss, T. M., Ma, X., Zhao, T. Hu, X., & Behrens, T. E. J. (2008). The evolution of the arcuate fasciculus revealed with comparative DTI. *Nature Neuroscience 11*, 426 – 428.

Roberts, M. J., & Russo, R. (1999). *A Student's Guide to Analysis of Variance*. London and New York: Routledge.

Rohde, T. E., & Thompson, L. A. (2007), Predicting academic achievement with cognitive ability. *Intelligence, 35*, 1, 83–92.

Rolls, E. T. (2005). *Emotions explained*. Oxford: OUP

Rose, H. & Rose, S. (Eds.) (2000). *Alas, Poor Darwin: Arguments against evolutionary psychology*. London: Jonathon Cape.

Rose, S., Spinks, N., & Canhoto, A. I. (2014). Tests for the assumption that a variable is normally distributed. *Note to Chapter 13. Management Research: Applying the Principles.* UK: Routledge. Downloaded from http://documents.routledge-interactive.s3.amazonaws.com/9780415628129/Chapter 13 - Tests for the assumption that a variable is normally distributed final_edited.pdf on 24/2/2016.

Rosner, B**.,** Glynn, R. J., & Lee, M-L. T. (2006). The Wilcoxon Signed Rank Test for Paired Comparisons of Clustered Data. *Biometrics, 62,* 185-192.

Rubin, P. H., & Somanathan, E. (1998). Humans as Factors of Production: An Evolutionary Analysis. *Managerial & Decision Economics, 19*, 7/8, 441-455.

Rugg, G., & McGeorge, P. (1997). The sorting techniques: a tutorial paper on card sorts, picture sorts and item sorts. *Expert Systems, 14*, 2. 80-93.

Santos, E. Lu, Q., Lenci, A., & Huamg, C-R. (2014). *Unsupervised Antonym-Synonym discrimination in Vector Space.* Downloaded from www.researchgate.net 10.12871/CLICIT2014163

Scher, S., & Cooper, J. (1989). Motivational basis of dissonance: The singular role of behavioral consequences. *Journal of Personality and Social Psychology, 56,* 6, 899-906.

Scherer, C. W. (1993). Fastest Wing Beat. In Thomas J. Walker (Ed.) *Book of Insect Records*. Chapter 9. Gainsville, FL: University of Florida.

Schmalhofer, F., McDaniel, M. A., & Keefe, D. (2002). A Unified Model for Predictive and Bridging Inferences. *Discourse Processes, 3,* 2, 105-32.

Schoenemann, P. T. (2012). Evolution of brain and language. In Hofman, M. A., & Falk, D. (Eds.) *Progress in Brain Research: Evolution of the Primate Brain from Neuron to Behavior, 195,* 443 – 459.

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences, 4,* 8, 299-309.

Shackman, A. J., Salomons, T.V., Slagter, H. A., Fox, A. S., Winter, J. J., & Davidson, R. J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Naturel Reviews Neuroscience, 12,* 154-167.

Shallice, T. & Cooper, R.P. (2012). The organisation of mind. *Cortex, 48,* 1366-70.

Shallice, T. & Cooper, R.P. (2011). *The organisation of mind*. Oxford, UK: OUP.

Sherman, S. J., & Gorkin, R. B. (1980). Attitude bolstering when behavior is inconsistent with central attitudes. *Journal of Experimental Social Psychology, 16,* 388-403.

Silverman, I., Choi, J., & Peters, M. (2007). The hunter-gatherer theory of sex differences in spatial abilities: Data from 40 countries. *Archives of sexual behavior*, *36*, 2, 261-268.

Simon, L., Greenberg, J., & Brehm, J. (1995). Trivialization: The forgotten mode of dissonance reduction. *Journal of Personality and Social Psychology*, *68*, 2, 247-60).

Singer, M. (2007). Inference processing in discourse comprehension. In Gaskel, M.G. (Ed.) *The Oxford Handbook of Psycholinguistics*. Oxford, OUP.

Singer, M., & Doering, J. C. (2014). Exploring Individual Differences in Language Validation. *Discourse Processes, 51*, 1-2, 167-188.

Singer, M., & Ferreira, F. (1983). Inferring Consequences in Story Comprehension. *Journal of Verbal Learning and Verbal Behaviour*. *32*, 437-448.

Singer, R., & Henderson, E. (2015). Object permanence in marine mammals using violation of expectation procedure. *Behavioural Processes, 112*, 108-113.

Snowball, D., & Bolton, F. (1999). *Spelling K – 8: Planning and Teaching*. Portland Maine: Stonehouse Publishers.

Somanathan, E., & Rubin, P. H. (2004). The evolution of honesty. *Journal of Economic Behavior and Organization, 54*, 1 – 17.

Soressi, M., McPherron, S. P., Lenoir, M., Dogandžić, T., Goldberg, P., Jacobs, Z., Maigrot, Y., Martisius, N. L., Miller, C. E., Rendu, W., Richards, M., Skinner, M. M., Steele, T. E., Talamo, S., Texier, J. P. (2013). Neanderthals made the first specialized bone tools in Europe. *PNAS, 110*, 35, 14186-90.

Sourse , D. L., Lennon, G. C., Lemon, R. P., & Leandro, S. A. (2015). The relationship between intelligence and academic achievement throughout middle school: The role of students' prior academic performance. *Learning and Individual Differences 41*, 73-78,

Sparks, J. R., & Rapp, D. N. (2010). Discourse processing - examining our everyday language processes. *Wiley Interdisciplinary Reviews- Cognitive Science, 1,* 3, 371.

Stadtler, M., Scharrer, L., Skodzik, T. & Bromme. R. (2014). Comprehending Multiple Documents on Scientific Controversies: Effects of Reading Goals and Signaling Rhetorical Relationships. *Discourse Processes 51,* 1-2, 93-116.

Starzyk, K. B., Fabrigar, L. F., Soryal, A. S., & Fanning, J. J. (2009). A Painful Reminder: The role of levels and salience of attitude importance in Cognitive Dissonance. *Personality and Social Psychology Bulletin, 35,* 1, 126 - 137.

State, J. E. (2004). The property 'instinct'. *Philosophical Transactions of the Royal Society B Biological Sciences.* 359 (1451) pp. 1763 - 1774.

Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning memory and Cognition*. *21*, 5, 1140 - 1154.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics*. Boston: Pearson Education.

Tallerman, M. (2014). No syntax saltation in language evolution. *Language Sciences, 46,* 2017-2019.

Tauer, J., Devine, P., & Elliot, A. (1998). *Attitude importance, dissonance, and resistance to attitude change reduction.* Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago.

Tharoor, I. (2015). Saudi princesses languish under house arrest. *The Dominion Post* 26/1/2015.

Thompson, P. M., Ge, T., Glahn, D. C. Jahanshad, N., Nichols, T. E. (2013). Geneics of the Connectome. *NeuroImage, 80*, 475-88.

Tooby, J., & Cosmides, L. (1992). The Psychological Foundation of Culture. In J.H Barkow, L. Cosmides & J. Tooby. (Eds.) *The Adapted Mind* (pp. 19-136). Oxford: OUP.

Toups, M. A., Kitchen, A., Light, J. E., & Reed, D. L. (2011). Origin of clothing lice indicates early clothing use by anatomically modern humans in Africa. *Molecular Biology and Evolution. 28*, 1, 29-32.

Treves, A., & Naughton-Treves, L. (1999). Risk and opportunity for human coexisting with large carnivores. *Journal of Human Evolution*, 36, 3, 275-82.

Treves, A., & Palmqvist, P. (2007). Reconstructing hominin interactions with mammalian carnivores (6.0–1.8 Ma). In Sharon Gursky-Doyen and K. A. I. Nekans (Eds.) *Primate anti-predator strategies*. New York: NY.

Trivers, R. L. (1971). The Evolution of Reciprocal Altruism. *Quarterly Review of Biology*, 46, 1, *35-57*.

Trochim, W., Donnelly, J. P., & Arora, K. (2016). *Research Methods: The Essential Knowledge Base. Cengage Learning.* Australia: Sydney.

Trope, Y., & Liberman, N. (2010). Construal level theory of psychological distance. *Psychological Review, 117,* 440–463.

Tullis, T., & Wood, L. (2004), "How Many Users Are Enough for a Card-Sorting Study?" *Proceedings UPA 2004* (Minneapolis, MN, June 7-11, 2004).

University of Colorado Cancer Center. (2015, April 20). Excessive use of dietary supplements linked to increase cancer risk. *ScienceDaily*. Retrieved from www.sciencedaily.com/releases/2015/04/150420182403.htm

Uttal, W. R. (2001). *The New Phrenology: The Limits of Localizing Cognitive Processes in the Brain.* Cambridge, MA: MIT Press.

Uttal, W. R. (2011). *Mind and Brain: A Critical Appraisal of Cognitive Neuroscience*. Cambridge, MA: MIT Press.

Uttal, W. R. (2013). *Reliability in Cognitive Neuroscience: A Meta-Meta Analysis* [e-book]. Cambridge, MA: MIT Press. Available from: eBook Collection (EBSCOhost), Ipswich, MA. Accessed July 21, 2015.

Van Den Broek, P., & Gustfson, M. (1999). Comprehension and Memory for Texts: Three Generations of Reading Research. In S.R. Goldman, A.C. Graesser, & P. Van den Broek (Eds.) *Narrative, Comprehension, Causality and Coherence: Essays in Honor of Tom Trabasso*. Mahwah, NJ. Lawrence Erlbaum Associates.

Van den Broek, P., Rapp D. N. & Kendeou, P. (2005). Integrating Memory-Based and Constructivist Process in Accounts of Reading Comprehension. *Discourse Processes, 29*, 2 & 3, 299-316.

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, E. J., Yacoub, E., Ugurbil, K. (2013). The WU-Minn Human Connectome Project: An Overview. *Neuroimage*, *80*, 62–79. DOI: 10.1016/j.neuroimage.2013.05.041

Van Essen, D (2015). The Human Connectome Project. Video of part of the "Neuroscience Virtual Conference, 18th March, 2015" downloaded on 18/7/2015. from https://www.youtube.com/watch?feature=player_embedded&v=dEk5h_n_IsA

van Harreveld, F., van der Pligt, J., & de Liver, Y. N. (2009). The agony of ambivalence and ways to resolve it: Introducing the MAID model. *Personality and Social Psychology Review, 13*, 45–61.

Van Veen, V., Kruk, M. K., Schooler. J. W., & Carter, C. S. (2009). Neural activity predicts attitude change in cognitive dissonance. *Nature Neuroscience 12*, 11, 1470-4.

Vernes, S. C., Spiteri E., Nicod, J., Groszer, M., Taylor, J. M., Davies, K. E., Geschwind, D. H., Fisher, S. E. (2007). High-throughput analysis of promoter occupancy reveals direct neural targets of FOX P2, a gene mutated in speech and language disorders. *American Journal of Human Genetics, 81*,1232-50.

Vickaryous, M. K., & Hall, B. K. (2006). Human cell type diversity, evolution, development, and classification with special reference to cells derived from the neural crest. *Biological Reviews 81*, 3, 425–455.

Villa, P., & Roebebroek, W. (2014). Neandertal Demise: An Archaeological Analysis of the Modern Human Superiority Complex. *PLOSone, April 30th 2014,* DOI: 10.1371/journal.pone.0096424.

Wadley, L. (2013). Recognizing complex cognition through innovative technology in Stone Age and Palaeolithic sites. *Cambridge Archaeological Journal*, *23*, 2, 163-83.

Wakslak, C. J. (2012). The experience of cognitive dissonance in important and trivial domains: A Construal-Level Theory approach. *Journal of Experimental Social Psychology, 48,* 1361-1364.

Wallace, B. (2010). *Getting Darwin Wrong: Why Evolutionary Psychology Won't Work.* Exeter, UK: Imprint Academic.

Washington State University. (2008). The Physical characteristics of humans. Retrieved from http://www.wsu.edu/gened/learn-modules/top_longfor/phychar/culture-humans-2two.html.

Waterman, C. K. (1969). The Facilitating and Interfering Effects of Cognitive Dissonance on Simple and Complex Paired Associated Learning Tasks. *Journal of Experimental Social Psychology, 5*, 31-42.

Wedekind, C., & Milinski, M. (2000). Cooperation through image scoring in Humans. *Science*, 288, 850-852.

West, S.G. Finch, J. F., & Curran, P. J. (1905). Structural equations models with nonnormal variables: Problems and remedies. In Hoyle, R.H. (Ed.) Structure equation modelling: Concepts, issues and Applications. Newberry Park, CA: Sage.

Whoriskey, P. (2015). Dietary cholesterol warnings set to change. Retrieve from *http://www.stuff.co.nz/life-style/well-good/teach-me/66042934/dietary-cholesterol-warnings-set-to-change*.

Wilkins, J., Schoville, B. J., Brown, K. S., & Chazan, M. (2012). Evidence for early hafted hunting technology. *Science*, *338*, 6109, 942-946.

Wrangham, R.W. (2009). *Catching Fire: how cooking made us hu*man. New York: Basic Books.

Zanna, M. P. & Cooper, J. (1974). Dissonance and the pill: An attribution approach to studying the arousal properties of dissonance. *Journal of Personality and Social Psychology, 29*, 5, 703-9.

Zanna, M. P., Higgins, E. T., & Taves, P. A. (1976). Is dissonance phenomenologically aversive? *Journal of Experimental Social Psychology*, *12*, 530-8.

Zuwerink, J. R., & Devine, P. G. (1996). Attitude importance and resistance to persuasion: It's not just the thought that counts. *Journal of Personality and Social Psychology, 70*, 5, 931-944. http://dx.doi.org/10.1037/0022-3514.70.5.931

Zwaan, R. A., & Radvinsky, G.A. (1998). Situation Models of Language Comprehension and Memory. *Psychological Bulletin, 123*, 2, 162-85.

Zwaan, R, A, Stanfield, R. A., & Yaxley, R. H. (2002). Do language comprehenders routinely represent the shapes of objects? *Psychological Science, 13*, 168–71.