Errors in moral forecasting: The importance of affect for the relationship between moral behaviours and moral forecasts

by

Rimma Teper

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Department of Psychology
University of Toronto

© Copyright by Rimma Teper 2014
Abstract

In recent years, the field of moral psychology has been heavily dominated by studies on moral judgment that have relied on hypothetical moral dilemmas. But how do individuals’ responses to such dilemmas map on to real-life moral behaviour? The aim of this dissertation was to explore this association, as well as to explore the role that affective experience plays in the relationship between peoples’ actual and forecasted moral behaviours. The results of five experiments suggest that people might act more morally than they would predict and that affective experience plays in important role in motivating moral behaviours, as well as accurate forecasts.

In Studies 1 and 2, participants acted more morally than they predicted they would in a series of moral dilemmas including a one-shot Dictator Game (Study 1), as well as in a math task where they had the chance to cheat on a math test through either commission or omission (Study 2). In Study 3, participants who actually had the chance to cheat on this same math task displayed significantly more autonomic nervous system arousal than participants forecasting their behaviour in this same dilemma. Critically, the moral underestimation effect in this study was mediated by autonomic arousal. In Study 4, I found that inducing affective salience by providing false somatic feedback indicative of arousal increased forecasting accuracy among participants for the same math task dilemma. Finally, in Study 5, I found that individual differences in emotional awareness moderate moral forecasting accuracy in this same moral dilemma, such that individuals low in emotional awareness exhibit exacerbated forecasting errors. This
research suggests that the affective arousal present during real-life moral dilemmas may not be fully engaged during moral forecasting, and that this may account for the moral forecasting errors that individuals make. Given that the majority of moral psychologists rely on hypothetical scenarios and self-report to study the nature of morality, I propose that investigating the way in which affective experience interacts with and shapes individuals’ moral forecasts is an important and meaningful pursuit.
Acknowledgments

I am infinitely grateful to my advisor, Michael Inzlicht, who has shaped the researcher I am today. I am confident that he will continue to influence my thinking, and inspire me as I move forward. This journey would have impossible without his unfailing support and encouragement. I would also like to thank Elizabeth Page-Gould who I have learned from immensely. I often marvel at her wealth of knowledge and her willingness to share it with others. I would also like to thank Jason Plaks for sharing his insights and providing support throughout my time in graduate school. My labmates are among the brightest and enthusiastic people I know. They have all made graduate school a true pleasure by inspiring me with their ideas and by creating a warm environment in which to work. Specifically, I want to thank Alexa Tullett, Jennifer Gutsell and Shona Tritt. They have become dear friends who provide me with an immense amount of positivity on a daily basis. I feel extremely lucky to have met them. I am exceptionally thankful for my family who has supported me unconditionally, and am especially grateful to my daughter, Isabel Pogossian, who has taught me the importance of patience and balance without even trying to. My friends, Julia Bossis, Simona Hiuitin, and Erin Goldvasser have consistently provided me encouragement and perspective, and I am very thankful for our friendship. I would also like to thank Spike Lee and Chen-Bo Zhong, who have selflessly provided advice and guidance to me. They have also set examples of brilliant scientists that I will always aspire to. I am grateful to Francesca Gino, Wil Cunningham, and Erika Carlson for generously offering their time and expertise in helping me complete this dissertation. Finally, I am thankful to the SPRG at the University of Toronto for allowing me to develop as a scientist in such a warm, accepting, and intellectually stimulating environment.
# Table of Contents

Chapter 1: General Introduction ............................................................................................................. 1  
  1 Objectives ........................................................................................................................................... 1  
  2 The Attitude-Behaviour Discrepancy ................................................................................................. 3  
  3 Emotions and Moral Decision-Making ............................................................................................... 5  
  4 Emotions May Drive Moral Behaviour ............................................................................................... 10  
  5 Summary .............................................................................................................................................. 14  

Chapter 2: Errors in Moral Forecasting in a Prescriptive Moral Dilemma ............................................ 15  
  1 Introduction .......................................................................................................................................... 15  
  2 Methods ............................................................................................................................................... 16  
  3 Results ................................................................................................................................................ 17  
  4 Discussion .......................................................................................................................................... 17  

Chapter 3: Errors in Moral Forecasting in Proscriptive Moral Dilemmas ............................................ 19  
  1 Introduction .......................................................................................................................................... 19  
  2 Methods ............................................................................................................................................... 20  
  3 Results ................................................................................................................................................ 21  
  4 Discussion .......................................................................................................................................... 22  

Chapter 4: Why We Make Moral Forecasting Errors: The Role of Affective Experience ............... 24  
  1 Introduction .......................................................................................................................................... 24  
  2 Method ................................................................................................................................................ 27  
  3 Results ................................................................................................................................................ 29  
  4 Discussion .......................................................................................................................................... 30  

Chapter 5: Perceptions of Arousal Influence Moral Forecasting ......................................................... 32  
  1 Introduction .......................................................................................................................................... 32  
  2 Method ................................................................................................................................................ 33  
  3 Results ................................................................................................................................................ 34  
  4 Discussion .......................................................................................................................................... 34  

Chapter 6: Individual Differences in Moral Forecasting Accuracy: The Role of Emotional Awareness ................................................................................................................................................... 36  

<table>
<thead>
<tr>
<th></th>
<th>Chapter 1: Introduction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Chapter 2: Methods</td>
</tr>
<tr>
<td></td>
<td>Chapter 3: Results</td>
</tr>
<tr>
<td></td>
<td>Chapter 4: Discussion</td>
</tr>
<tr>
<td></td>
<td>Chapter 5: General Discussion</td>
</tr>
<tr>
<td></td>
<td>Chapter 6: Summary</td>
</tr>
<tr>
<td></td>
<td>Chapter 7: Limitations</td>
</tr>
<tr>
<td></td>
<td>Chapter 8: Future Directions</td>
</tr>
<tr>
<td></td>
<td>Chapter 9: Conclusion</td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th></th>
<th>Chapter 7: General Discussion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Summary</td>
</tr>
<tr>
<td></td>
<td>Limitations</td>
</tr>
<tr>
<td></td>
<td>Future Directions</td>
</tr>
<tr>
<td></td>
<td>Conclusion</td>
</tr>
</tbody>
</table>
List of Tables

Table 1. Bivariate correlations for main study variables. ** denotes $p < .01$, * denotes $p < .05$

Table 2. Indirect effects of action/forecasting on cheating through psychophysiological arousal.

* denotes $p < .05$ (significant indirect effect).
List of Figures

Figure 1. Number of coins left/number of coins forecasted leaving for confederate as a function of condition.

Figure 2. Number of times cheated/forecasted cheating as a function of condition.

Figure 3. Number of times cheated/forecasting cheating as a function of condition.

Figure 4. Change in PEP (ms) from baseline as a function of condition.

Figure 5. Change in RSA from baseline as a function of condition.

Figure 6. Change in number of NSSCRs from baseline as a function of condition.

Figure 7. A multiple mediation model of action/forecasting and cheating through psychophysiological arousal. Unstandardized regression coefficients from a bootstrap procedure are provided along the paths along with their associated standard errors.

Figure 8. Number of times cheated/forecasted cheating as a function of condition.

Figure 9. The number of times cheated/forecasted cheating (represented with log values) as a function of condition, as moderated by trait alexithymia.
Chapter 1: General Introduction

1 Objectives

Within the last decade, social psychologists have started to elucidate the processes underlying moral decision-making. This trend is not surprising, given the potential applicability and consequential nature of moral psychology research. Indeed, moral transgressions can cause detrimental social and economic outcomes. In 2007, the U.S. Internal Revenue Service estimated that tax evasion cost the American economy $345 billion, or about 14% of federal revenues for the fiscal year (Kaufman, 2007). At a local level, schools often suffer from academic misconduct such as cheating and plagiarism (Stern & Havlicek, 1986), while large corporations often witness employee theft (Hollinger & Clark, 1983) and sexual abuse in the workplace (Richman et al., 1999). Moral stability is vital to a well-functioning society. As such, psychologists have begun to explore the mechanics behind moral decision-making; recently focusing on the integral role the emotions may play in influencing our moral judgments and decisions in hypothetical moral dilemmas. Although these studies have undoubtedly helped us learn a great deal about the basic processes involved in moral reasoning, it is not entirely clear how individuals’ responses to hypothetical moral dilemmas map onto real-life moral decision-making. In other words, do individuals’ predicted moral behaviours map on to their actual moral behaviours? Overwhelmingly, the answer seems to be “no”. Indeed, both classic research on attitudes, as well as more contemporary research suggest that individuals’ forecasts, or predictions about what they might do, may not accurately reflect their actual behaviours (Blasi, 1980; Festinger, 1957; LaPiere, 1934). But why is this the case? And more importantly, what can we do to increase the accuracy of individuals’ moral forecasts?

From the outset, I would like to define what I mean by “moral behaviour”. Unlike what has been the tradition in much of the research on moral judgments, the goal here is not to disentangle the processes underlying competing moral principles (e.g., deontology, utilitarianism, etc.). For the sake of this dissertation, I will define “moral behaviour” as a set of rules that is collectively agreed upon by society. These rules can be either proscriptive or prescriptive in nature (Janoff-Bulman, Sheikh, & Hepp, 2009), but most members of society should agree that breaking them would go against normative expectations (i.e. cheating, lying, denying help to someone in need). Since these are the types of moral situations that people are most likely to encounter in everyday life, this “you know it when you see it” operationalization may be most appropriate when
discussing the ways in which affective processes drive us toward moral or prosocial action. “Moral forecasting”, on the other hand, will be defined as the prediction of one’s behaviour in a hypothetical moral dilemma. Because a large majority of moral psychologists rely on hypothetical scenarios to study the processes involved in moral decision-making (e.g. Greene, Nystrom, Engell, Darley, & Cohen, 2004; Rozin, Lowery, Imada, & Haidt, 1999; Haidt, 2001), I propose that clarifying the relationship between peoples’ moral forecasts and moral behaviours, as well as discovering methods by which we can increase the accuracy of these measures is an important and worthwhile pursuit.

The overarching aim of this dissertation is to investigate the relationship between individuals’ moral forecasts and their moral behaviours, and to answer the following questions: Might people be more moral than they think across a variety of different moral dilemmas? If people are in fact more moral than they predict, what factors might account for this dissociation between actions and forecasts? Can the differences in affective experience help to explain the nature of the dissociation between actions and forecasts? If engaging in moral forecasting is less affectively arousing than being an in actual moral dilemma, might experimentally manipulating affective experience increase moral forecasting? Finally, do individual differences in emotional awareness influence moral forecasting? The hypotheses for the current dissertation are thus as follows:

Hypothesis 1. Individuals faced with an actual moral dilemma will behave more morally than individuals presented with the same hypothetical moral dilemma will forecast behaving.

Hypothesis 2. Individuals faced with an actual moral dilemma will experience more affective arousal, as measured by autonomic nervous system activation, than will participants responding to a hypothetical moral dilemma.

Hypothesis 3. This hypothesized difference in affective experience between actual and hypothetical moral dilemmas should explain the discrepancy between individuals’ actions and forecasts.

Hypothesis 4. Experimentally increasing individuals’ perceptions of their arousal should result in moral forecasts that more closely match moral behaviours.
Hypothesis 5. Dispositional emotional awareness should be influence moral forecasting, such that individuals higher on trait emotional awareness should produce moral forecasts that more closely match moral behaviours.

Through a series of five studies, I will explore the dissociation between moral action and moral forecasting. Furthermore, I will investigate the role that affective experience plays in this relationship by manipulating emotional salience and by exploring the role of emotional awareness. The first hypothesis will be tested in Studies 1 and 2 where I will test the relationship between individuals’ moral behaviours and moral forecasts using a one-shot Dictator Game (Study 1), and in a math task dilemma on which participants have the opportunity to cheat through either commission (by committing an explicit action) or omission (by passively allowing the transgression to occur; Study 2). The second and third hypotheses will be tested in Study 3, where I will explore the affective factors that may be responsible for this dissociation (using a similar math task), by measuring autonomic nervous system arousal while participants complete the test with the chance to cheat, and also while they forecast their behaviour in this moral dilemma. The fourth hypothesis will be tested in Study 4, where I will explore the influence that affective salience has on moral forecasting by providing participants with false somatic feedback, which signals a state of arousal. Finally, the fifth hypothesis will be tested in Study 5, where I will investigate how trait alexithymia (i.e. a sub-clinical condition related to low levels of emotional awareness) moderates individuals’ moral forecasts. In sum, I propose that people reason about hypothetical moral scenarios in a more purely logical way than they respond to moral scenarios in the moment. I hypothesize that the affective intensity of choosing a morally relevant behaviour is underestimated during moral forecasts. Consequently, the discrepancy between moral forecasting and moral behaviour should be reduced to the degree that affect is considered during the moral forecasting process.

2 The Attitude-Behaviour Discrepancy

The work that has been conducted thus far on moral judgments is of great importance. Understanding the way in which individuals arrive at moral judgments sheds light on the broader scope of individuals’ beliefs about what is right and wrong; an action or decision can only be conceptualized as moral/immoral if a moral judgment about whether this action is morally right or wrong precedes it. However, decades of research imply that there exists a
significant discrepancy between individuals’ beliefs and individuals’ actions (Festinger, 1957; LaPiere, 1934), suggesting that utilizing hypothetical moral dilemmas may in actuality tell us very little about real-life moral decision-making.

2.1 Attitudes Do Not Reflect Behaviours

Perhaps most notably, Festinger’s work on cognitive dissonance (1957) has shown that individuals often engage in behaviour that is incongruous with their attitudes (Ajzen & Fishbein, 1977). Although this “dissonance” creates psychological tension, there are apparently much less costly ways of relieving this tension than by adjusting either the attitude or the behaviour so that they are in line with one another. Empirical evidence for this theory, however, existed as early as the 1930s, when Richard LaPiere conducted his famous field study on discrimination. He found that although 92% of hotel personnel said they would deny accommodations to Chinese guests when asked over the phone, nearly 100% accepted Chinese guests in person (1934). The classic Good Samaritan study also illustrates the dissociation between behaviours and attitudes, revealing that actively thinking about helping does not always increase helping behaviour (Darley & Batson, 1973). This discrepancy was then investigated in a review of numerous studies, which examined the relationship between self-reported morality and actual morality (i.e. honesty). While approximately half of the studies reviewed showed a positive association between self-reported and actual morality, the other half found no significant relationship, suggesting that judgment and behaviour may not have a one-to-one relationship (Blasi, 1980). In short, the social psychology literature gives us good reason to believe that predictions about one’s behaviour might not reflect actual behaviour.

2.2 Affective Forecasting and Empathy Gaps

Research on affective forecasting has found that individuals have poor insight about their future affective states. In particular, people tend to overestimate their negative affect after certain events such as romantic breakups, being denied tenure, and moving to an undesirable location (Wilson & Gilbert, 2003; Wilson & Gilbert, 2005). In other words, individuals may forecast responding to a hypothetical situation with one emotion, but may in actuality respond with a different emotion. Thus, if emotions are important for driving actions and decisions (Loewenstein, 1996; Loewenstein, Weber, Hsee, & Welch, 2001; Zeelenberg, Nelissen,
Breugelmans, & Pieters, 2008), it is easy to see why individuals might not be able to accurately predict their behaviour. Essentially, there is little reason to believe that peoples’ attitudes or beliefs about what they might do in a moral dilemma would accurately reflect their behaviour in real life.

Related work on the hot-cold empathy gap (a bias that causes people to underestimate the role of visceral experience in decision-making) has repeatedly found that people fail to appreciate the extent to which affective experiences fuel their behaviours (Van Boven, Loewenstein, & Dunning, 2005). For instance, during states of “low-craving”, smokers tend to underestimate the extent to which their future cravings will affect their behaviour (Sayette, Loewenstein, Griffin, & Black, 2008). Such research is directly applicable to the somatic markers hypothesis, which has shown that physiological responses to risky stimuli precede the behavioral decision to avoid those stimuli, and people who do not exhibit heightened physiological arousal to risky stimuli are unsuccessful at determining risk (Damasio, 1994). According to this hypothesis, there exist primary and secondary emotional inducers. Primary inducers are stimuli that are present within the immediate environment and cause pleasurable or aversive states (e.g. experiencing a break-up). Secondary inducers are generated by recalling or imagining an emotional event (e.g. imagining breaking up). While real-life experiences typically produce activations in the amygdala, a brain region implicated in primal affective responses, secondary inducers cause activations in the ventromedial prefrontal cortex, a brain area implicated in decision-making and emotion regulation (Bechara & Damasio, 2005). Because secondary inducers fail to mirror the neurological, and therefore somatic state produced by real-life experiences (specifically, they replicate it at a fainter level), it is likely that individuals might be prone to making errors when predicting their behaviour in hypothetical moral dilemmas. It seems then, that errors in affective forecasting may reflect an inability to access the intense emotional state inherent to a real-life situation.

3 Emotions and Moral Decision-Making

3.1 What is an Emotion?

Before discussing the way in which emotions inform moral decisions, it is useful to specify precisely what “emotion” means. Though numerous definitions have been proposed throughout the decades (e.g., Izard, 2010; Lazarus, 1991; Zajonc, 1984), the current consensus is that
emotion consists of three main components: subjective experience, changes in physiological arousal, and behavioural expression (Ekman, 1992; Izard, 1977; Lang, 1988; Lazarus, 1991; Levenson, 1994; Scherer, 1984; Tomkins, 1962). Importantly, research now suggests that the physiological and behavioural expression of affective states can occur without conscious experience of either the cause of the affective state or of the affective state itself (Winkielman & Berridge, 2004; Winkielman, Berridge, & Wilbarger, 2005; Zemack-Rugar, Bettman, & Fitzsimons, 2007). This notion is important when considering just how powerful emotions might be in influencing our behaviour.

3.2 The Utility of Emotions

Emotions are important tools. For instance, it is adaptive to recoil in fear upon seeing a bear, or to cringe with disgust upon encountering rotten food. The way in which such affective experiences promote adaptive behaviour is likely a neurologically primitive mechanism (Panksepp & Biven, 2012). In other words, emotions have likely evolved to deal with evolutionary recurrent situations and may represent “best guesses” as to what one ought to do in those situations (Tooby & Cosmides, 2008). This view suggests that emotion is information (Schwarz, 2001; Schwarz & Clore, 1988), and that it evolved because it allowed us to survive our ancestral environments (e.g. by avoiding the consumption of rotten food or by avoiding bears). However, there is wide consensus that beyond its informational role, emotion also serves as a motivator (Frijda, 1986; Kroll & Egan, 2004; Zeelenberg & Pieters, 2006). In other words, emotion feeds energy and vigor into the motivational system by amplifying basic drives and needs, thus making them more likely to translate into behaviour (Tomkins, 1982).

3.3 Are Emotions Important for Morality?

Although the idea the role of emotions for moral decisions may seem apparent today, this notion has been battled quite extensively historically. Indeed, the idea that emotions are simply artifacts of our animal pasts is one that has pervaded philosophical traditions for thousands of years. This belief system is well-depicted in the various writings of Plato where he symbolizes humans as being governed by two opposing forces – their rational heads and their passionate bodies (4th century B.C./1949), and where he claims that “outrage” is the product of desire without reason (4th century B.C./1997). Later philosophers such as Rene Descartes further endorsed such sentiments and warned against trusting one’s passions or intuitions, stating that “…from time to
time I have found that the senses deceive, and it is prudent never to trust completely those who have deceived us even once” (1641/1996). David Hume attempted to “rescue” emotions in the 18th century by providing an alternative theory of rationalism (1739/1969). He believed that emotions motivate behaviour and that without them we would lack all impulse to act or even to reason. Adam Smith (1759/2002) offered a more moderate account of the role of emotions for moral behaviour, distinguishing between social, unsocial, and selfish passions. With these classifications, he argued that although some emotions can elicit “improper” behaviour, other emotions can prompt us to act prosocially. Though this view gained some popularity with philosophers from the same time period, it was ultimately Immanuel Kant who left a lasting imprint on moral psychology, espousing that emotions are unreliable and capricious, and therefore cannot be moral motives (1785/1964).

The influence of these rationalist philosophies can be seen in the theories of moral psychology pioneers such as Piaget (1932), Kohlberg (1969), and later Turiel (1983), who studied the development of moral reasoning in children through the use of hypothetical scenarios and interviews. These psychologists adopted an approach that excluded emotional processes almost entirely. They believed that moral judgments are arrived at through step-by-step reasoning that is both conscious and reflective. It is only recently that psychologists have begun to reconsider emotions as an important ingredient for moral decision-making (Greene & Haidt, 2002; Haidt, 2001; Haidt, 2003).

How Emotions Interact with Cognitions and Motivations

A plethora of research suggests that no one single psychological process drives moral decision-making on its own. Specifically, both cognitive forces (Greene, Morelli, Lowenberg, Nystrom & Cohen, 2008) and conscious motivations (Kroll & Egan, 2004) interact interchangeably and iteratively with emotional processes to shape moral decisions. Cognitions feed into our moral decision-making processes by shaping emotional responses (Pizarro & Bloom, 2003). On a broader scale, we now know that emotions influence a variety of basic cognitive operations, such as attention, memory, control, and even basic perception (Fredrickson & Branigan, 2005; Gable & Harmon-Jones, 2008; Harmon-Jones, Gable, & Price, in press; Inzlicht & Al-Khindi, 2012; Schmeichel & Inzlicht, 2013).
There also exists evidence that supports the interactive relationship between emotional experience and conscious motivational processes (e.g., desires and goals), suggesting that emotions feeds into our motivations, making them more likely to translate into actual behavior (Tomkins, 1982). These motivations can then be important in shaping the way in which we act (Mazar & Ariely, 2006), but may also shape the emotions that then influence moral decision-making “in the moment”. Although these cognitive and motivational processes are important for moral behaviors, there is reason to believe that the transient experience of emotion drives moral actions on a moment-by-moment basis.

3.4 What is a Moral Emotion?

Haidt (2003) has suggested that a moral emotion should be “linked to the interests or welfare either of society as a whole or at least of persons other than the judge or agent” (p. 276), and has argued that moral emotions have evolved as commitment devices that motivate behaviour that may be costly in the short-term but beneficial in the long-term (see also Rand & Nowak, 2013). The multifaceted nature of moral emotions has prompted theorists to create classifications and categories for them. When anticipating self-conscious emotions such as guilt and shame, for instance, people are less likely to commit transgressions. At the same time, experiencing guilt after a transgression is likely to influence future behaviors both in terms of deterring future transgressions and motivating restitutions (Tangney, Stuewig, & Mashek, 2007). On the other hand, positive moral emotions such as love, pride, and elevation (Tangney, 1991; Haidt, 2000) are theorized to be instrumental in driving ethical behaviors because they may inspire individuals to act prosocially. Although conceptualized as an emotional process, rather than an emotion proper, empathy is also categorized as a positive, other-oriented moral force that motivates prosocial behavior (Eisenberg, Valiente, & Champion, 2004; however, see Tullett, Harmon-Jones, & Inzlicht, 2012). Finally, condemning emotions, such as anger and disgust, are thought to play an important role in reinforcing moral norms in the society as they are usually directed towards a third party (Rozin, Lowery, Imada, & Haidt, 1999). Individuals who are angered by witnessing injustice or defection, for example, are more likely to take actions to punish the transgressor and reaffirm the norm of cooperation (Fehr & Gachter, 2002).

Independent of these broad categories, researchers have also suggested different classifications for self (e.g. guilt, pride) versus other-oriented (e.g. love, anger) moral emotions (Tangney et al.,
Within the past decade, moral psychologists have begun to discuss the role that these discrete moral emotions such as guilt and love have in motivating moral behavior (Pfister & Bohm, 2008). Furthermore, psychologists have theorized about the ways in which these “feelings” might drive moral decision-making, in both their real and anticipated forms (Baron, 1992; Tangney et al., 2007). For instance, if an individual anticipates tremendous guilt and shame upon stealing, it is likely that these anticipated emotions will deter them from transgressing. Similarly, if an individual experiences a surge of guilt or fear before stealing a candy bar off a store shelf, this moral emotion might prevent them from transgressing (Tangney et al., 2007). In these cases, moral emotions may serve as both motivation and information. In the stealing example, we can see how the psychosomatic feelings of guilt and shame might act as information, signaling that stealing is morally wrong or is typically not advantageous for long-term self-interest (Rand & Nowak, 2013). At the same time, emotion’s role as a moral motivator is more complex because individuals might be motivated by a number of factors, such as to relieve the pre-decisional negative affect (e.g. guilt), avoid post-decisional anticipated negative affect (e.g. shame), or achieve post-decisional positive affect (e.g. pride; Baumeister, Vohs, DeWall, & Zhang, 2007). In other words, moral emotions can provide both the information and motivational force to do the “right thing” (Kroll & Egan, 2004).

Interestingly, most theoretical accounts of moral emotions have been largely based on studies that have relied on self-report measures of morality such as hypothetical moral scenarios (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Greene, 2007; Sheikh & Janoff-Bulman, 2010), biographical recall (Keltner & Buswell, 1996; Tracy & Robins, 2006), and perhaps most importantly, moral judgments (Haidt, 2001; Inbar, Pizarro, Knobe, & Bloom, 2009; Rozin et al., 1999). More often that not, the hypothetical scenarios that are utilized in such studies are extreme and lack ecological validity (Greene et al., 2001; 2004). Although these studies have been integral in placing morality on the social psychology map, it is difficult to conclude precisely what individuals’ responses to such dilemmas imply for real-life moral behaviour. However, because the predominant theories of moral emotions have almost entirely excluded behavioural measures from their investigations, there exist very real drawbacks to relying on these theories alone when attempting to discuss the role of emotions for real-life
moral behaviours (see Blasi, 1980 for an early review of studies comparing moral behaviour to moral attitudes).

Conceptual Problems with Self-Report in Moral Psychology

Although the over-reliance on self-report measures is a concern in almost all areas of social psychology (Baumeister, Vohs, & Funder, 2007), I propose that this “behavioural gap” in the moral psychology literature is especially problematic for reasons I will outline below. The issue with relying on self-report measures in moral psychology depends on the type of measure being used. Currently the field has been dominated by two different types of measures – 1) scenarios that require a judgment or reaction (i.e. how morally acceptable is it to burn your national flag?), and 2) hypothetical dilemmas that require a decision (i.e. would you push the man off the footbridge in order to save five people?; Monin, Pizarro, & Beer, 2007). In the former case, it is likely that judgments will not always align with behaviors for the simple reason that people often transgress even when they recognize that their actions are morally “wrong”. Although most people would agree that lying and cheating are morally wrong, we engage in these behaviors quite frequently by justifying them in a variety of ways (Festinger, 1957; Mazar, Amir, & Ariely, 2008). This attitude-behavior dissociation presents a challenge for applying findings in the domain of moral judgments to real-life behaviour. In the case of hypothetical moral decision-making, the self-report problem is two-fold. First and foremost, many of the scenarios that are used are extreme and lack ecological validity (e.g. Greene et al, 2001; 2004; 2008; Haidt, 2001). Second, the emotions that may be present during hypothetical moral decision-making are likely quite different (and certainly not as intense) than those elicited during real-life moral decision-making. For this reason, I propose that studying the relationship between actual and hypothetical moral decision-making is an important endeavour.

4 Emotions May Drive Moral Behaviour

Although the majority of research in moral psychology has focused on hypothetical moral dilemmas, several recent studies have begun to investigate the role that emotional experience has on motivating real-life moral behaviour. The results of these experiments overwhelmingly point to the same inference – that emotions or affective experiences can motivate people to do the “right thing”. I make the distinction here between full-blown emotional responses that are typically conscious, and more sustained, and affect, which may be unconscious and fleeting, but
may still have the potential to influence behaviour (Zajonc, 1980). I believe that both are important for influencing moral behaviour. For instance, a recent study found that participants act more morally when they are simply led to believe that they are physiologically aroused, supporting the notion that affective experience can serve as information (Gu, Zhong, & Page-Gould, 2013). Participants who were presented with false somatic feedback (i.e. an audio recording of a speedy heartbeat) and were led to believe that it was their own heart beating were significantly more likely to volunteer their time and less likely to cheat than participants who listened to a calm heartbeat (Gu et al., 2013). The results of such studies suggest that individuals rely heavily on affective cues when engaging in real-life moral decision-making, and that these cues may deter transgressions (in both their perceived and actual forms) when they signal states of high arousal.

Within recent years, researchers in the field of moral psychology have theorized about why these visceral states motivate moral behaviour. Do emotions motivate moral behaviours because ultimately those moral behaviours benefit people in the end? Or, do people engage in moral behaviours and refrain from immoral ones because they develop affective affinity or aversion to those behaviours, independent of what consequences those behaviours may cause? These are two general approaches towards explaining why emotions motivate morally charged actions.

4.1 Intuitively Prosocial: An Evolutionary Approach

Within the past two decades or so, social scientists have begun to theorize about the utility of emotional processes for moral decision-making. For instance, Robert Frank has posited that emotions serve as commitment devices (also see Haidt, 2003), deterring us from self-interested behaviour. Our emotions, he argues, are what push us to act prosocially, even if the immediate consequences of prosocial behaviour are less favorable than that of selfish behaviour. From an evolutionary perspective, the presence of moral emotions can help explain why we live in mainly cooperative societies (Rand & Nowak, 2013), why people will punish transgressors even when this comes at a personal cost (Feinberg, Willer, Stellar, & Keltner, 2012), and why most individuals do not behave according to the clear-cut rules of self-interest in economic decision-making games (Frank, 1988). For instance, the anger or indignation that presumably drives individuals to reject unfair offers in economic games like the Ultimatum Game, despite incurring a personal cost, likely reflects an evolutionary mechanism that favors punishing
transgressors. This sort of costly punishment is arguably necessary for deterring future transgressions, and for the maintenance of societal order (Fehr & Gachter, 2002). Other studies have found that prosocial behaviour is frequently observed in isolated and anonymous interactions, where the prospect of direct reciprocity is virtually nonexistent (Frank, Gilovich, & Regan, 1993).

Rand, Greene, and Nowak (2012) have suggested that prosocial cooperation may act as our default state because cooperation is not only advantageous, but also critical for the societal norm of reciprocity (Nowak & Sigmund, 2005). Defections, on the other hand, are normally associated with negative consequences (i.e. punishment). In a series of studies, Rand and colleagues (2012) found that individuals make more generous contributions in various economic decision-making games when they take less time to make the decision, and also when they are put under time constraints. Similar effects have been found among cognitively “depleted” participants. In one study, depleted participants made significantly more generous offers in an economic decision-making game than their non-depleted counterparts. In a second study, depleted participants were significantly more likely to reject unfair monetary offers as a form of punishment, even if this meant that they themselves would have to incur a financial loss (Halali, Bereby-Meyer, & Meiran, in press). All of these seemingly “irrational” behaviours help to maintain societal order by which all members reap greater overall benefits. In other words, when individuals lack the time and resources to engage in effortful, reflective processing, they rely on automatic intuitions or basic “System 1” emotional processes (Shiffrin & Schneider, 1977; Stanovich & West, 1998; Chaiken & Trope, 1999), which in these cases led to prosocial choices.

A related set of studies demonstrated that framing moral decisions as “intuitive”, as opposed to “deliberative” served to encourage ethical behaviour (Zhong, 2011). It seems that simply activating an intuitive (by writing about how you feel about a specific issue), vs. deliberative (by solving math problems) mindset, or simply asking participants how much they “feel” they want to donate as opposed to asking them to “decide” how much they want to donate, increases prosocial behaviour. Importantly, this same study found that the presence of aversive emotions such as fear and disgust predicted ethical behaviour. A related set of experiments investigated the role of guilt in motivating prosocial behaviour, and found that a guilt induction served to increase prosocial behaviour, but only when this behaviour didn’t come at a personal cost (de
Hooge, Nelissen, Breugelmans, & Zeelenberg, 2011). In another study, participants’ self-reported guilt and empathy predicted helping behaviour toward a partner in a lab task (Gino & Pierce, 2009). This work implies that at least in some cases, our feelings, or intuitions may drive prosocial behaviours.

4.2 Aversion to Acts of Transgression: A Non-Consequentialist Account

Evolutionary accounts of moral psychology posit that emotions deter transgressions because they help us make decisions that will maximize positive consequences in the long-term (Rand & Nowak, 2013). Recent empirical data, however, suggests that our emotions may be active at a more basic level. Specifically, the results of several studies suggest that the act of transgressing in and of itself is aversive, and that the emotions associated with aversion to immoral actions are what may drive individuals to act prosocially. For instance, a recent study found that willingness to induce harm in a moral dilemma was predicted by individual differences in psychophysiological arousal (Cushman, Gray, Gaffey, & Mendes, 2012). The results of this study also revealed that psychophysiological arousal was higher for participants simulating harmful actions (i.e. hitting a plastic baby doll), than for participants witnessing others perform these actions, suggesting that aversion to harmful actions extends beyond consequentialist considerations (Cushman et al., 2012). Related research on cheating behaviour suggests that the best way to decrease cheating behaviour is to make the act of cheating more aversive, by making it less abstract. For example, forcing participants to pay themselves in real money instead of tokens causes participants to “steal” less. Interestingly, manipulating consequences (i.e. likelihood of getting caught) does not affect cheating rates (Mazar, Amir & Ariely, 2008).

The idea that people’s emotional reaction to the behaviour itself may be a critical driver of the behaviour also has been supported on a neural and physiological level. Research conducted with psychopaths, for instance, reveals that they may lack empathy (Blair, Jones, Clark, & Smith, 1997) and that they may exhibit decreased physiological responses to affective stimuli, allowing them to transgress “with no remorse” (Williamson, Harpur, & Hare, 1991; Kiehl et al., 2001; Vaidyanathan, Patrick, & Cuthbert, in press). These ideas are directly applicable to the somatic markers hypothesis, which suggests that sometimes our emotional systems are able to react to an action before we cognitively assess the consequences of the action. There is also a plethora of
evidence showing that individuals with damage to brain regions that are involved in emotional processing (e.g. prefrontal cortex, amygdala) exhibit atypical behaviour in social as well as morally charged situations (Bechara, Damasio, Damasio, & Anderson, 1994; Bechara, 2004; Bechara, Damasio, Damasio, & Lee, 1999). Critically, these same individuals also exhibit decreased physiological processing of emotional stimuli (Damasio, Tranel, & Damasio, 1990). Work of this nature is what may have allowed moral psychologists to rethink the notion that emotions cloud or interfere with sound moral decisions and recognize the important role that emotional experience plays in self-regulation and moral decision-making.

5 Summary

Decades of research in social psychology suggests that peoples’ attitudes are often incongruous with their behaviours, with more recent research revealing that dissociation may be attributed to the fact that people have poor insight into their future emotional states. If these emotional states are as critical for driving moral actions as both evolutionary and non-consequentialist accounts of moral behaviour suggest they are, it is then likely that people might underestimate their propensity to act morally. I predict that people will act more morally in actual moral dilemmas than they will forecast acting in hypothetical versions of these same dilemmas. Critically, I predict that this incongruence will be explained by differences in affective experience between individuals who are acting vs. forecasting in moral dilemmas and that this gap can be attenuated to the degree that affect is considered during the moral forecasting process. A large majority of moral psychologists rely on hypothetical scenarios to study the nature of moral decision-making. Thus, I propose that clarifying the relationship between actual and hypothetical moral decision-making, as well as uncovering methods by which we can align peoples’ moral forecasts more closely with actual moral behaviour are important and worthwhile pursuits.
Chapter 2: Errors in Moral Forecasting in a Prescriptive Moral Dilemma

1 Introduction

1.1 Prescriptive vs. Proscriptive Morality

Janoff-Bulman and colleagues (2009) have posited that an important distinction in the realm of moral behaviour is one between prescriptive and proscriptive morality. In prescriptive moral dilemmas, the rules center around what one should do in order to be a moral person (i.e. helping those in need, giving to charity, donating blood etc.). Thus, the rules for these types of moral dilemmas are more abstract, but also approach-based. Adhering to rules of prescriptive morality results in moral credit. Proscriptive morality, on the other hand, centers around what one should not do in order to be a moral person (i.e. lying, cheating, stealing etc.). Consequently, the rules in these sorts of moral dilemmas are more explicit, concrete, but also inhibition-based, and breaking these rules typically results in moral blame (Janoff-Bulman et al., 2009; Janoff-Bulman & Sheikh, 2010). Because both of these moralities are valued and well-represented in individuals’ moral repertoires, I aimed to investigate the relationship between peoples’ moral behaviours and moral forecasts for both types of moral situations. In Study 1, I focus on prescriptive morality.

The Dictator Game as a Measure of Prescriptive Moral Behaviour

In order to tap into prescriptive morality, I used a one-shot Dictator Game in which participants have to split $10 between themselves and a confederate (Hoffman, McCabe, & Smith, 1994). This paradigm has been previously used to study moral behaviour (Caruso & Gino, 2010; Shariff & Norenzayan, 2007) and is successful because it captures the moral foundation of fairness – a facet of morality that is widely endorsed (Graham, Haidt & Nosek, 2009). Unlike other previously examined prescriptive moral behaviours such as charity donation, the Dictator Game also eliminates the potential for diffusion of responsibility (Darley & Latane, 1968). Finally, the Dictator Game presents an advantage over other economic decision-making games,
such as the Prisoner’s Dilemma or Ultimatum Game, because the participant is fully responsible for the outcome of the game and does not need to worry about potential free riders.

1.2 Hypotheses

If emotions such as guilt that may fuel moral actions in prescriptive moral dilemmas (Pfister & Bohm, 2008), and if these emotions are less intense during hypothetical considerations of these same moral dilemmas, it is likely that a dissociation between individuals moral behaviours and forecasts will be observed. Specifically, I predicted a main effect of action versus forecasting, such that individuals’ predictions about how much money they would leave for the confederate will differ significantly from their actual behaviour. In other words, I predicted that participants who actually play the Dictator Game would leave significantly more money for the confederate than participants in the forecasting condition would forecast leaving.

2 Methods

2.1 Participants

One hundred and twenty two participants were recruited from the University of Toronto Scarborough participant pool to participate in a psychology study in return for course credit. A total of 2 participants were excluded from all analyses because they failed to properly follow experiment instructions¹, and one additional participant was excluded because he/she guessed the purpose of our study during the suspicion probe. This left 119 participants in the sample (71% female; $M_{age} = 18.70$).

2.2 Procedure

Participants were randomly assigned into one of two conditions in a 2-factor (behaviour vs. forecasting) between-subjects design. Participants were led to private cubicles where all questionnaires and information was presented to them on a computer screen. After providing demographic information, participants had to either play the Dictator Game (moral action condition) or predict their behaviour in the Dictator Game (moral forecasting condition).

Dictator Game
Participants in the action condition proceeded to play a one-shot version of the anonymous Dictator Game (Hoffman et al., 1994). Participants were given two envelopes, one marked “giver” and the other “receiver.” The giver envelope contained 10 one-dollar coins. Participants were given written and verbal instructions explaining that they had randomly been assigned as the giver in this game and that their job was to divide the 10 coins however they liked. They were told that the receiver was also a participant in the experiment whom they would not meet. Participants were assured that only the “receiver” would know about their decision and that their identity would be kept hidden; so they were assured that they would be completely anonymous. In reality, there was no receiver. Participants in the moral forecasting condition were presented with a written description of the Dictator Game, and also listened to a prerecorded audio of the instructions over a set of headphones. Participants were then asked to predict how many coins (out of 10) they would keep for themselves.

3 Results

The primary interest of this study was to examine differences in money given when participants were engaged in moral action vs. moral forecasting. I conducted an analysis of variance (ANOVA) and found a main effect for action vs. forecasting, such that action participants left significantly more money for the “receiver” \( (M = 4.47, SD = 2.62) \) than forecasting participants predicted leaving \( (M = 3.49, SD = 2.36) \), \( F(1, 117) = 4.53, p = .035, d = .39 \) (see Figure 1). These results suggest that individuals may underestimate their propensity to act morally in a prescriptive moral dilemma.

4 Discussion

In Study 1, I found evidence for a dissociation between peoples’ moral forecasts and peoples’ moral behaviours. Specifically, the data indicate that individuals in a real-life moral dilemma gave more money to an anonymous “stranger” than their counterparts who were presented with the same hypothetical dilemma predicted they would give. These results imply that individuals may underestimate their moral capacities in prescriptive moral situations where the moral rules are centered on fairness (Haidt & Graham, 2007). Although the results of Study 1 reveal an incongruence between peoples’ behaviours and forecasts in a prescriptive moral dilemma, it is
not clear whether the same result will be observed in different types of situations – specifically in proscriptive moral dilemmas, where the rules are more explicit and inhibition-based (Janoff-Bulman et al., 2009).
Chapter 3: Errors in Moral Forecasting in a Proscriptive Moral Dilemmas

1 Introduction

In Study 1, I found that individuals may underestimate their propensity to act morally in a prescriptive moral dilemma. In Study 2, I aimed to extend these findings to a proscriptive moral dilemma. In proscriptive moral dilemmas, the rules are centered around what one should not do in order to be a moral person. The rules surrounding proscriptive moral dilemmas are more explicit, and the consequences of transgressing are typically more severe. In addition, proscriptive immorality results in more moral blame than prescriptive immorality (Janoff-Bulman et al., 2009). For these reasons, I suspected that the emotions elicited during proscriptive moral dilemmas might be more intense. Thus, if individuals misforecast their behaviours due to an inability to tap into future emotional states (Van Boven et al., 2005), the behaviour-forecasting gap should be exacerbated for proscriptive moral dilemmas.

1.1 The Omission Bias

Another factor I was interested in exploring in Study 2 was the “omission bias”, whereby transgressions that are caused by omission, or inaction are typically judged as less severe than those that are caused by commission, or explicit actions. Past research has suggested that this difference in moral judgments across dilemmas of omission and commission occurs because it is typically more difficult to infer intent for transgressions of omission (Baron & Ritov, 2004; Ritov & Baron, 1999). In addition, there seems to be something specifically aversive about committing the physical action that constitutes a transgression that goes above and beyond the consequences that this action results in (Cushman et al., 2012). If transgressing through commission might be more emotionally aversive than transgressing through omission, I wondered whether the forecasting-behaviour discrepancy might be affected by the nature of the transgression at hand (commission vs. omission). Because past work suggests that people are conscious of their own “omission bias” (Cushman, Young, & Hauser, 2006), I suspected that this bias would translate to behaviours, but also to forecasts.
1.2 Hypotheses

For this study, I used a math task on which participants had the opportunity to cheat as the measure of morality. This sort of moral dilemma provides a clear comparison to that of Study 1 because although concerns of cheating also fall into the moral domain of fairness (Haidt & Graham, 2007), rules surrounding cheating are conceptualized as a proscriptive moral rule. In line with the results of Study 1, I predicted a main effect of action versus forecasting in both the commission and omission conditions, such that participants in the action condition would cheat less than their counterparts in the forecasting condition would predict cheating. I also predicted a main effect of omission vs. commission such that participants in the omission condition would cheat more, but would also predict cheating more than those in the omission condition.

2 Methods

2.2 Participants

One hundred and sixty participants were recruited for this study from the University of Toronto Scarborough for course credit. Twelve participants were excluded from all analyses because they guessed the purpose of our experiment during the suspicion probe, and an additional 3 participants were excluded from all analyses because they were extreme outliers. This left 145 participants in our sample (58% female; $M_{age} = 18.94, SD = 1.76$). Including these participants in our sample did not statistically change our results.

2.3 Procedure

Upon arrival to the lab, participants were led to private cubicles where all information was provided on the computer. After completing several demographic questionnaires, all participants were randomly assigned to one of four conditions in a 2 (action vs. forecasting) X 2 (omission vs. commission) between subjects design.

Math Task

Participants assigned to the action conditions were required to complete a math test on the computer, consisting of 15 simple, but tedious, arithmetic problems (e.g. $45 + 679 + 8 + 11 +$
234 + 50 – 71 – 1 – 524 – 25 =?), without the use of pencil or paper. Participants were informed of a “glitch” in the software, which differed depending on which condition they were in. In the commission condition, participants were informed that if they pressed the spacebar, the answer would appear on the screen. The experimenter then asked participants not to press the spacebar, to ensure their honest participation. In the omission condition, participants were informed that the answer to each question would appear on the screen if they did not press the enter key within 5 seconds of the question appearing. In this condition, the experimenter asked participants to press the enter key as soon as each question appeared on the screen. In this way, participants could cheat through omission, by simply waiting for the answer to appear on the screen. Importantly, participants in both conditions were told that we would have no way of knowing whether or not they pressed the spacebar/enter key, but that we would appreciate their honest participation. In reality, key presses were recorded. Finally, the experimenter informed participants that they would be rewarded with $5.00 if they answered 10 or more questions correctly. In the two forecasting conditions, participants were asked to forecast how many times they would cheat if they were in the above situations, by entering a number from 0-15 into a text box. They were then presented with an example of a question from the math task in order to allow them to assess the difficulty of the test.

3 Results

In order to correct for violations of normality within our data, I conducted a Poisson regression to explore the effect of cheating/predicted cheated (-1 = action, 1 = forecasting) within the commission condition (commission = 0, omission = 1), and then within the omission condition (commission = 1, omission = 0). I report estimated marginal means throughout. I found that there was marginally significant interaction between action vs. forecasting and commission vs. omission, $b = .29, SE = .15, \text{Wald } \chi^2 = 3.68, p = .055, d = .32$. As hypothesized, I found a significant main effect of action vs. forecasting, whereby participants in the action condition cheated significantly less than participants in the forecasting condition predicted cheating, $b = -1.81, SE = .27, \text{Wald } \chi^2 = 46.49, p < .001, d = 1.40$. There was also a main effect of omission vs. commission, whereby participants in the omission condition cheated more and predicted cheating more than those in the commission condition, $b = .38, SE = .12, \text{Wald } \chi^2 = 10.86, p = .055, d = .58$. I probed this interaction term using the methods of Aiken and West (1991; West, Aiken, & Krull, 1996). Consistent with the results of Study 1, I found that participants in the
commission condition exhibited a significant discrepancy between how many times they actually cheated ($M = 0.53$), and how many times they predicted cheating ($M = 3.26$), $b = .61$, $SE = .08$, $Wald \chi^2 = 63.14$, $p < .001$, $d = 1.76$. I found the same dissociation in the omission condition, such that participants in the moral action condition cheated significantly less ($M = 1.41$) than participants in the moral forecasting condition predicted cheating ($M = 4.79$), $b = .91$, $SE = .13$, $Wald \chi^2 = 46.49$, $p < .001$, $d = 1.37$ suggesting that individuals may underestimate their moral capacities regardless of whether the moral transgression is commission or omission-driven. Finally, the results of this study suggest that the “omission bias” translates to actual behaviour, such that participants in the commission condition cheated significantly less ($M = 0.53$) than participants in the omission condition ($M = 1.41$), $b = -.49$, $SE = .14$, $Wald \chi^2 = 11.72$, $p = .001$, $d = .360$. The “omission bias” was observed within the forecasting conditions as well, whereby participants in the commission condition predicted cheated significantly less ($M = 3.26$) than participants in the omission condition ($M = 4.79$), $b = -.19$, $SE = .06$, $Wald \chi^2 = 10.86$, $p = .001$, $d = .58$ (see Figure 2).

4 Discussion

In Study 2, I replicated and extended the results of Study 1. Specifically, the results of this study suggest that individuals may underestimate their likelihood to act morally not only in prescriptive moral situations, but also in proscriptive moral situations. Specifically, I found that participants tend to underestimate their morality in proscriptive moral dilemmas where transgressions require explicit actions (i.e. pressing a button to cheat), and are commission-based, but also in moral dilemmas where individuals can transgress by omission (i.e. passively waiting for the response to appear on the screen). Because transgressions caused by omission are usually viewed as less severe (Baron & Ritov, 2004), than those caused by commission, this suggests that there is an important distinction between these two types of moral dilemmas. The results of this study seem to corroborate this idea, since participants were more likely to cheat and also predicted cheating more in the omission condition.

Overall, the results of Study 2 suggest that individuals may underestimate their propensity to act morally in a wide range of moral dilemmas. Interestingly, the marginal interaction also suggests that the moral underestimation effect may be marginally stronger in the commission condition than in the omission condition. If transgressions of commission are more emotionally aversive
than transgressions of omission, this may explain why people have an even harder time simulating the emotional experience of transgressions caused by commission. Interestingly, the discrepancy between peoples’ behaviours and forecasts in this study was stronger than the discrepancy found in Study 1. One possibility is that people have an even harder time simulating the emotions present in proscriptive moral dilemmas because they are more intense than those present during prescriptive moral dilemmas. Finally, these findings suggest it is unlikely that the physiological stress that sometimes drives moral behaviours (e.g. Kiehl et al., 2001) causes a “freezing” response that then prevents people from transgressing. Rather, it seems possible that this affective experience motivates ethical action tendencies (i.e. by pressing the spacebar to avoid cheating). Although it seems plausible that affect plays an important role in motivating moral actions, the results of Study 2 do not allow for such conclusions. It is also possible, for instance, that participants misforecasted their behaviour due to a lack of information about the math task. In Study 3, I address these issues by simulating the entire math task for participants in the forecasting condition, and also by measuring participants’ affective arousal.
Chapter 4: Why We Make Moral Forecasting Errors: The Role of Affective Experience

1 Introduction

In Studies 1 and 2, I found that individuals may underestimate their likelihood to act morally. The goal of Study 3 was investigate the cause of this moral underestimation effect. Because past work has suggested that people may not appreciate the extent to which affective experience drives decision-making (Van Boven et al., 2005), and also that affective arousal may actually drive individuals toward moral action (Damasio, Tranel, & Damasio, 1990), I posited that the dissociation between peoples’ moral forecasts and behaviours might be driven by differences in affective arousal. Even though moral emotions are present during moral forecasts (Greene et al., 2001), if these emotions are less intense than during actual moral dilemmas (Bechara & Damasio, 2005), then individuals may underestimate the strength of their emotions when they are making predictions.

1.1 Physiological Arousal and Affect

Although there is disagreement on the nature of emotion (c.f., Ekman, 1992; Frijda, Mesquita, Sonnemans, & van Goozen, 1991; Izard, 1977; Russell, 1980; Scherer, 1993), most researchers agree that physiological arousal is a hallmark of emotional experience (e.g., Schacter & Singer, 1962). Measuring affective experience with psychophysiological measures has two advantages; 1) it will allow for me to measure aspects of affective experience that are unconscious, but that may be important in influencing behaviour (Winkielman & Berridge, 2004), and 2) it will allow for me obtain measures of affective experience in the moment of decision-making. This is advantageous because individuals often have difficulty accurately reporting their emotions after the fact (Kahneman, Fredrickson, Schreiber, & Redelmeier, 1993). For the purpose of the present study, I will focus on physiological measures that have been related to the affective/motivational states of approach and avoidance, social engagement, as well as generalized physiological arousal. As such, I will measure activity of three physiological responses that are governed by the autonomic nervous system: Pre-Ejection Period (PEP), Respiratory Sinus Arrhythmia (RSA), and Skin Conductance Response (SCR).
Emotionality in the Autonomic Nervous System (ANS)

The ANS is the peripheral stress system that governs cardiovascular, pulmonary, and electrodermal physiological responses, in addition to other gastrointestinal functions. As I will now elaborate, there is reason to believe that two main branches of the ANS, the sympathetic nervous system (SNS) and the parasympathetic nervous system (PNS), may be reactive during moral dilemmas. Importantly, I hypothesize that the intensity of these reactions will drive moral behaviour.

The SNS and Moral Behaviour

At a very basic level, emotional experience can be differentiated into approach and avoidance tendencies (Davidson, Ekman, Saron, Senulis, & Friesen, 1990; Harmon-Jones, 2003). In a comparison of a number of ANS responses, Brenner, Beauchaine, and Sylvers (2005) systematically examined the physiological correlates of approach and avoidance behaviour during reward and extinction, and found that the SNS — as indexed through a decrease in PEP — was only active during reward trials. PEP reflects ventricular contractility, and is the only non-invasive measure that distinctly reflects SNS activity (Berntson, Cacioppo, Binkley, Uchino, Quigley, & Fieldstone, 1994; Schachinger, Weinbacher, Kiss, Ritz, & Langewitz, 2001). Specifically, PEP is the time from the beginning of electrical stimulation of the ventricles to the opening of the left aortic valve. A smaller PEP indicates a stronger heart contraction and greater sympathetic arousal. Research on the biopsychosocial model of motivated performance incorporates SNS activation as a component of the physiological pattern of “challenge” (Blascovich, Mendes, Tomaka, Salomon, & Seery, 2003), which has been directly compared to approach motivation (Blascovich, Mendes, Hunter, & Salomon, 1999). Taken together, these findings suggest that sympathetic activity is indicative of approach, and I thus hypothesize that it will become activated when participants are faced with the opportunity to cheat for monetary reward.

The PNS and Moral Behaviour

The PNS theorized to be directly involved in social engagement and social cooperation (Porges, 2006), perspective taking and compassion (Oatley, Keltner, & Jenkins, 2006), and behavioural flexibility (Berntson et al., 2007). The heart beats faster during inspiration and slower during
expiration (Hon, 1958; Lacey, 1967; Obrist, 1981), and the PNS directly regulates this synchrony of cardiopulmonary activity. Porges’ (2001) polyvagal theory posits that this regulation of cardiovascular output evolved among mammals to adapt to the complex demands of social coordination. RSA is a measure derived from heart and respiration rates, and is the only non-invasive index of PNS activity (Berntson, Quigley, & Lozano, 2007). Increases in RSA represent greater PNS activity, whereas decreases in RSA represent that PNS activity has halted. Decreases in RSA reliably covary with attentional focus and cognitive load (Kassam, Koslov, & Mendes, 2009; Tattersall & Hockey, 1995). To the extent that parasympathetic activity reflects social engagement and prosocial behaviour -- emotional processes I suspect are occurring for participants who decide to act morally -- I hypothesize that greater parasympathetic activity should underlie moral action.

**ANS Coupling**

I focused on PEP and RSA as two key measures of interest, because these measures uniquely reflect the activation of the sympathetic and parasympathetic branches of the ANS, respectively. Although traditional models of ANS activation have posited that the sympathetic and parasympathetic systems lie along a one-dimensional continuum, with these two systems acting reciprocally (e.g. Gaskell, 1900), more recent formulations suggest a two-dimensional model with non-reciprocal co-activation of both systems (Bernston, Caccioppo, & Quigley, 1991). In order to capture an appropriately nuanced understanding of physiological arousal, activity of both the SNS and PNS must be measured concurrently (Cacioppo, Uchino, & Berntson, 2007). In line with this modern view, I expect that the SNS and PNS will be coupled when participants are in a situation where they must act morally or immorally. I hypothesize that the SNS will be activated by the opportunity to cheat for money, and I expect that the PNS will be activated due to the socially relevant nature of moral decision-making.

**Affective Intensity**

Finally, I will also measure SCR as a correlate of generalized arousal. SCR reliably correlates with emotional arousal, and has been previously used in a number of emotion studies (e.g. Demaree, Schmeichel, Robinson, & Everhart, 2004; Gross & Levenson, 1997). For example, Demaree and colleagues (2004) found increased skin conductance level in participants that were asked to exaggerate their emotions. SCR is dually innervated by the sympathetic and
parasympathetic nervous systems, so it reflects generalized arousal across both autonomic systems. For this reason, I will measure SCR as correlates of generalized emotional arousal. I expect that the more aroused participants become when faced with a moral dilemma, the more likely they will be to act morally.

1.2 Hypotheses

For this study, I will use the “commission version” of the math task from Study 2 as my measure of morality, and will measure participants’ affective arousal while the complete this task, or forecast their behaviour on this task. I hypothesize that: 1) participants in the action condition will cheat significantly less than participants in the forecasting condition will forecast cheating, b) that the intensity of participants’ arousal will drive moral behaviour, and c) that differences in arousal between conditions will mediate the effect of action vs. forecasting on cheating behaviour/forecasts.

2 Method

2.1 Participants and Procedure

Sixty-seven participants (54% female; M_{age} = 20.03) from the University of Toronto participated for course credit. Two participants were excluded from psychophysiological analyses due to messy physiological data. Physiological sensors were applied, and participants were asked to sit still and relax for 30 seconds, which comprised the physiological baseline. After these 30 seconds, participants saw a “Recording Stopped” message on the computer screen. At this point, the experimenter asked participants if they would be willing to participate in another unrelated experiment. The experimenter explained that the recording had been turned off, but that it would be more convenient if the electrodes were removed at the end of the experiment. In fact, physiological responses were recorded throughout. Participants were then randomly selected into one of 3 conditions: moral action, moral forecasting, or control. Participants in the moral action had to complete a math test in which they had the opportunity to cheat. Participants in the moral forecasting condition had to predict how many times they would cheat if they were to complete such a math test. Finally, participants in the control condition had to complete the math test, but with no option of cheating.
Math Task

Participants assigned to the action condition were required to complete a math test on the computer, consisting of 15 simple, but tedious, arithmetic problems (e.g. $45 + 679 + 8 + 11 + 234 + 50 - 71 - 1 - 524 - 25 = ?$), without the use of pencil or paper. This task mirrored the “commission version” of the math task used in Study 2. In the forecasting condition, participants were presented with the same 15 math problems one by one, and indicated whether they would reveal the answer for each question. Presenting participants with each individual problem eliminated the possibility that participants would make forecasting errors due to a lack of information about the situation. In the control condition, participants had to complete the math task with no opportunity to cheat, thereby allowing us to separate the arousal elicited by a moral dilemma from the arousal elicited by solving difficult math problems.

2.2 Physiological preparation and recording

Physiological measures were obtained simultaneously using a BioNex impedance cardiograph and GSC amplifier (Gahanna, OH), sampling at a frequency of 1000 Hz. Four pre-gelled spot Ag-AgCl electrodes were attached to the neck and torso, and 2 pre-gelled Ag-AgCl electrodes were attached in a Standard Lead-II chest formation for measures of PEP and RSA. The cardiopulmonary data were cleaned and analyzed in 30-second epochs. These epochs were then averaged across the first two minutes of the stimulus task, and the 30-second baseline was subtracted from this average. PEP was calculated as the time in milliseconds between the depolarization of the left ventricle and the opening of the left aortic valve to eject blood. RSA was derived from a power spectral analysis of high frequency heart rate variability (Berntson et al, 2007). SCR was acquired through 2 pre-gelled Ag-AgCl electrodes that were placed on the thenar eminence of participants’ non-dominant palm. Within the SCR data, I was interested in the number of non-specific skin conductance responses (NSSCRs) within each epoch of the baseline and task blocks. NSSCRs are observed when skin conductance level increases by 0.05 µS.
3 Results

3.1 Behavioural: Action vs. Forecasting

I conducted a Poisson regression (action vs. forecasting), which revealed that participants who completed the math task cheated significantly less ($M = .96, SD = 1.65$) than participants in the forecasting condition predicted cheating ($M = 4.82, SD = 5.32$), $b = .97, SE = .10, Wald \chi^2 = 97.15, p < .001$ (see Figure 3). In other words, people acted more morally than they would have predicted. I next investigated the role of affective experience.

3.2 Psychophysiological: Action vs. Forecasting vs. Control

Pre-Ejection Period

I hypothesized that participants who had the option to behave immorally would exhibit greater PEP decreases during the math task than participants in the other two groups. Four participants were excluded from the analysis because they were extreme outliers\(^2\). I found significant differences in PEP reactivity between conditions, $F(2, 40) = 3.64, p < .05$ (Figure 4). Simple effects testing revealed that participants in the action condition showed greater decreases in PEP ($M = -9.50, SD = 7.65$), than participants in the forecasting condition ($M = -2.40, SD = 10.31$), $p < .05, d = .78$, and participants in the control condition ($M = -2.42, SD = 6.17$), $p < .05, d = 1.02$. The latter two groups did not differ from one another, $p > .90$.

Respiratory Sinus Arrhythmia

I expected participants in the action condition to exhibit greater RSA than participants in the other two groups. Three participants were excluded from the analysis because they were extreme outliers\(^2\). Figure 4 illustrates a significant difference in RSA reactivity among the conditions, $F(2, 42) = 3.77, p < .05$, with a simple effects test revealing that participants in the action condition displayed significantly greater increases in RSA ($M = .36, SD = .72$), compared to participants in the forecasting condition ($M = -.38, SD = .33$), $p < .05, d = 1.32$, or control condition ($M = -.36, SD = 1.13$), $p < .05, d = .76$, who exhibited decreases in RSA. As with PEP, these latter two groups did not differ from one another, $p > .90$. 
Skin Conductance Response

Finally, I examined NSSCRs as a correlate of arousal intensity. Figure 5 illustrates significant differences in NSSCR frequency among the conditions, $F(2, 41) = 4.59, p < .05$, with participants in the action condition exhibiting more NSSCRs ($M = 2.33, SD = 1.34$) than participants in the forecasting condition ($M = .20, SD = 2.01$), $p < .01, d = 1.25$. With SCR, however, we found no significant differences between the control ($M = 1.26, SD = 1.98$) and action conditions, $p > .10$. There was no significant difference between the control and forecasting conditions, $p > .10$.

3.3 Process: The Mediation of Action/Forecasting Dissociation

A multiple mediation model was analyzed, which involves simultaneous mediation by multiple variables. Preacher and Hayes (2008) have recommended that testing a multiple mediation model involves an analysis of the total indirect effect and an analysis of the specific indirect effects of each mediator. I tested the relationship between action/forecasting and cheating as mediated by PEP, RSA and SCR. PEP did not constitute a significant mediator, so I focused solely on RSA and SCR. A bootstrap analysis was used to test the significance of the indirect effects using 5,000 samples to yield a parameter estimate for both total and indirect effects. If the 95% bias-corrected confidence interval for the parameter estimate do not contain zero, then the indirect effect are statistically significant and mediation is demonstrated (Preacher & Hayes, 2008). Table 2 contains the confidence intervals for the tota and indirect effects of RSA and SCR on the relationship between action/forecasting and cheating. The absence of zero in the confidence interval for the total indirect effect indicates that the mediating effect of the combination of the two variables was significant. An examination of the confidence intervals for specific indirect effects indicates that RSA was a significant mediator. SCR, however, was not. Specifically, condition was negatively related to RSA ($b = -.71$) that, in turn, was negatively related to cheating ($b = -1.94$; see Figure 6).

4 Discussion

Taken together, the results of Study 3 suggest that being in an actual moral dilemma elicits more affective arousal than does simply predicting behaviour in a moral dilemma. Compared to the
moral forecasting condition, participants in the moral action condition showed greater sympathetic and parasympathetic activity, as well as greater generalized arousal. The consistency of this pattern across a constellation of physiological measures provides compelling evidence that participants facing a moral dilemma were more aroused than those who were forecasting moral behaviour. This study also provides evidence that those facing a moral dilemma—facing the possibility of cheating on a math test—showed greater sympathetic and parasympathetic arousal than those taking the same math test, but without the option of cheating. The sympathetic activation findings suggest that participants were more reward-focused in the cheating condition than the math condition. The opposite directions of the RSA reactivity between the two conditions suggest that the parasympathetic nervous system was activated among participants facing a moral dilemma, and deactivated among participants completing a math task with no moral component. Decreases in RSA among participants in the math condition likely reflect the increased cognitive load and attentional resources required by the math task. There were no differences between the cheating and math tasks in generalized arousal, however, as these two groups did not differ in SCR reactivity. It is important to note that the math task was difficult and stressful for all participants, even those that did not have the opportunity to cheat. Thus, it is not surprising that control participants would display significant arousal. Nevertheless, GSR is not the most reliable measures of arousal (e.g. Demaree et al., 2006), and taken as a whole, our results seem to indicate that there is something unique about being in an actual moral dilemma that seems to call upon emotional processes above and beyond those that may be called upon during a similarly stressful mental task that does not involve a moral dilemma, or during a hypothetical situation where people are asked to make moral forecasts.

In addition to suggesting that individuals seem to display more emotionality during moral behaviour than during moral prediction, the data indicate that these affective processes drive moral behaviour, thus accounting for the dissociation between actual behaviour and reported behaviour. Both parasympathetic arousal and arousal intensity mediated the relationship between action/forecasting and cheating, which supports my hypothesis that people make moral forecasting errors because they do not have access to how they would feel in the actual moral situation.
Chapter 5: Perceptions of Arousal Influence Moral Forecasting

1 Introduction

In Studies 1, 2 and 3, I found that individuals may underestimate their likelihood to act morally. Because the results of Study 3 suggest that the dissociation between moral actions and moral forecasts may be driven by differences in affective experience, I wondered whether the nature of this relationship could be altered by experimentally manipulating perceptions of physiological arousal among participants forecasting their behaviour in a moral dilemma. In other words, might simulating a state of physiological arousal lead to forecasts that more closely align with actual behaviours?

1.1 Affective Cues as Information

Damasio (1994) has proposed that attention to bodily cues (i.e. heart beat) is fundamental for effective decision-making. For instance, individuals with damage to the ventromedial prefrontal cortex, who have difficulty perceiving their emotional states are unable to properly navigate social contexts (Anderson, Barrash, Bechara, & Tranel, 2006; Bechara, 2004; Damasio, Tranel, & Damasio, 1990). Similarly, Schwarz and Clore (1988) have posited that emotions can serve as information in decision-making contexts, and that feelings facilitate decisions by informing us about our current state or situation (Schwarz, 2001). Interestingly, somatic feedback has been previously linked with increases in self-directed attention (Fenigstein & Carver, 1978) as well as increases in prosocial behaviour (Gu et al., 2012). If such feedback influences participants’ responses, this would suggest that the perceptions of affective experience play an important role in moral forecasting, above and beyond physiological arousal. As such, I wondered whether cues in the form of somatic feedback might be effective in influencing moral forecasting.

1.2 Hypotheses

In Study 4, I plan to use the same math task from Study 3 to test for assess differences in moral forecasting. If individuals make moral forecasting errors due to an inability to access the affective state inherent to real-life moral dilemmas, then simulating such a state should shift individuals’ forecasts to more closely match behaviours. In line with the view that emotions provide people with information in decision-making contexts, I suspect that providing
forecasts with false somatic feedback might serve to align moral forecasts more closely with moral behaviours (i.e. forecast cheating less). I predict that this manipulation may be effective in influencing individuals’ moral forecasts because the false somatic feedback provided would reflect the feeling state that is ostensibly experienced by individuals who are actually completing the math task.

2 Method

2.1 Participants

One hundred and twenty six participants were recruited for this study from the University of Toronto Scarborough for course credit. Upon arrival, each participant was led to a computer station in a cubicle. Two participants were excluded from all analyses because they guessed the purpose of our experiment, an additional participant was excluded because they questioned the authenticity of the heartbeat audio, and two additional participants were excluded because they were extreme outliers. This left 121 participants in the sample (72% female; M-age = 18.57).

2.2 Procedure

Participants were randomly assigned to one of three conditions in a between subjects design: 1) moral action, 2) moral forecasting with normal heartbeat, or 3) moral forecasting with fast heartbeat. In the moral action, participants completed the same math task (commission version) as in Study 3, on which they had to chance to cheat. Participants in this condition were not provided with any somatic feedback. Participants in the two forecasting conditions were asked to test a heart monitor for an unrelated study, which included a bogus electrode that was attached to their non-dominant inner wrist, and a headset. To manipulate perceived heartbeat, we played a prerecorded heartbeat sound through the headset. Participants in the normal heartbeat condition listened to a heartbeat of 60 beats/min, while participants in the fast heartbeat condition listened to a heartbeat of 96 beats/min. These two paces were selected according to the American Heart Association’s (2012a, 2012b) definition of relaxed and high heart rates. Participants were asked to listen to this heartbeat audio while completing the moral forecasting part of the experiment. For this study, I modified the forecasting task in order to ensure that participants in Study 2 were not making inaccurate predictions simply because they
did not possess enough information about the math task. In this study, I presented each of the 15 math questions to participants on the computer, and asked them to indicate whether or not they would reveal the answer to each question, one by one.

3 Results

The main objective of Study 4 was to explore the influence that providing somatic feedback might have shifting individuals’ moral forecasts to more closely match behaviours. Given that I wanted to test a 3-level categorical variable with Poisson regression, I used the model comparison approach formalized by Cohen and Cohen (1983, Ch. 4) that yields chi-square statistics. I conducted a Poisson regression of cheating/predicted cheating with condition as the predictor term. First, results revealed an overall effect of condition on cheating/predicted cheating, \( \text{Wald } \chi^2 = 75.42, p < .001, d = 2.57 \). The three experimental conditions were represented in the linear model with contrast codes that compared the two forecasting conditions to the action condition and the two forecasting conditions to each other. Next, I conducted a priori contrasts which revealed that participants in the moral action condition cheated significantly less (\( M = .63 \)) than participants in the two forecasting conditions taken together, \( b = 1.65, \text{Wald } \chi^2 = 63.52, SE = 0.21, p < .001, d = 2.10 \), but also that participants in the normal heartbeat condition predicted cheating significantly more (\( M = 3.89 \)) than participants in the fast heartbeat predicted cheating (\( M = 2.78 \)), \( b = 0.34, \text{Wald } \chi^2 = 6.71, SE = 0.13, p = 0.01, d = .48 \) (see Figure 3). Interestingly, participants in the fast heartbeat condition still predicted cheating significantly more than participants in the moral action condition actually cheated, \( b = -0.67, \text{Wald } \chi^2 = 6.71, SE = 0.26, p = 0.01, d = .48 \) (see Figure 7). Taken together, the results of this study imply that providing participants with somatic feedback that signals a state of high arousal serves to significantly shift individuals’ moral forecasts to more closely match actual behaviour, but that it does not completely eliminate the action-forecasting discrepancy.

4 Discussion

The findings of Study 4 suggest that although participants tend to predict behaving less morally than they might actually behave in a moral dilemma, their moral forecasts can be significantly influenced when they are provided with somatic information indicative of physiological arousal. Specifically, participants who listened to a fast heartbeat audio that they were made to believe
was their heartbeat forecasted cheating significantly less than participants who listened to a relaxed heartbeat. These data provide further evidence for the notion that the dissociation between moral behaviours and moral forecasts can be at least partially explained by differences in affective experience, and more specifically by the inability of forecasters to tap into the affective experience that is inherent to real-life moral decision-making.

Interestingly, participants who listened to the fast heartbeat still forecasted cheating significantly more than participants in the action condition actually cheated. These results suggest that even when participants are able to access certain aspects of the physiological experience inherent to moral dilemmas, this may not be sufficient in completely eliminating the behaviour-forecasting discrepancy. Preliminary data (not reported here) suggesting that this sort of somatic feedback does not alter the individuals’ actual somatic experience may support the notion that the perceived physiological experience was not fully simulated in the body and may help explain why even individuals who listened to the speedy heartbeat were not able to produce completely accurate moral forecasts.
Chapter 6: Individual Differences in Moral Forecasting Accuracy: The Role of Emotional Awareness

1 Introduction

The results of the aforementioned experiments have thus far provided evidence for the idea that access to affective experience might be a necessary prerequisite to accurate moral forecasting. This pattern of results begs the question – are there individual differences in moral forecasting accuracy? In other words, are some people better able to forecast their behaviours than others? And if so, what are the traits that may be particularly important? Specifically, I wondered about the role that individual differences in emotional awareness might play in moderating moral forecasting accuracy. Because my findings in Study 4 imply that some forms of emotional induction serve to improve individuals’ moral forecasts, I wondered whether individuals that are particularly low in emotional awareness might exhibit exacerbated forecasting errors, and whether people high in emotional awareness might be particularly good at predicting their behaviour.

1.1 Emotional Awareness and the Toronto Alexithymia Scale

Alexithymia is a sub-clinical trait that varies in magnitude among individuals. It is characterized by the inability to identify emotional feelings, a difficulty distinguishing between and describing feelings, as well as an externally-oriented cognitive style and impoverished fantasy life (Nemiah, Freyburger, & Sifneos, 1976). Individuals who score high on trait alexithymia typically exhibit low “psychological mindedness”, or personal insight (Taylor, Bagby, & Parker, 1997). They also display lower levels of emotional intelligence (Parker, Taylor, & Bagby, 2001), which is thought to reflect the ability to use emotion as information to guide one’s thinking and actions (Salovey & Mayer, 1989). In other words, alexithymia seems to tap into emotional awareness – and for this reason, I wondered whether exploring variations in this personality trait might reveal the mechanism by which the experience and awareness of emotional states influences moral forecasting accuracy.
1.2 Hypotheses

In this study, I employed the same math task paradigm from Studies 3 and 4, but also assessed participants’ emotional awareness using the Toronto Alexithymia Scale (Bagby, Parker, & Taylor, 1994). Because the results of my previous studies suggest that affective experience may be important for accurate moral forecasting, I hypothesize that participants high in trait alexithymia (low in emotional awareness) will exhibit exacerbated forecasting errors, while participants low in trait alexithymia (high in emotional awareness) will be significantly more accurate at forecasting their moral behaviour.

2 Methods

2.2 Participants

One hundred and twenty one participants were recruited for this study from the University of Toronto Scarborough for course credit. A total of seven participants were excluded from the sample because they guessed the purpose of our experiment during the suspicion probe, and one additional participant was excluded because they were an extreme outlier, leaving 113 participants in the sample (60% female; $M_{\text{age}} = 18.61, SD = 2.55$).

2.3 Procedure

Upon arrival, each participant was led to a private cubicle, where all study materials were presented on a computer. Participants were randomly assigned to one of two conditions: 1) moral action, or 2) moral forecasting, in which they either had to complete the same math task we utilized in Study 3, or forecast their behaviour in this moral dilemma. Upon completion of the math task, participants filled out Toronto Alexithymia Scale (TAS-20; Bagby et al., 1994; $\alpha = .84$), which has been previously shown to be a valid and reliable measure of trait alexithymia (Bagby et al., 1994; Bagby, Taylor, & Parker, 1994a). For this scale, participants had to indicate the degree to which they agreed with each of the 20 statements on a 5-point Likert scale, ranging from “strongly disagree” to “strongly agree” (e.g. I am often confused about what emotion I am feeling, I don’t know what’s going on inside me). Because alexithymia captures metacognitive emotional awareness, as opposed to actual emotional experience, I predicted that participants who scored high on trait alexithymia would make more extreme moral forecasting errors, but
that variations in this personality trait would not predict actual cheating rates in the moral action condition.

### 3 Results

In order to assess how trait alexithymia moderates moral behaviour and moral forecasting, I conducted a Poisson regression analysis of cheating/predicted cheated with condition (-1 = action, 1 = forecasting), alexithymia (mean centered), and the interaction term as predictors. I report actual means, rather than the log of means for ease of interpretation. This analysis revealed a significant main effect of condition on cheating, where participants in the action condition cheated significantly less than participants in the forecasting condition predicted cheating, \( b = .42, SE = .06, Wald \chi^2 = 45.11, p < .001, d = 1.62 \). There was no main effect of alexithymia on cheating, \( b = .09, SE = .13, Wald \chi^2 = .49, p = .49 \). However, I did find a significant interaction between condition and alexithymia, \( b = .46, SE = .13, Wald \chi^2 = 12.36, p < .001, d = .70 \). Simple slopes analyses revealed that although trait alexithymia had no influence on cheating rates in the moral action condition, \( b = -.37, SE = .23, Wald \chi^2 = 2.58, p = .11 \), it did moderate predicted cheating in the moral forecasting condition, such that participants high in trait alexithymia predicted cheating significantly more (\( M = 5.44 \)), than participants low in trait alexithymia (\( M = 3.18 \)), \( b = .56, SE = .13, Wald \chi^2 = 19.34, p < .001, d = .90 \). Finally, I wanted to explore whether trait alexithymia influenced the relationship between moral action and moral forecasting. Simple slopes analyses suggest that participants high in trait alexithymia exhibit a strong discrepancy between actual cheating (\( M = 1.41 \)) and predicted cheating (\( M = 5.09 \)), \( b = .64, SE = .08, Wald \chi^2 = 55.54, p < .001, d = 1.95 \), but that participants low in alexithymia exhibit a discrepancy between actual behaviour (\( M = 2.02 \)) and predicted behaviour (\( M = 2.97 \)) as well, \( b = .19, SE = .09, Wald \chi^2 = 4.30, p = .038, d = .40 \) (see Figure 8). Taken together, these findings suggest that individuals who are particularly high in alexithymia (i.e. low on emotional awareness) may be especially prone to making moral forecasting errors, presumably because they are not able to properly utilize their feelings as information. However, these findings also suggest that individuals who are low in alexithymia (i.e. high on emotional awareness) are also not entirely immune to such forecasting errors.
4 Discussion

The results of Study 5 imply that individual differences in emotional awareness or emotional acuity may moderate moral forecasting accuracy. Specifically, it seems that individuals who score high on trait alexithymia (i.e. display low emotional awareness) exhibit exacerbated moral forecasting errors. Interestingly, this data seem be consistent with the results of Study 4. In particular, I observed a significant behaviour-forecasting dissociation even among individuals who scored low on trait alexithymia (i.e. display high emotional awareness). I suspect this is the case because alexithymia measures one’s ability to identify and describe feelings, as well as their propensity for externally-oriented thinking, rather than their propensity to keenly feel their emotions. In other words, these data suggest that while having high emotional insight may provide individuals with a forecasting “edge”, it is not sufficient in eliminating forecasting errors altogether. It seems that even individuals who exhibit high emotional awareness are not fully immune to making such errors because they may be missing an integral piece of the emotional puzzle. It is likely that one must actually experience these feelings in order to fully simulate the emotional experience of a real-life moral dilemma. Indeed, past research has suggested that the visceral experience of emotion (i.e. psychophysiological arousal) is an important determinant in driving moral behaviour (e.g. Kiehl et al., 2001). If alexithymia taps into the metacognitive aspect of emotional experience, rather than visceral aspect of emotional experience, this might also explain why variations in this personality trait had no influence on actual cheating behaviour.
Chapter 7: General Discussion

Contrary to the widespread notion that people’s self-views are clouded by positive illusions (e.g. Baumeister, 1998; Sedikides, 1993; Epley & Dunning, 2000; Taylor & Brown, 1988), here I report findings showing that people sometimes underestimate their own morality. In both a prescriptive dilemma where the moral behaviour is to give money, and a proscriptive dilemma where the moral behaviour is to not cheat, I find that forecasters anticipate acting more immorally than people experiencing the dilemma actually act. Moreover, I find evidence that this gap between forecasting and action is explained, at least in part, by the inability of forecasters to simulate the emotional experience of committing an immoral act. I find that autonomic arousal predicts less cheating, but that is also predicts less forecasted cheating. Importantly, this difference in arousal between conditions seems to mediate the relationship between what people think they will do in a moral dilemma and what they actually do. Furthermore, I find that when participants’ perceptions of affective experience are enhanced – either by false somatic feedback or by high trait levels of emotional awareness – this gap is attenuated. These findings suggest that because people don’t fully invoke the affective experience of going through the dilemma, they underestimate the forces that keep them from transgressing.

1 Summary

In Study 1 I found evidence for this gap between forecasting and action in a Dictator Game, showing that people can also underestimate their morality in prescriptive dilemmas where moral behaviour is accomplished by doing something good rather than resisting something bad (Janoff-Bulman et al., 2009). The findings of Study 2 extend this effect further by showing that moral underestimation can also occur in proscriptive moral dilemmas where the rules are more explicit and inhibition based. I find evidence for this in situations where the transgression results from commission, but also from omission. Transgressions of omission are typically judged less harshly (Baron & Ritov, 2004) presumably because they are seen as less deliberate and it is more difficult to infer intent. My findings from Study 2 suggest that individuals are also more likely to commit such transgressions, and also to forecast committing such transgressions. These data suggest that there are different processes at play in these situations, or that perhaps transgressions caused by omission are less emotionally charged. Although further research is
needed to explore this possibility, Study 2 suggests that the tendency to underestimate moral behaviour extends across both types of moral dilemmas.

In Study 3 I tested the idea that this discrepancy between action and forecasting might arise because people are unable to accurately simulate the affective character of a moral dilemma. In order to tap into participants’ affective experience, I measured autonomic nervous system arousal while participants completed the study tasks and found evidence for the hypothesis that being in actual moral dilemma is more affectively intense than forecasting behaviour for that same dilemma. These results also suggest that arousal fuels moral behavior and moral forecasting, thus accounting for the dissociation between actual behavior and reported behavior.

Then, in Studies 4 and 5 I tested the possibility moral forecasting accuracy differs as a function of perceived affective experience. If this is the case, enhancing affective experience should lead to moral forecasts that more closely map on to the way that people actually behave. In Study 4, people were given false somatic feedback that indicated that their heart was beating rapidly. Compared to people who thought that their heart was beating at a normal rate, this feedback led people to predict that they would cheat less often. Thus, it seems that falsely mimicking part of the emotional experience of being in the dilemma made participants produce forecasts the more closely matched actual behaviour.

In Study 5 I further examined this explanation by looking at alexithymia as a potential moderator of forecasting accuracy. I reasoned that if moral forecasting errors stem from an inability to accurately forecast the emotional state of the dilemma, emotional awareness should then improve moral forecasting accuracy. This seems to be the case, as illustrated by the finding that lower alexithymia (i.e. higher emotional awareness) was associated with predictions that more closely matched the behaviour of people in the action condition.

2 Limitations

2.1 Achieving Full Forecasting Accuracy

Although I suggest that there are ways to improve the accuracy of people’s moral forecasts, it is important to note that here these efforts at improvement still fall short of matching actual behaviour. The fact that this discrepancy remains even with high emotional salience and
awareness suggests that there is still something missing when people make forecasts. Even when people are high on emotional awareness, or think that they are experiencing physiological arousal in the form of a rapid heart rate, they are still not simulating the experience with perfect accuracy. I suggest that this occurs because these manipulations affect the informational, but not motivational (or visceral) aspects of emotion. This interpretation is corroborated by the fact that false somatic feedback does not induce physiological changes such as increases in actual heart rate. Similarly, alexithymia measures awareness of emotion but not intensity of emotional experience, and as such the forecasts of people low on alexithymia may still lack the physiological impact of the actual situation.

2.2 Discrete Emotions

Taken together, the findings reported herein provide support for the idea that people underestimate their own morality because, in part, they fail to imagine how emotions will influence their behaviour when they are actually faced with a moral dilemma. One limitation of the studies reported here is that I do not have self-report measures of emotion, and thus cannot be certain that the conscious awareness of emotion is what drives moral behaviour (Study 3) and is what influences moral forecasts in Studies 3, 4 and 5. I made this methodological choice because of the difficulties of measuring emotional experience without interrupting it. Previous work has demonstrated that labeling an emotion or making attributions about its source can change the experience of that emotion as well as its influence on subsequent behaviour (Lieberman, Eisenberger, Crockett, Tom, Pfeifer, & Way, 2007; Schachter & Singer, 1962). An alternative to measuring self-reported emotion while it is being experienced would be to ask participants to reflect on their emotions after the fact, but there is reason to believe that people may not be able to do this accurately (Kahneman et al., 1993). Future work will need to explore ways in which we can effectively isolate the specific emotions that may be responsible for moral behaviour and moral forecasting.

2.3 The Self-Serving Bias

Although here I document people’s tendency to underestimate their morality, there also appear to be times when people overestimate their positive qualities (e.g. Epley & Dunning, 2000; Pronin, Lin, & Ross, 2002; Svenson, 1981). The large majority of research on the self-serving
bias effect utilizes comparisons between self and others. In other words, participants must rate or predict their behavior in comparison to the average person. Such methods, for example, have led to participants to claim that they are less biased than the average person (Pronin et al, 2002) or better drivers than the average person (Svenson, 1981). In all of the current studies, participants simply had to predict their own actions in a moral dilemma. Perhaps if participants were asked to predict their behavior relative to the average person, their predictions of their own morality would have increased. It seems possible that social comparison would incite a competitive drive within individuals, causing them inflate their positive qualities. Similarly, if participants were asked to predict how morally another person would behave, they may predict even less moral behavior for others than they do for themselves. Applied to my findings, self-enhancement research suggests that moral forecasting errors will be amplified when predicting the moral behavior of others.

Perhaps most applicable to the current research, is the work of Epley and Dunning (2000) who examined the dissociation between actual behavior and prediction in a series of dilemmas. Across several different scenarios, they found that individuals were less altruistic than they predicted, a finding they dubbed the “holier than thou effect”. Although this work is quite informative, I would like to address why it differs significantly from the current research.

Primarily, it seems like a number of the studies conducted by Epley and Dunning (2000) did not involve “pure” measures of morality. For example, some of their studies did not allow for a “fair” option, forcing participant to allot all benefits to themselves or all benefits to the confederate. In addition, one of their dependent variables was the Prisoner’s Dilemma game, in which trust plays a larger role than prosociality. More generally, however, it seems as if many of the paradigms used in such self-serving bias studies are not as emotionally evocative as the paradigms we employed in our research. I feel that this can partially be attributed to the fact that in both the math task and Dictator Game, participants were very close in social distance to the recipient of their potential transgressions (the experimenter and a fellow student, respectively). And social distance, after all, is recognized as an important factor in motivating moral behavior (Eckel & Grossman, 1996; Frohlich & Oppenheimer, 2001; Burnham, 2003; Charness & Gneezy, 2003). Given my findings, it seems that less social distance may actually elicit higher affect in moral dilemmas.
Pfister and Böhm (2008) point out that emotion is largely responsible for motivating behavioral decision-making. They also argue that there are often competing emotional drives that are present when individuals are faced with decisions (e.g. the guilt of cheating vs. the potential pleasure of obtaining $5.00). The parasympathetic activation among participants in the cheating condition in Study 3 suggests that the social nature of the emotions dominant in the dilemmas I employed may be fundamentally different from those invoked in the studies conducted by Epley and Dunning (2000). Overall, it is evident that the self-serving bias exists, especially when we compare ourselves to others. As such, conditions that are high in social distance (e.g. charity), and thus low in affect, seem to produce the self-serving bias effect, especially when the situation does not allow for a “fair” or purely moral decision. However, when we are not engaging in social comparison and the situation is low in social distance (e.g. cheating), and thus high in affect, it seems that we may actually underestimate our tendency to act morally. Exploring these variables as they relate to moral behavior and moral forecasting may be a fruitful avenue for future research.

In sum, the current findings do not necessarily imply that people will always act more morally than they think; rather, I suggest that people have trouble simulating the emotions that they will experience in a moral situation. In some cases, those emotions might include guilt or shame, and thus encourage people to act morally as we see in the studies documented here. In other cases, however, those emotions might include a strong desire for the outcomes of an immoral act (e.g. stealing money or committing adultery), and in these cases people might end up acting more immorally than they would have expected if they had imagined the situation in the absence of those emotions. For instance, other moral overestimation effects that have been previously reported might be a function of the high monetary payoff for immoral behavior (approximately 4 times higher than what I offered in my experiments), and the ostensible approach-related emotions associated with this reward that may have been underestimated by the forecasters (FeldmanHall et al., 2012). Another possibility is that certain transgressions are construed as less morally acceptable than others. For example, delivering electric shocks to a confederate for monetary gain might sound less morally acceptable than cheating on a test for monetary gain, and might also explain why people were less likely to forecast immoral behavior in the former context (FeldmanHall et al., 2012). Although we do not know exactly how morally acceptable our participants found cheating on a math test to be, recent work suggests that undergraduate
study participants place cheating for monetary gain on the higher end of the immorality spectrum (Meindl & Graham, in press).

3 Future Directions

3.1 Eliminating the Behaviour-Forecasting Gap Through Visceral Experience

Although the results of the current research suggest that it is possible to increase the accuracy of the individuals’ moral forecasts, the manipulations and individual difference measures explored herein were not successful at fully attenuating the gap between moral behaviour and moral forecasting. The nature of these independent variables (i.e. false somatic feedback and trait alexithymia) implies that they may have only tapped into the cognitive or informational aspects of affective experience. It seems that this may not be sufficient if the goal is to fully simulate the affective experience inherent to real-life moral decision-making. In order to explore this possibility, it might be beneficial to manipulate the visceral aspects of emotional experience and explore the effect that this manipulation might have on individuals’ moral forecasts. Perhaps reproducing a physiological or bodily state of arousal in forecasters will help to fully attenuate the gap between moral behaviours and moral forecasts.

3.2 Tapping Into Discrete Emotions

The results of the current research make a strong case for the role of affective experience in moral decision-making. However, these data do not allow for claims about the nature of the emotions that might be responsible for relationship between moral behaviour and moral forecasting, leaving certain questions unanswered. For instance, are people misforecasting anxiety in the math task? Or are they misforecasting anticipated guilt? Given how difficult it is to accurately capture the nature of emotions via self-report, one possibility for future research is to explore other tools such as rating dials for measuring discrete emotions. Past work suggests that this may be a reliable method to measure emotions “in the moment” without altering emotional experience (Mauss, Levenson, McCarter, Wilhelm, & Gross, 2005). This method could also potentially help to answer whether it was the conscious experience of emotions that led to greater forecasting accuracy in Study 4. Coupled with psychophysiological measures, using a rating dial might make it possible to uncover whether inducing emotional salience
actually changes physiological expressions of emotion, or whether it simply provides individuals with emotional information.

3.3 The Boundary Conditions of Moral Underestimation

Another direction for future research is to examine the generalizability and boundary conditions of the moral underestimation effect. As previously mentioned, research suggests that situations low in affect and high in social distance could lead people to predict acting more morally than they actually do (Epley & Dunning, 2000); but what about other factors? Here, I used moral dilemmas that involved little potential gain for the participant. These dilemmas, in other words, allowed participants to avoid moral transgression rather easily because the profit gained by transgressing was not very high (up to $10). It would be useful to test whether the obtained results would hold consistent in moral dilemmas where the potential for profit is considerably higher, and where the emotions elicited by the potential reward might override any moral emotions that might arise.

4 Conclusion

The findings presented here document an interesting deficit in predicting moral behaviour, highlighting a potential problem with neglecting behavioural measures in research on morality – people are often unable to accurately report how they would behave in a moral dilemma. Importantly, this work suggests that this deficit might result from faulty affective forecasts and that there might be methods through which we can attenuate the behaviour-forecasting discrepancy. It seems that measuring moral behaviour has the potential to provide information that would be unattainable with self-report measures alone (Baumeister, Vohs, & Funder, 2007). The findings here suggest that there may be systematic and predictable biases in the way that people forecast their moral behaviour, whether they are responding to a self-report measure of morality or planning their day-to-day lives. A greater understanding of this process has the potential to lead people to make better forecasts by, for instance, making a greater effort to simulate the emotional experience that would characterize the actual situation. Importantly, however, these results suggest that efforts to simulate the affective nature of a situation can sometimes fall short, perhaps because these simulations are lacking the visceral aspects of experience. As a consequence, people’s forecasts may never be a perfect proxy for their behaviours.

American Heart Association. (2012a). *All about heart rate (pulse)*. Retrieved from http://www.heart.org/HEARTORG/Conditions/More/MyHeartandStrokeNews/All-About-Heart-Rate-Pulse_UCM_438850_Article.jsp


Copyright Acknowledgments

The content of Chapter 3 is currently published in Social Psychological and Personality Science:
doi: 10.1177/1948550610389338

The content of Chapter 4 is currently published in Psychological Science:
Notes

1. Two participants left the entire $10 behind because they did not realize that they were allowed to actually keep the money. Including these participants in our analysis, changed the significance value to $p = .083$.

2. All outliers across all studies were determined using the same criteria, namely the Grubbs’ ESD test. All excluded outliers were deemed statistically significant ($p < .05$), as determined by Grubbs’ ESD test.
Table 1. Bivariate correlations for main study variables.

<table>
<thead>
<tr>
<th></th>
<th>Cheating</th>
<th>PEP</th>
<th>RSA</th>
<th>SCR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cheating</strong></td>
<td>-</td>
<td>.22</td>
<td>-.44*</td>
<td>-.59**</td>
</tr>
<tr>
<td><strong>PEP</strong></td>
<td>-</td>
<td>-</td>
<td>-.26</td>
<td>-.09</td>
</tr>
<tr>
<td><strong>RSA</strong></td>
<td>-</td>
<td>-</td>
<td>.02</td>
<td>-</td>
</tr>
<tr>
<td><strong>SCR</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

** denotes \( p < .01 \)

* denotes \( p < .05 \)

Table 2. Indirect effects of action/forecasting on cheating through psychophysiological arousal.

<table>
<thead>
<tr>
<th>Mediator</th>
<th>SE</th>
<th>95% Bias Corrected CI Lower</th>
<th>95% Bias Corrected CI Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TOTAL</strong></td>
<td>2.65</td>
<td>.45</td>
<td>11.89*</td>
</tr>
<tr>
<td><strong>RSA</strong></td>
<td>1.05</td>
<td>.17</td>
<td>5.22*</td>
</tr>
<tr>
<td><strong>SCR</strong></td>
<td>2.15</td>
<td>-.11</td>
<td>8.66</td>
</tr>
</tbody>
</table>

* denotes \( p < .05 \) (significant indirect effect)