

**The role of unique personal representations in understanding and responding to the emotions of others**

by

Alicia J. Hofelich

A dissertation submitted in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
(Psychology)  
in The University of Michigan  
2012

Doctoral Committee:

Associate Professor Stephanie D. Preston, Chair  
Professor Patricia Reuter-Lornez  
Associate Professor Rachael Seidler  
Assistant Professor Nicholas Giardino

© Alicia J. Hofelich

2012

To Matthew

## **Acknowledgements**

I want to thank my committee for all your support and feedback (and the kind responses to all the emails I fill your inboxes with ☺). A special, heartfelt thank you goes to Stephanie – it’s been a long journey, and you have helped me so much along the way. I still don’t always see the forest through the (statistical) trees, but I have most definitely grown as both an academic and a person. Thank you.

Thank you to my fellow lab mates, Brian Vickers and Dan Porter, and all the wonderful undergraduates who have worked with me on these projects over the years – Mansi Mehta, Melissa Fasteau, Eugenia Chen, Kendall Arslanian, Khari Jones, Rebecca Semel, Courtney Hsing, Jamie Cohen – you guys are awesome. Thank you also to Ethan Kross and Daniel Weissman for your feedback on these projects and support.

A huge thanks to my family – especially my mom and my future husband Matthew Mohr– you have always been rocks in my life, and the past five years were no different. Thanks for always being there and for always believing in me. Thank you to my fellow graduate students and friends – you were always there to offer support, advice, empathy, and wine when I needed it. Thanks also to my cat, Marm. Although your insatiable love of power cords drives me crazy, you remind me to take joy in the little things in life, like breakfast, naps, and chipmunks.

## Table of Contents

<b>Dedication .....</b>	<b>ii</b>
<b>Acknowledgements .....</b>	<b>iii</b>
<b>List of Figures.....</b>	<b>v</b>
<b>List of Tables.....</b>	<b>vi</b>
<b>Abstract.....</b>	<b>vii</b>
<b>Chapter 1: General Introduction.....</b>	<b>1</b>
References .....	16
<b>Chapter 2: Implicit Emotion Associations Affect the Prosocial Response.....</b>	<b>23</b>
Abstract .....	23
Introduction.....	24
Methods.....	28
Results.....	32
Discussion .....	38
References.....	40
<b>Chapter 3: Priming Associations Affect the Prosocial Response.....</b>	<b>44</b>
Introduction.....	44
Methods.....	46
Results.....	50
Discussion .....	59
General Discussion .....	62
References.....	65
<b>Appendix 3.A .....</b>	<b>68</b>
<b>Appendix 3.B .....</b>	<b>76</b>
<b>Chapter 4: The Effects of Remitted and Current Depression on Empathy for Another's Distress .....</b>	<b>79</b>
Abstract .....	79
Introduction.....	80
Methods.....	87
Results.....	94
Discussion .....	109
Follow-up Study .....	117
References.....	123
<b>Appendix 4.A .....</b>	<b>130</b>
<b>Chapter 5: General Discussion .....</b>	<b>137</b>
References.....	146

## List of Figures

<b>Figure 2.1.</b> Scatterplot of individuals' IAT score by degree of preferential help offered to <i>distraught</i> over <i>resilient</i> patients with best-fit lines.....	37
<b>Figure 3.1.</b> Scatterplot of female-sadness associations and preferential giving to <i>distraught</i> over <i>resilient</i> female targets with best-fit line. ....	56
<b>Figure 3.2.</b> Scatterplot of female-stereotype associations and preferential giving to <i>distraught</i> over <i>resilient</i> female targets with best-fit line. ....	57
<b>Figure 4.1.</b> Histograms of BDI scores by group with standard cut-offs for depression severity .....	95
<b>Figure 4.2.</b> Change in agglomerative coefficient for Ward's clustering when additional groups are added. Displayed by depression history group. ....	97
<b>Figure 4.A.1.</b> Displayed emotions and emotional reactions to the patients in Preston et al., under review. ....	132

## List of Tables

<b>Table 1.1.</b> Definition of common empathy terms.....	2
<b>Table 2.1.</b> Description of the blocks for the Implicit Associations Test (IAT). Participants received either order A or B (counterbalanced). .....	29
<b>Table 2.2.</b> Mean responses (SE) to <i>distraught</i> and <i>resilient</i> targets by observer gender..	35
<b>Table 2.3.</b> Mean ratings by patient type, patient gender, and observer gender.....	54
<b>Table 3.A.1.</b> Mean ratings by display type and individual patient for the transcripts.....	70
<b>Table 3.A.2.</b> Mean ratings from Chapter 2 by display type and individual patient for the videos. ....	71
<b>Table 4.1.</b> Comparison of videos in clustering solutions by depression history and current depression. ....	97
<b>Table 4.2.</b> Means (SD) for variables by depression history group and cluster and beta values for the relationship with BDI (bold indicates significant cluster x current depression interactions) .....	104
<b>Table 4.A.1.</b> Mean (standard error) emotional reactions to the common target types in control and depressed participants in pilot study. ....	134

## **Abstract**

### **The role of unique personal representations in understanding and responding to the emotions of others**

**by**

**Alicia J. Hofelich**

**Chair: Stephanie D. Preston**

We encounter others in need everyday. Often we may feel a twinge of their suffering, and sometimes we offer help or support. However, not every individual in need elicits these feelings, and people vary widely in the types of targets they choose to help. According to the Perception-Action Model of empathy (PAM; Preston & de Waal, 2002), the neural substrates that give rise to first-hand experiences of an emotion also underlie the understanding of that emotion in another. Therefore, an observer's ability to understand and empathize with a target depends on the extent to which they have compatible representations of the target's state. This dissertation investigates these predicted interactions between the personal representations of observers and the various emotional displays of targets, focusing on how these interactions impact empathic, physiological, and prosocial responses. Chapter 2 and Chapter 3 demonstrate that individual differences in associations between females and sadness predict preferential giving to sad, distraught



female hospital patients compared to more positive resilient females. Chapter 4 examines empathic, prosocial, and physiological responses in females with and without a history of depression, determining whether observers who have experienced intense negative affect empathize and give the most help to patients who express the most negative emotion. Overall, the findings suggest emotion representations differ across individuals and these differences influence the types of targets people empathize with and help.

## Chapter 1: General Introduction

We encounter others in need everyday, from the homeless man on the street corner, to the hungry child on a television commercial, to the distressed student in our office. Often we may feel a twinge of their suffering, and sometimes we offer support—a dollar, a donation, a few reassuring words. However, not every individual we encounter elicits these feelings, and people vary widely in the types of targets they choose to help. For example, while the images of starving children may cause one person to pick up the phone and donate, it may cause another to become overwhelmed and turn the channel. Similarly, if a distressed student is quiet and unexpressive, some may think their need is nonexistent, while others may interpret their situation as quite serious. Understanding these individual differences in empathy is crucial for understanding the presence and absence of prosocial behavior in everyday life.

This dissertation presents three studies that investigate the bases of such individual differences in the empathic response. All look at the ways in which personal, experience-dependent representations of *observers* (people exposed to the need of others) affect empathy for and understanding of different people in need (referred to throughout as *targets*). Specifically, these studies address how individual differences in observers' representations of sadness predict responses to various targets, whose emotional displays range from overt distress to more positive resilience. Chapters 2 and 3 assess

continuously varying differences in observers' associations between females and sadness, and demonstrate that observers preferentially help targets who display emotions in ways that are consistent their associations. Chapter 4 investigates changes in representations of sadness across observers with past or current depression experience, and demonstrates that these experiences change both perceptions of and empathic responses to different types of targets. Before describing each study in more detail, I will begin with an overview of the relevant theories and literature motivating this dissertation.

### **Empathy and the Perception-Action Model**

Because the field of empathy is filled with inconsistent definitions for the words used to describe various phenomena, I want to first clarify the terms I will use throughout this dissertation (Table 1.1).

**Table 1.1. Definition of common empathy terms**

Term	Definition	Similar terms
Empathy	Broad term encompassing any kind of resonating or feeling “with” a target	
Affective Resonance	Feeling the same emotion as the target	Emotional contagion
Neural Resonance	Activating representations of the emotion perceived in the target	Self-other neural overlap
Sympathy	Tender feelings towards or for the other	Empathic concern, compassion
Personal Distress	Self-focused distress from viewing the distress of a target	

*Empathy* is used broadly to encompass any type of resonating or feeling “with” the target, which can be experienced as a subjective, affective state (i.e., feeling the same emotion as another; sometimes called *affective resonance* or *emotional contagion*) or can simply reside at the level of the mental or neural representation (i.e., activating personal representations of experienced emotion when perceiving and understanding the state of the target). Empathy can lead to *sympathy* or *empathic concern*, which involve tender, compassionate, other-oriented feelings, or to *personal distress*, which is a self-focused response to another’s need (Batson, Fultz, & Schoenrade, 1987; Eisenberg & Fabes, 1990; Preston & Hofelich, 2012). In terms of prosocial behavior, empathy and sympathy are thought to underlie altruistic responses to others, while intense feelings of personal distress can result in self-focus, which can lead to egoistically motivated or inhibited helping, especially when avoidance of the other is easy (Batson et al., 1987).

Many lay definitions of empathy highlight the importance of shared experience, requiring that an observer has previously experienced a similar situation as another, or has “walked a mile in their shoes,” in order to truly empathize. This is consistent with mechanistic understandings of empathy, as many theories propose that understanding another relies on the observer’s own representations of the target’s state. From a neuroscientific perspective, understanding another by using one’s own knowledge and personal representations of affective experiences is the proximate mechanism for empathy, and is how empathy is instantiated in the nervous system (e.g., Decety & Jackson, 2004; Preston & Hofelich, 2012; Preston & de Waal, 2002). The perception-action model of empathy (PAM; Preston & de Waal, 2002), which is the prominent model of empathy discussed here, assumes that when we perceive emotion in another, we

spontaneously activate our own neural representations of that emotion, situation, state, and person. This provides a direct mechanism for empathy and understanding that can also result in felt affective resonance depending upon the degree of similarity of our experiences, the salience of the other's emotion, or the attention paid to the other (Preston & Hofelich, 2012). Such a mechanism is based on principles of motor perception-action overlap, which involves similar neural coding between the execution and perception of an action (Preston & de Waal, 2002; Prinz, Heuer, & Sanders, 1987).

### **Evidence for the Perception-Action Model**

Perhaps the most notable evidence for perception-action mechanisms is the existence of “mirror neurons” located in frontal F5 and parietal regions in macaque monkeys (e.g., Gallese, Fadiga, Fogassi, & Rizzolatti, 1996). These neurons show mirror properties, responding both to the execution and observation of known motor actions or goals, which suggests a direct link between perception and action understanding. The mirror responses of these neurons are only observed during actions that have previously been performed, suggesting the response reflects recruitment of stored knowledge about the specific action. As a result, mirror neurons have been proposed to underlie action understanding, and even complex functions such as empathy and language acquisition (Gallese, 2001; Rizzolatti & Arbib, 1998).

However, the centrality of mirror neurons in these extended functions such as empathy and language has come under scrutiny (e.g., Hickok, 2009). Claims about mirror neurons have been difficult to test in humans, as most studies have used functional neuroimaging, which captures the combined activity of thousands of neurons, rather than

single cells. An exception is a recent recording of single-cell activity epileptic patients (Mukamel, Ekstrom, Kaplan, Iacoboni, & Fried, 2010), which found “mirror” neurons in the supplementary motor area (SMA) as well as in the hippocampus, parahippocampal gyrus, and entorhinal cortex. Similarly, human neuroimaging studies have found neural regions with mirroring properties outside of the areas where mirror neurons were initially identified in non-human primates, and in response to non-motor tasks. For example, regions of the insula are activated both during the feeling and observation of disgust (Wicker et al., 2003), and regions of the insula, anterior cingulate, and even somatosensory cortices are active during the observation and experience of pain (Cheng, Yang, Lin, Lee, & Decety, 2008; Jackson, Brunet, Meltzoff, & Decety, 2006; Lamm, Decety, & Singer, 2011; Morrison, Lloyd, Di Pellegrino, & Roberts, 2004; Singer & Frith, 2005). These data suggest the brain may employ a broad perception-action mechanism that involves the activation of distributed neural representations, which are not restricted to individual “mirror neurons” per se.

This distributed, neural-level overlap between the representations for perceiving a target’s state and experiencing the same state is the mechanism for empathy according to the PAM; however, it is important to note that the presence of such neural resonance does not necessitate subjective feelings of empathy (Preston & Hofelich, 2012). For example, some amount of neural resonance is thought to spontaneously occur when perceiving another, because of the way that the brain represents information (Preston & Hofelich, 2012; Preston & de Waal, 2002). In this context, representations are specific patterns of neural activity activated when a concept, experience, emotion, or situation is brought to mind, either by direct perception or from top-down generation. These representations are

similar to distributed perceptual knowledge representations (e.g., Barsalou, 1999), and include conceptual knowledge, motor and sensory information, memories, emotions, and physiological responses associated with an emotion or experience. However, activation of these representations does not require an observer to feel emotion or to demonstrate an overt motor response. Rather, these distributed components are what underlie our understanding of a concept.

For example, our knowledge of a hammer includes visual information about what a hammer looks like, likely stored in primary and associative visual regions (Roland & Gulyás, 1994); motor actions associated with a hammer, such as swinging or pounding, represented in the premotor cortex (Grafton, Fadiga, Arbib, & Rizzolatti, 1997), and semantically related words, like nail or screwdriver, which are primed when thinking about the word “hammer” (e.g., Meyer & Schvaneveldt, 1971). Together, this network of activated representations *is* our representation of “hammer” and allows us to understand what a hammer is. Similar distributed representations give rise to our understanding of an emotion or situation. For example, our understanding of concepts like “sadness” may include motor codes related to the formation of sad facial expressions (Lundqvist, 1995), action tendencies such as slowed walk or hunched posture (e.g., Duclos, Laird, Schneider, & Sexter, 1989), physiological bodily states related to sadness, and episodic memories of sad events. These representations are activated when we think about or experience that emotion, but are also spontaneously activated when we see the same emotion in another or imagine how they must be feeling.

### **Representations are shaped by experience**

Individuals have unique experiences and interactions with the world, leading to different representations of events or concepts, and therefore differences in how a target is represented and understood by a particular observer. For example, someone who passes a begging homeless man on the street corner each day may think of him as a nuisance, while a person who encounters begging for the first time may feel more alarmed or concerned. Experience shapes responses to others because it changes how a concept is represented and affects associations between concepts, situations, emotions, or people. To draw from the hammer example above, if someone had never used a hammer, but was familiar with music from the 1980's, they may more readily activate representations of musician M.C. Hammer, or his famous "Hammer Pants" instead of tools or nails when thinking about the word "hammer." Their associations between "hammer" and "pants" would then be stronger than their associations between "hammer" and "nail." Similarly, our ability to understand the emotion of another depends on personal experience. The richer our own representations of the other's state, the greater our capacity to have empathy for them. Someone who had actually lived on the streets, for example, may be even more likely to offer assistance to the begging homeless man.

Evidence for experience-dependent modulation in felt empathic and neural resonance has been found at many levels, from basic motor mirroring to altruistic giving. Some theories in the motor domain, for example, posit that automatic neural mirroring is completely experience dependent and is due to correlated experiences of executing and observing actions (e.g., Heyes, 2001). These theories have been supported with findings that action mirroring is dependent on the specific expertise of the observer (Calvo-Merino, Glaser, Grèzes, Passingham, & Haggard, 2005; Calvo-Merino, Grèzes, Glaser,



Passingham, & Haggard, 2006; Haslinger et al., 2005), and experimental inductions of “counter-imitation.” In these studies, participants are trained to execute opposite movements when observing an action and later show opposite motor mirroring of that action (Catmur, Walsh, & Heyes, 2007; Heyes, Bird, Johnson, & Haggard, 2005). In the empathy literature, familiarity and similar experience are among the strongest predictors of empathic concern and prosocial behavior (Barnett, 1984; Barnett, Tetreault, Esper, & Bristow, 1986; Batson, Eklund, Chermok, Hoyt, & Ortiz, 2007; Batson et al., 1996; Cialdini, Brown, Lewis, & Luce, 1997; Hodges & Wegner, 1997; Stotland, 1969). However, recent studies have shown that people with similar experiences, such as divorce or childbirth, are not necessarily more accurate at understanding the emotions of another (Hodges, Kiel, Kramer, Veach, & Villanueva, 2010). This suggests that experiences alone may not instill uniform representations across individuals. Rather, variations in the specific context and appraisals of the experience will produce differences in the representations associated with an individual’s understanding of that situation. For example, someone whose parents divorce after many years of tense fighting may see it as a welcome event, whereas someone who experiences their parents’ sudden divorce as emotionally shattering would understand it very differently. Therefore, the specific content of an observer’s representations is important in determining their understanding of and responses to others.

### **Cognitive versus emotional empathy**

It is important to note that the PAM is not the only way empathy is conceptualized in the literature. For example, many make the distinction between cognitive and

emotional empathy (e.g., Decety & Jackson, 2004; Nummenmaa, Hirvonen, Parkkola, & Hietanen, 2008; Shamay-Tsoory, Aharon-Peretz, & Perry, 2009; Zaki, Hennigan, Weber, & Ochsner, 2010; Zaki, Weber, Bolger, & Ochsner, 2009), where emotional empathy involves emotional resonance and more automatic responses to another (often the PAM is characterized as a mechanism for this), while cognitive empathy is more related to perspective taking and theory of mind. Because of the dynamic and distributed nature of the brain and representations (e.g., Preston & Hofelich, 2012), it is not often useful to impose strict segregation between “cognitive” and “affective” processes or components. For example, even more “cognitive” attempts at mentalizing can activate “affective” regions and produce emotional responses (Schnell, Bluschke, Konradt, & Walter, 2011), and spontaneous reactions to emotional faces also include semantic representations of emotion (Hofelich & Preston, 2011). Because of the distributed and dynamic make up of the brain, it is not likely that the any process remains purely affective or cognitive, regardless of the initial mechanism.

Rather, it is more useful to make distinctions between the processes through which the representations are activated, whether through more spontaneous bottom-up processes in response to a salient target, or through more effortful top-down processes. According to the PAM, levels of attention, engagement with the other, and similar experience can affect the extent to which people need to engage in more effortful perspective taking in order to understand the other (Preston & Hofelich, 2012). Because prior experience forms richer representations of a situation, the PAM predicts observers with similar experience have greater opportunity for spontaneous neural self-other overlap (activating more of their own representations when seeing that state in another).

This often leads to a better understanding of another's state and increased empathy. However, observers without such experience can still sympathize with another's state, by drawing on their own representations of sadness, which may include different personal situations or memories. Neural self-other overlap is still involved in this case, but is restricted to more abstract or general understandings of their emotion (i.e., that they are sad; Preston & Hofelich, 2012). Therefore, not having an experience does not preclude empathy or sympathy, if other relevant representations are activated enough to give an observer an idea of what the other is feeling, but the relative empathic response should be stronger with increasing correspondence between the observer's representations and the target's state. The studies presented here focus on this specific prediction of the PAM.

It is also important to note that while effortful, top-down processes can increase attention and feelings of empathy when an observer does not have similar representations, they can also serve to limit empathy when observers have a competing goal state or when help is costly (Preston & de Waal, 2002). Competing goal states, such as induced competition with the target (Lanzetta & Englis, 1989) or perceptions that the target behaved unfairly (Singer et al., 2006), can result in reduced or even counter-empathic responses. The PAM predicts that if a competing goal state is present, cognitive processes can be engaged to direct attention away from the target and suppress empathic processing (Preston & de Waal, 2002). To limit the extent to which this occurs, the current research does not induce competing goal states, and although these studies measure help that comes at some cost to observers (i.e., losing money), the cost of helping is not extreme.

## **The current research**

Since the representations activated depend specifically on the emotion, situation, or target that is observed, the best way to study these individual differences in empathy is by using a wide range of targets. This is in contrast to the majority of empathy studies, which typically employ one or two prototypical targets in obvious need, and are designed to elicit personal distress or empathic concern (e.g., Coke, Batson, & McDavis, 1978; Eisenberg & Fabes, 1990). Such paradigms have served well to isolate specific situational or personality factors that influence empathy, but often mask important interactions than can arise between the experience of the observer and the emotional display of the target. Such interactions abound in real-world situations, as people display their need in various, and sometimes divergent, ways, which elicit wide-ranging responses from observers. The studies presented in this dissertation focus on the ways in which differences in personal representations result in specific observer-target interactions, which have been largely unstudied in the empathy literature. Personal representations can vary across individuals in different ways, through differences in experience or in mental associations. People may differ in the extent to which they associate certain concepts; for example, knowing whether someone more strongly associates “hammer” with “nail” or with “pants” provides information about that individual’s representation of hammer. People can also vary in the richness of their representations, with specific experiences providing more extensive representations of that state (e.g., Preston & Hofelich, 2012). This may occur through the formation new episodic memories, the strengthening of associations between feelings and situations, or the recruitment of additional cortical space devoted to an action or sensation.

The work presented in this dissertation is motivated by findings of a previous study done in collaboration with Stephanie Preston. In this research, we used videos of real hospital patients suffering from serious chronic or terminal illnesses as targets, and found that despite being in similar situations, these targets responded to their need in a wide variety of ways (Preston et al., under review). The emotions expressed by the fourteen hospital patients were characterized by participants' ratings. Using mathematical clustering analyses, the patients were grouped into similar emotional "types" based on the emotions they expressed. These display types included *distraught* patients who overtly displayed distress and sadness, *resilient* patients who displayed positive emotion and resilience, *sanguine* patients who made jokes and were overly positive, a *reticent* patient who did not show any emotion, and sad but quiet *wistful* patients. Interestingly, these different display types also evoked various reactions in observers, with the *distraught* and *resilient* patients eliciting the most divergent responses. Some participants offered the most help to *distraught* patients and others to *resilient* patients despite the fact that everyone rated the *distraught* women as needing more help (Preston et al., under review). One of the goals of this dissertation is to uncover factors that can explain these divergent responses to *distraught* and *resilient* targets.

Chapters 2 and 3 of this dissertation examine how learned implicit associations between emotion and gender can affect prosocial responses to different types of targets. Such associations are formed through social and cultural experience, but also can vary widely across individuals depending on their own experiences. These chapters will be submitted together as one study with two experiments and is co-authored with Stephanie Preston. Therefore, a single abstract is presented in Chapter 2 for both chapters and a

general discussion in Chapter 3 will summarize both experiments. In Chapter 2, participants completed an Implicit Associations Test (IAT; Greenwald, McGhee, & Schwartz, 1998) that measured the extent to which they associated females (compared to males) with sadness (compared to anger). They then watched videos of female hospital patients who had been previously classified as *distraught* or *resilient*, and rated their emotional responses, empathy, willingness to help and the amount of time or money they were willing to donate to each patient. People with stronger associations between females and sadness offered more practical support and donated more money to the sad, *distraught* female patients, compared to the more positive *resilient* female patients, suggesting that observers are more likely to help others who express their need in ways that conform with their expectations.

In Chapter 3, participants completed a different measure of implicit associations. A priming version of the lexical decision task (e.g., Dijksterhuis, Aarts, Bargh, & van Knippenberg, 2000) was used to specifically assess participants' associations between females and sadness, males and sadness, as well as female and male stereotypical associations. Participants then read video transcripts from *distraught* and *resilient* patients, with the gender of the patient randomized to measure responses to both female and male patients. Male and female patients of each type were rated similarly, and participants with strong female-sadness associations again donated more money to *distraught* female patients over *resilient* female patients. Chapter 3 demonstrates this effect is specific to affective associations, rather than more general associations about the target or about who is more deserving of need (i.e., people who are distressed), as female-

stereotype associations did not predict donation behaviors and female-sadness associations did not predict donations to *distraught* male patients.

Chapter 4 of the dissertation also examines how differences in observers' representations influence responses to various targets, but instead of using a continuous measure of association across participants as in the previous chapters, observers were selected who either had or did not have past experiences of intense negative affect. Because past life experiences can differ widely in their context and appraisals across individuals, the past history of an affective experience like depression was chosen in attempt to minimize these differences. Although the specific presentation of depression can vary widely, at the core of all depressive experiences is an intense negative state. Additionally, initial pilot data with depressed women suggested that empathy was higher for negative, *distraught* targets, whose displayed affect was congruent with the prior emotional experiences of those observers (Appendix 4.A). Chapter 4 tested this in a larger sample of observers, who varied in both past and current experiences of depression, as assessed by diagnosis history and current scores on depression scales. The findings suggest that observers with a depression history give more help to targets they rate as having similar emotional and life experiences, although these were not always the *distraught* patients as initially expected. Unexpectedly, participants who were currently the most depressed gave the least to *distraught* targets, perceiving them to be less in need of help than non-depressed observers.

Together the studies presented in this dissertation demonstrate the importance of considering the individual attributes of both targets and observers in the empathic response. The use of naturalistic targets of need highlight the fact that people respond

differently even when in very similar situations, such as being hospitalized for chronic or terminal illness. These subtle differences are not captured when responses are only studied from a single target of need or from several artificial targets that vary only on certain specific dimensions. Beyond demonstrating that various targets elicit divergent responses in observers, this work also suggests that the extent of these effects are predicted by individual differences in the representations of observers. Therefore, it is critical to account for both the displays of the target and the individual representations of observers in order to begin to fully understand what contributes to the presence and absence of empathy.



## References

- Barnett, M. A. (1984). Similarity of experience and empathy in preschoolers. *Journal of Genetic Psychology, 145*(2), 241-250.
- Barnett, M. A., Tetreault, P. A., Esper, J. A., & Bristow, A. R. (1986). Similarity and empathy: The experience of rape. *Journal of Social Psychology, 126*, 47-50.
- Barsalou, L. W. (1999). Perceptual symbol systems. *The Behavioral and Brain Sciences, 22*(4), 577-609.
- Batson, C. D., Eklund, J. H., Chermok, V. L., Hoyt, J. L., & Ortiz, B. G. (2007). An additional antecedent of empathic concern: Valuing the welfare of the person in need. *Journal of Personality and Social Psychology, 93*(1), 65-74.  
doi:10.1037/0022-3514.93.1.65
- Batson, C. D., Fultz, J., & Schoenrade, P. A. (1987). Distress and empathy: Two qualitatively distinct vicarious emotions with different motivational consequences. *Journal of Personality, 55*(1), 19-39.
- Batson, C. D., Sympson, S. C., Hindman, J. L., Decruz, P., Todd, R. M., Weeks, J. L., Jennings, G., et al. (1996). "I've Been there, Too": Effect on empathy of prior experience with a need. *Personality and Social Psychology Bulletin, 22*(5), 474-482.  
doi:10.1177/0146167296225005
- Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: An FMRI study with expert dancers. *Cerebral Cortex, 15*(8), 1243-9. doi:10.1093/cercor/bhi007

- Calvo-Merino, B., Grèzes, J., Glaser, D. E., Passingham, R. E., & Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Current Biology*, 16(19), 1905-10. doi:10.1016/j.cub.2006.07.065
- Catmur, C., Walsh, V., & Heyes, C. (2007). Sensorimotor learning configures the human mirror system. *Current biology : CB*, 17(17), 1527-31. doi:10.1016/j.cub.2007.08.006
- Cheng, Y., Yang, C.-Y., Lin, C.-P., Lee, P.-L., & Decety, J. (2008). The perception of pain in others suppresses somatosensory oscillations: a magnetoencephalography study. *NeuroImage*, 40(4), 1833-40. doi:10.1016/j.neuroimage.2008.01.064
- Cialdini, R. B., Brown, S. L., Lewis, B. P., & Luce, C. (1997). Reinterpreting the empathy-altruism relationship: When one into one equals oneness. *Journal of Personality and Social Psychology*, 73(3), 481-494.
- Coke, J. S., Batson, C. D., & McDavis, K. (1978). Empathic mediation of helping: A two-stage model. *Journal of Personality and Social Psychology*, 36(7), 752-766. doi:10.1037//0022-3514.36.7.752
- Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews*, 3(2), 71-100. doi:10.1177/1534582304267187
- Dijksterhuis, A., Aarts, H., Bargh, J. A., & van Knippenberg, A. (2000). On the relation between associative strength and automatic behavior. *Journal of Experimental Social Psychology*, 36(5), 531-544. doi:10.1006/jesp.2000.1427

- Duclos, S. E., Laird, J. D., Schneider, E., & Sexter, M. (1989). Emotion-specific effects of facial expressions and postures on emotional experience. *Journal of Personality and Social Psychology*, 57(1), 100-108. doi:10.1037//0022-3514.57.1.100
- Eisenberg, N., & Fabes, R. A. (1990). Empathy: Conceptualization, measurement, and relation to prosocial behavior. *Motivation and Emotion*, 14(2), 131-149.
- Gallese, V. (2001). The “shared manifold” hypothesis: From mirror neurons to empathy. *Journal of Consciousness Studies*, 8(5-7), 33-50.
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain: A Journal of Neurology*, 119(2), 593-609.
- Grafton, S. T., Fadiga, L., Arbib, M. A., & Rizzolatti, G. (1997). Premotor cortex activation during observation and naming of familiar tools. *NeuroImage*, 6(4), 231-236. doi:10.1006/nimg.1997.0293
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6), 1464-1480.
- Haslinger, B., Erhard, P., Altenmüller, E., Schroeder, U., Boecker, H., & Ceballos-Baumann, a O. (2005). Transmodal sensorimotor networks during action observation in professional pianists. *Journal of Cognitive Neuroscience*, 17(2), 282-93. doi:10.1162/0898929053124893
- Heyes, C. (2001). Causes and consequences of imitation. *Trends in Cognitive Sciences*, 5(6), 253-261.

- Heyes, C., Bird, G., Johnson, H., & Haggard, P. (2005). Experience modulates automatic imitation. *Cognitive Brain Research*, 22(2), 233-240.  
doi:10.1016/j.cogbrainres.2004.09.009
- Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience*, 21(7), 1229-43.  
doi:10.1162/jocn.2009.21189
- Hodges, S. D., & Wegner, D. M. (1997). Automatic and controlled empathy. *Empathic Accuracy* (pp. 311-340).
- Hodges, S. D., Kiel, K. J., Kramer, A. D. I., Veach, D., & Villanueva, B. R. (2010). Giving birth to empathy: the effects of similar experience on empathic accuracy, empathic concern, and perceived empathy. *Personality & Social Psychology Bulletin*, 36(3), 398-409. doi:10.1177/0146167209350326
- Hofelich, A. J., & Preston, S. D. (2011). The meaning in empathy: distinguishing conceptual encoding from facial mimicry, trait empathy, and attention to emotion. *Cognition & Emotion*, 26(1), 119-28. doi:10.1080/02699931.2011.559192
- Jackson, P. L., Brunet, E., Meltzoff, A. N., & Decety, J. (2006). Empathy examined through the neural mechanisms involved in imagining how I feel versus how you feel pain. *Neuropsychologia*, 44(5), 752-61.  
doi:10.1016/j.neuropsychologia.2005.07.015
- Lamm, C., Decety, J., & Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *NeuroImage*, 54(3), 2492-502. doi:10.1016/j.neuroimage.2010.10.014

- Lanzetta, J. T. & Englis, B. G. (1989). Expectations of cooperation and competition and their effects on observer's vicarious emotional responses. *Journal of Personality and Social Psychology*, 56(4), 543-554.
- Lundqvist, L. O. (1995). Facial EMG reactions to facial expressions: A case of facial emotional contagion? *Scandinavian Journal of Psychology*, 36(2), 130-141.
- Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 90(2), 227-234.
- Morrison, I., Lloyd, D., Di Pellegrino, G., & Roberts, N. (2004). Vicarious responses to pain in anterior cingulate cortex: is empathy a multisensory issue? *Cognitive, Affective, & Behavioral Neuroscience*, 4(2), 270-278.
- Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., & Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Current Biology*, 20(8), 750-756. doi:10.1016/j.cub.2010.02.045
- Nummenmaa, L., Hirvonen, J., Parkkola, R., & Hietanen, J. K. (2008). Is emotional contagion special? An fMRI study on neural systems for affective and cognitive empathy. *NeuroImage*, 43(3), 571-80. doi:10.1016/j.neuroimage.2008.08.014
- Preston, S. D., & Hofelich, A. J. (2012). The many faces of empathy: Parsing empathic phenomena through a proximate, dynamic-systems view of representing the other in the self. *Emotion Review*, 4(1), 24-35. doi:10.1177/1754073911421378
- Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *The Behavioral and brain sciences*, 25(1), 1-20; discussion 20-71.

- Preston, S. D., Hofelich, A. J., & Stansfield, R. B. (under review). The ethology of empathy: Emotions displayed and elicited by real targets of need.
- Prinz, W., Heuer, H., & Sanders, A. F. (1987). *Ideomotor action*. (H. Heuer & A. F. Sanders, Eds.) *Perspectives on Perception and Action* (pp. 47-76). Erlbaum.
- Rizzolatti, Giacomo, & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*, 21(5), 188-194. doi:10.1016/S0166-2236(98)01260-0
- Roland, P. E., & Gulyás, B. (1994). Visual imagery and visual representation. *Trends in Neurosciences*, 17(7), 281-287.
- Schnell, K., Bluschke, S., Konradt, B., & Walter, H. (2011). Functional relations of empathy and mentalizing: an fMRI study on the neural basis of cognitive empathy. *NeuroImage*, 54(2), 1743-1754.
- Shamay-Tsoory, S. G., Aharon-Peretz, J., & Perry, D. (2009). Two systems for empathy: a double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain : A Journal of Neurology*, 132(3), 617-27. doi:10.1093/brain/awn279
- Singer, T., & Frith, C. (2005). The painful side of empathy. *Nature Neuroscience*, 8(7), 845-846. doi:10.1093/brain/awh500
- Singer, T., Seymour, B., O'Doherty, J. P., Stephan, K. E., Dolan, R. J., & Frith, C. D. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature*, 439(26), 466-469. doi: 10.1038/nature04271.
- Stotland, E. (1969). Exploratory investigations of empathy. *Advances in Experimental Social Psychology*, 4, 271-313.

- Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., & Rizzolatti, G. (2003). Both of us disgusted in my insula: The common neural basis of seeing and feeling disgust. *Neuron*, 40(3), 655-64. doi:10.1016/S0896-6273(03)00679-2
- Zaki, J., Hennigan, K., Weber, J., & Ochsner, K. N. (2010). Social cognitive conflict resolution: Contributions of domain-general and domain-specific neural systems. *The Journal of Neuroscience*, 30(25), 8481-8. doi:10.1523/JNEUROSCI.0382-10.2010
- Zaki, J., Weber, J., Bolger, N., & Ochsner, K. (2009). The neural bases of empathic accuracy. *Proceedings of the National Academy of Sciences*, 106(27), 11382-7. doi:10.1073/pnas.0902666106

## Chapter 2: Implicit Emotion Associations Affect the Prosocial Response

### Abstract

Emotional expressions like sadness and crying are thought to signal need and adaptively motivate others to help. Such clear displays often do elicit empathy and prosocial behavior, but can also elicit personal distress, aggression, and even horror. The perception-action model (PAM) assumes that such varying responses reflect individuals' differing internal representations for affective states and situations derived from prior experience in life, family, and culture. This assumption was tested across two studies by predicting the prosocial response to sad (*distraught*) versus more positive (*resilient*) female targets through the strength of individual associations between females and sadness. Participant associations were measured with two different methods: a modified Implicit Associations Test (IAT) that measured female-sadness and male-anger associations (Chapter 2) and a priming task that measured associations between gender and sadness as well as gender role stereotypes (Chapter 3). Across studies, participants with stronger female-sadness associations offered more help to *distraught* over *resilient* females, an effect that was specific to the affective associations and not due to more general stereotyping. People appear to behave more prosocially towards those who express emotion in ways they believe are normative, but these beliefs differ across individuals, with profound implications for interpersonal exchanges.



## **Introduction**

Displays of distress, such as sadness and crying, are thought to have evolved to ensure rapid responding to helpless and vulnerable offspring (e.g., Eibl-Eibesfeldt, 1972; Marsh, Kozak, & Ambady, 2007; Preston & Hofelich, 2012). While these displays can engender empathy and helping among unrelated adults, they can also elicit negative responses that inhibit prosocial behavior, such as personal distress (Batson, Fultz, & Schoenrade, 1987), aggression (Zahn-Waxler, Radke-Yarrow, Wagner, & Chapman, 1992), and even anger and horror (Preston, Hofelich, & Stansfield, under review).

How can such diverging responses to distress be explained? Unsympathetic responses may adaptively prevent exploitation (Axelrod & Hamilton, 1981), and surely result when observers cannot regulate their own emotion in the face of distress (e.g., Eisenberg, Wentzel, & Harris, 1998). However, the perception-action model of empathy (PAM) additionally predicts that responses differ as a function of observers' internal representations for the state, because the PAM assumes that others' states are understood through partial activation of the observers' own neural representations for producing and experiencing those states (Preston & Hofelich, 2012; Preston & de Waal, 2002; Preston et al., 2007). According to this model, people should be more empathic and altruistic towards others that are similar, familiar, or experiencing states that are relatable to the observer from prior experience. Evidence for this assumption is often found (reviewed in Preston & de Waal, 2002), but not always. For example, sometimes observers who have experienced a similar event to the target (like divorce or childbirth) do not show enhanced understanding for how the other feels (Hodges, Kiel, Kramer, Veatch, & Villanueva, 2010).

Presumably, we don't always have greater empathy or understanding for similar experiences since people can appraise and respond to even the same situation in very different ways. For example, women who were distressed and scared during childbirth would not necessarily understand or empathize with a woman who went through the process with a "stiff upper lip" or a serene smile. Thus, it is important for an empathic exchange that the target and observer not only have a shared event in common, but that the actual content of their representation for the state and situation overlap—that their subjective experience of the event was similar. However, this more precise prediction of the PAM has yet to be tested, allowing controversy to persist over the importance of prior experience for empathy and helping (Batson et al., 1996; Hodges et al., 2010; Preston & Hofelich, 2012).

One way to examine person-specific representations of emotion is to use known variation in the association between gender and sadness. Sadness often occurs in situations of need, but the response to open sadness or crying is highly variable across individuals and cultures. For example, Americans were surprised to see Russian President-elect Vladimir Putin shedding a tear while his display elicited little reaction from Russians, who consider "sparse masculine tears" appropriate, if not expected, from stoic men in their culture (Rath, 2012). Even within American democrats, individuals responded differently to the tears of Hillary Clinton in the 2008 presidential primaries, with some finally sympathizing with her after showing such a vulnerable, human response and others deriding her for losing control, presumably due to different beliefs about what is permissible for females or leaders.

In general, Americans appear to believe that overt sadness is more permissible for

females and openly discourage such displays in even very young males (Brody & Hall, 2000). People also believe that females are more emotional and sad than males, despite the fact that they appear to experience similar levels in daily life (Barrett & Bliss-Moreau, 2009; Barrett, Robin, Pietromonaco, & Eyssell, 1998). They also believe that females express sadness more often (Plant, Hyde, Keltner, & Devine, 2000) and they are more likely to perceive sadness in females than in males (Birnbaum, Nosanchuk, & Croll, 1980; Hess, Adams, & Kleck, 2008; Parmley & Cunningham, 2008; Plant, Kling, & Smith, 2004).

In our own research, we have found that patients respond to serious chronic or terminal illness in a wide variety of ways, from not showing any emotion (*reticent*) to overtly displaying of distress and sadness (*distraught*) or displaying highly positive emotion and resilience (*resilient*). These displays also produce different responses in observers, with some offering the most help to *distraught* patients and others to *resilient* patients despite the fact that everyone rates the *distraught* women as needing more help (Preston et al., under review). One of the goals of the current study is to determine whether individual differences in gender-emotion associations could explain these divergent responses to *distraught* and *resilient* patients. According to the PAM, those who associate women with expressed sadness may perceive the *distraught* affect as more normative, and thus offer those women more empathy and help.

However, this hypothesis needs to be experimentally tested, not only as key support for a specific tenant of the PAM and as an explanation for observers' divergent giving, but also because research and theory in social psychology may suggest the opposite. Those with strong female-sadness associations may perceive female sadness as

a dispositional trait (Barrett & Bliss-moreau, 2009), rather than a response to their current situation. If people assume sadness is characteristic of females, they might attribute the distress displayed by *distraught* female patients to their trait tendencies to be sad, rather than seeing their distress as a situational response evoked by their current need.

According to this prediction, those with strong female-sadness associations may actually have reduced empathy and helping for *distraught* females because they would perceive them as less in need (e.g., Betancourt, 1990; Weiner, 1980).

Two studies tested the hypothesis that individual variation in the association between women and sadness influences the prosocial response: Chapter 2 measured the strength of each participant's female-sadness association (contrasted against male-anger) using a modified Implicit Associations Test (IAT; Greenwald, McGhee, & Schwartz, 1998). The IAT measures differential associations between two concepts and two attributes using reaction time (RT) to categorize exemplars of each category when the response mappings are compatible with the predicted association ("female" concept and "sadness" attribute are mapped on to one button; "male" concept and "anger" attribute are mapped on to the second button) compared to when mappings are incompatible ("female" and "anger" on one button; "male" and "sadness" on the other). Chapter 3 used a modified priming task (e.g., Dijksterhuis, Aarts, Bargh, & van Knippenberg, 2000) to measure the association between each gender and sadness separately (i.e., without confounding female-sadness association scores with male-anger or lacking female-anger or male-sadness association). Chapter 3 also ruled out the alternative hypothesis that results in Chapter 2 emanated from more general stereotypes about women or need and not gender-affect associations *per se*. In both studies, the measure of individual gender-

affect association was followed by observation of videotaped female targets, two *distraught* and two *resilient* (Preston et al., under review). Female-sadness associations were then used to predict the prosocial response to *distraught* over *resilient* targets.

## **Methods**

### **Participants**

Fifty-seven students (30 women and 27 men, mean age = 19.2, range 18-27) from the University of Michigan participated for course credit. All participants were consented and tested individually in the laboratory using Eprime Version 2.0 (Psychology Software Tools, Inc., Pittsburgh, PA) on a Dell desktop PC. All procedures were approved by the Institutional Review Board of the University of Michigan.

### **IAT**

In the IAT, which was always performed first, participants categorized 10 names (from Sriram & Greenwald, 2009) as either female (*Amy, Lisa, Sarah, Diana, Kate*) or male (*John, Paul, Mike, Kevin, Steve*) and categorized emotion adjectives (from Shaver, Schwartz, Kirson, & O'Connor, 1987) as sad (*depressed, hopeless, grieving, sorrowful, miserable*) or angry (*enraged, furious, hostile, hateful, and scornful*). Participants pressed the button corresponding to the correct category for each exemplar as quickly and accurately as possible using the left and right index fingers. As in the standard IAT task (Greenwald et al., 1998), the stimuli, responses, and button mappings changed across five blocks (see Table 2.1). Participants first categorized the exemplars of the target alone (Block 1: “female” versus “male” names), then the attributes alone (Block 2: “sad” versus “angry” adjectives), and then received either the compatible (“female” and “sadness” on

the same button) or incompatible (“female” and “anger” on the same button) blocks where they categorized exemplars from each of the four categories (“female”, “male”, “sad”, “angry”). They categorized the target exemplars alone again with the reversed response mappings before completing the second combined block (incompatible or compatible). The order of the blocks were counterbalanced such that half the participants received the compatible block first and half received the incompatible block first.

**Table 2.1. Description of the blocks for the Implicit Associations Test (IAT). Participants received either order A or B (counterbalanced).**

Order	Block	Description	# of Trials	Response Label	Response Key
A	1	Compatible Associations	20	Female	p
				Male	q
	2		20	Sadness	p
				Anger	q
	3		20	Female/Sadness	p
	4	Incompatible Associations	20	Male/Anger	q
				Male	p
				Female	q
	5		20	Male/Sadness	p
				Female/Anger	q
B	1	Incompatible Associations	20	Male	p
				Female	q
	2		20	Sadness	p
				Anger	q
	3		20	Male/Sadness	p
	4	Compatible Associations	20	Female/Anger	q
				Female	p
				Male	q
	5		20	Female/Sadness	p
				Male/Anger	q

For analysis, incorrect trials and those with reaction times (RT) greater than three standard deviations (SD) from each participant’s mean were removed. To demonstrate the general-affect association, median RTs between compatible and incompatible blocks

were contrasted using a paired t-test; gender differences were examined using univariate-ANOVA. To test the hypothesis that stronger gender-affect associations would predict more prosocial responding to the *distraught* patients, IAT scores per participant were computed (median RT incompatible - compatible block) and correlated with prosocial difference scores for *distraught* minus *resilient* targets across helping measures (described in detail subsequently).

### **Prosocial Response Task**

All participants watched four short target videos of female patients (order randomized) with serious or terminal illness who answered the same four quality-of-life questions. Two of the targets were previously classified by observers as *distraught* (high distress, low positive emotion, crying) and two were *resilient* (low distress, high positive emotion, smiling) (see Preston et al., under review). Two targets were included for each type to increase validity, since these real patients differed on variables other than expressed affect (e.g., weight, appearance, age) that would confound responses if only two patients were compared.

Participants rated targets on a variety of dimensions after each video on seven-point Likert scales from 1 (not at all) to 7 (extremely). As a manipulation check, targets were directly rated as being *distraught* and *resilient*. Then, to narrow the source of the prosocial response to each type of patient, observers rated how sick the patient seemed, how much they liked them, and the emotion factors for self (the observer's felt emotion) and other (the target's displayed emotion) from the prior study (target emotion factors: softhearted, distressed, happy; observer emotion factors: empathic, distressed, pleased, and horrified; see footnote in Table 2.2 for definitions of the terms given to participants).

To assess altruism, participants self-reported how much support they would give each patient on an emotional (e.g., talking to them, giving advice, soothing, spending time with them) or practical (e.g., getting prescriptions, changing sheets, watering plants, grocery shopping) level. Participants additionally rated how much emotional and practical support each patient needed. We also included two real measures of giving to increase validity and avoid common floor and ceiling effects. Participants were given nine tokens per patient that they could donate to each patient, knowing that undonated tokens would later be converted to cash to take home (via an unknown exchange rate). After the experiment ended, participants were given a letter requesting help from each patient with a project to promote awareness for their specific illness. Participants then indicated their willingness to volunteer time to complete a letter mailing task on behalf of each patient (in units of: 0, .5, 1, 2, 3, or 4 hours; after Penner, Fritzsche, Craiger, & Freifeld, 1995). All participants were fully debriefed to explain that the letter was not real and that no one would contact them for assistance, but their monetary donations would be given to a related charity. In addition to receiving course credit for their participation, participants were given the money from any tokens they did not donate.

### **Trait Scales**

Before receiving the letter helping measure, participants completed the Mehrabian and Epstein Scale of Emotional Empathy (ME; Mehrabian & Epstein, 1972) and the Prosocial Personality Battery (PSB; Penner et al., 1995). Individual scores were computed for the ME and for the two relevant subscales of the PSB, Other-oriented Empathy and Helpfulness.



## Statistical Analyses

All measures were averaged between the two targets within each type (*distraught*, *resilient*) and were compared using mixed-ANOVA with participant gender as a between-subjects factor. We first confirmed emotional differences between the types of patients and then examined whether participants felt more empathy and offered more help to one target type over the other. IAT scores and the trait scales were correlated with emotional and prosocial responses to the targets. Alpha was set at 0.05 for all comparisons.

## Results

### IAT

Participants were highly accurate in this task, with an average error rate of 4%. After error trials were removed, participants responded faster to compatible (male-anger; female-sad;  $M = 661.0$  ms) than incompatible word pairings (male-sad; female-anger;  $M = 735.4$  ms),  $t(56) = 4.5$ ,  $p < .001$ , demonstrating an implicit female-sadness/male-anger association which was not affected by participant gender,  $t(55) = 0.43$ ,  $p = .66$ , or block order,  $t(55) = .86$ ,  $p = .40$ . Because the compatible and incompatible RTs included both female-sadness and male-anger combinations, the IAT block RTs were decomposed to determine the specific contribution of sad-female versus anger-female associations. To do this, compatible RTs were calculated from only the compatible block trials where a female name or sad word was presented, and incompatible RTs were calculated from trials in the incompatible block where a female name or angry word was presented. Participants responded faster to female-sad compared to female-anger combinations,

$t(55) = 3.66, p < .001$ , suggesting the IAT score was not solely driven by strong male-anger associations or a lacking male-sad association. This score also did not differ for by participant gender,  $t(53.45) = 1.01, p = .32$ .

To generate the individual total IAT scores, each participant's RT on the compatible block was subtracted from their RT on the incompatible block, with positive scores indicating an implicit female-sadness/male-anger association. Overall, this association was prevalent, as 43 participants (22 f) had positive IAT scores and only 13 had negative IAT scores (8 f).

### **Perception of *distraught* and *resilient* targets**

Before collapsing across similar display types, the two videos in each type were compared to determine whether they were perceived similarly. The two *distraught* females were perceived as equally distraught,  $t(54) = -0.83, p = .41$ , although one patient was seen as more distressed than the other,  $t(54) = -2.74, p = .008$ . One *resilient* patient was perceived as marginally more resilient than the other,  $t(54) = -1.91, p = .061$ , but both were perceived as equally happy,  $t(53) = 0.36, p = .72$ .

Table 2.2 summarizes mean responses per dependent measure, by participant gender and patient type. Confirming that the distraught and resilient targets were perceived in ways that were convergent with their a priori classification (and despite the slight differences between the two patients within each type), *distraught* targets appeared more distraught and distressed than *resilient* targets,  $F(1, 51) > 417.04, p < .001$ , and *resilient* targets appeared more resilient, softhearted, and happy,  $F(1, 51) > 47.33, p < .001$ ; Table 2.2. *Distraught* targets were also seen as more sick,  $F(1, 51) = 188.88, p <$

.001, and as needing more practical and emotional support than *resilient* targets,  $F(1,54) > 125.0, p < .001$ . Females and males rated targets similarly overall, main effects of gender:  $F(1, 51) < .858, p > .36$ , but females rated *resilient* targets as more resilient and *distraught* targets as less resilient than males, gender-by-type interaction for resilient,  $F(1, 51) = 13.09, p = .001$ ; all others,  $F(1, 51) < 2.74, p > .10$ .

Observers felt more empathic towards, but also more horrified and distressed by, *distraught* than *resilient* targets,  $F(1, 51) > 23.95, p < .001$ , while feeling more pleased by and liking *resilient* targets more,  $F(1, 51) > 33.24, p < .001$ . Overall, female observers felt more empathy,  $F(1, 51) = 6.03, p = .02$ , and liked both targets more than males,  $F(1, 51) = 4.54, p = .04$ ; all other main effects of gender:  $F(1, 51) < 1.24, p > .27$ . Females also felt more distress towards *distraught* targets than males, but felt similarly towards *resilient* targets as males, gender-by-type interaction for distressed:  $F(1, 51) = 4.05, p = .049$ ; all other interactions,  $F(1,51) < 3.36, p > .07$ .

### **Prosocial Response Task**

Participants offered more emotional and practical support, and donated more money to *distraught* compared to *resilient* targets,  $F_s(1, 51) > 15.85, p_s < .001$ ; time donations  $F(1, 51) = .68, p = .41$ . Overall, females offered more practical and emotional support than males, main effects of gender:  $F_s(1, 51) > 4.13, p_s < .047$  and offered marginally more money,  $F(1, 51) = 3.9, p = .054$ , but not time,  $F(1, 51) = 1.91, p = .17$ . Null effects for donations of time likely resulted from floor effects, with a mode of 0 hours donated. Females particularly reported offering more support than males to *resilient* patients, gender-by-type interactions for emotional and practical support:  $F(1,$

**Table 2.2. Mean responses (SE) to *distraught* and *resilient* targets by observer gender<sup>1</sup>**

		Females		Males	
		Distraught	Resilient	Distraught	Resilient
Patient Perception	Distraught	5.69 (.20)	1.93 (.15)	5.36 (.20)	1.92 (.16)
	Softhearted	4.72 (.18)	5.89 (.18)	4.63 (.19)	5.65 (.18)
	Distressed	5.98 (.16)	1.67 (.13)	5.64 (.16)	1.79 (.13)
	Happy	1.96 (.14)	5.37 (.20)	1.98 (.14)	5.02 (.20)
	Resilient	2.39 (.22)	5.69 (.23)	3.27 (.22)	4.89 (.23)
	Sick	4.98 (.17)	3.09 (.16)	4.87 (.17)	2.87 (.17)
Emotional Responses	Distressed	4.85 (.25)	1.88 (.19)	4.18 (.26)	2.00 (.20)
	Horried	3.22 (.26)	1.33 (.13)	2.62 (.28)	1.42 (.14)
	Pleased	1.47 (.12)	4.21 (.322)	1.58 (.13)	3.60 (.34)
	Like	4.93 (.19)	5.83 (.21)	4.40 (.20)	5.29 (.22)
	Empathy	5.48 (.20)	4.86 (.21)	5.31 (.21)	3.96 (.22)
Helping Responses	Practical Support	5.34 (.24)	4.98 (.27)	4.79 (.25)	3.78 (.29)
	Emotional Support	5.46 (.21)	5.00 (.25)	5.17 (.22)	4.06 (.26)
	Money Donation	7.07 (.49)	6.32 (.55)	6.06 (.52)	4.54 (.59)
	Time Donation	0.32 (.08)	0.36 (.09)	0.19 (.10)	0.16 (.09)

<sup>1</sup> Note. All means are in units of the Likert scale from 1-7, except for money and time donation, which are in units of tokens (out of 9) and hours (0-4). Target and Observer emotion factors were defined for participants with a list of adjectives that significantly loaded on the factor in the prior study (Preston et al., under review). Target factors (with adjectives): Softhearted (softhearted, compassionate, tender, warm, sympathetic, moved), Distressed (disturbed, upset, afraid, bothered, panicked, distressed, disconcerted, troubled, perturbed, worried, sad, horrified, angry, sorrowful, grieved, alarmed, concerned), Happy (happy, amused, pleased). Observer factors (with adjectives): Empathic (compassionate, sympathetic, softhearted, tender, warm, moved), Distressed (troubled, distressed, worried, upset, afraid, grieved, sad, disturbed, bothered, concerned, sorrowful, alarmed, disconcerted), Pleased (amused, pleased, and happy), and Horrified (horrified, perturbed, angry, panicked).

51)  $> 4.58, p < .04$ ; gender effect within *resilient*,  $F(1,55) > 5.87, p < .02$ ; gender effect within *distraught*,  $F(1, 55) < 2.81, p > .10$ ; all other interactions,  $F_s(1, 51) < 1.82, p_s > .18$ .

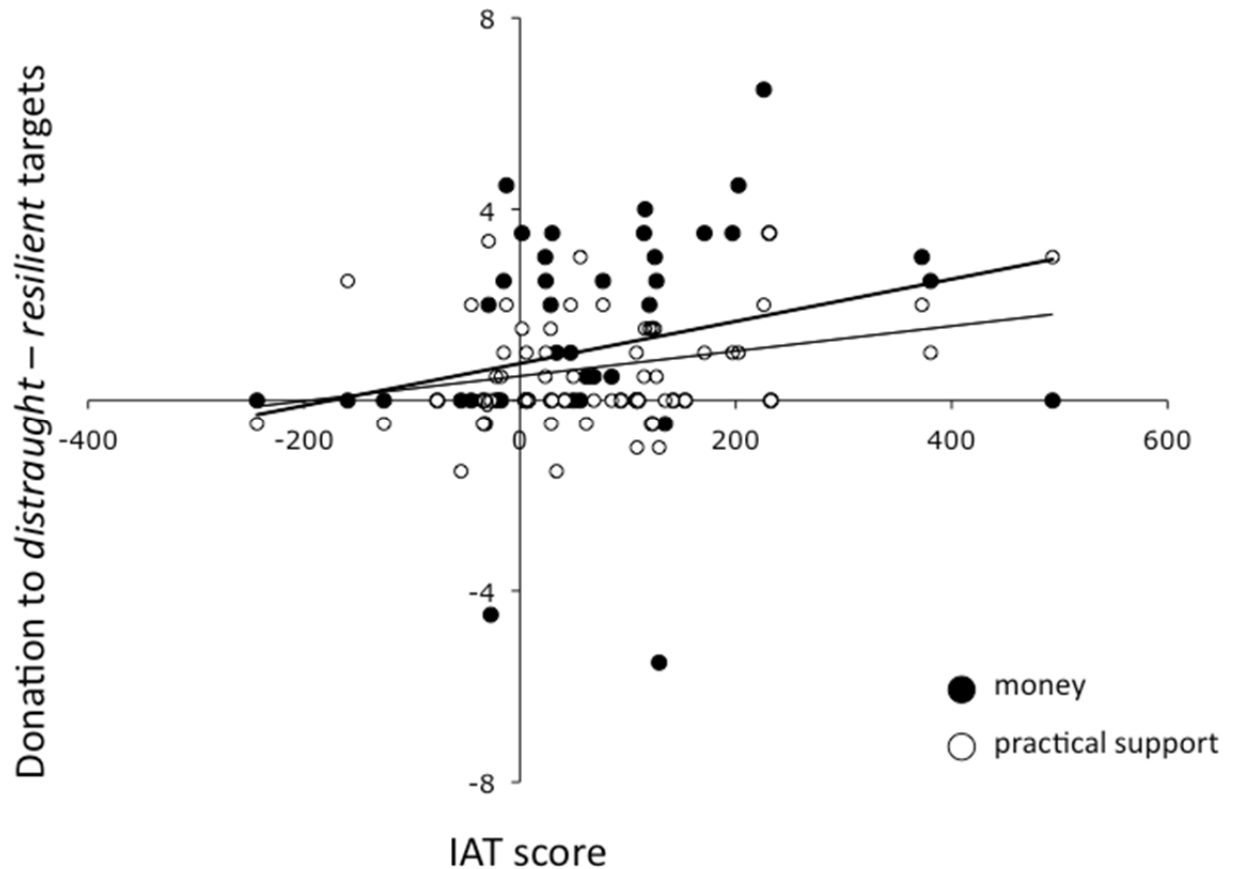
### Individual differences in the prosocial responses

Because we were most interested in whether differences in these associations would predict *differential* helping to each patient type, we examined the relationship between individual IAT scores and the difference in help given to *distraught* versus *resilient* patients. Participants with stronger female-sad associations on the IAT offered more help to *distraught* over *resilient* targets for practical support and monetary donation ( $r_s(57) > .26, p_s < .05$ ); emotional support, time donation,  $r_s(57) < .22, p_s > .1$ ; see Figure 2.1). The relationship between IAT score and preferential token donations to *distraught* patients did not differ by gender,  $\beta = 0.002, t = 0.52, p = .61$ , but the relationship with practical support did,  $\beta = -.006, t = -2.52, p = .015$ , being driven by females,  $\beta = .005, t = 3.78, p = .0008$ ; males only:  $\beta = -0.0009, t = -0.46, p = .65$ . Participants also perceived *distraught* patients as being more sick than *resilient*,  $r(53) = .27, p = .045$ . However, IAT scores did not predict perceptions that *distraught* patients needed more emotional or practical support than *resilient* patients,  $r(54) < .20, p > .14$ .

The IAT associations were not related to trait empathy or prosociality,  $r(52) < .21, p > .13$ , suggesting the association between females and sadness is not a result of an increased sensitivity to or ability to resonate with the emotions of others. Females had higher emotional empathy (ME) scores ( $M = 45.14$ ) than males ( $M = 21.28$ ),  $t(51.35) = -3.76, p < .001$ ,  $1.43, p = .16$ , while males ( $M = 90.96$ ) had higher Other-Oriented Empathy (PSB) scores than females ( $M = 81.57$ ),  $t(50.05) = 2.82, p = .01$ ; Helpfulness

scores did not differ between genders,  $t(50.44) = -1.43, p = .16$ . Trait empathy scores did not account for the relationship between IAT and increased ratings of sickness and help to *distraught* over *resilient* patients, as IAT was still a significant predictor of these variables when trait empathy scores were included as a covariate, IAT estimates:  $t(50) > 1.95, p < .057$ ; ME and PSB estimates:  $t(50) < 1.49, p > .14$ .

**Figure 2.1. Scatterplot of individuals' IAT score by degree of preferential help offered to *distraught* over *resilient* patients with best-fit lines<sup>2</sup>.**



<sup>2</sup> Linear regression lines are presented for money donations (darker line;  $\beta = 0.005, t = 2.13, p = .038$ ) and self-reported donations of practical support (lighter line;  $\beta = 0.003, t = 2.56, p = .013$ ).

## Discussion

While ideas about socially appropriate expressions in females and males are widely discussed and researched, to our knowledge this is the first study to demonstrate a continuously-varying, implicit, female-sadness/male-anger association. This association did not differ by participant gender, suggesting that even though males and females actually experience emotions like anger and sadness equally often (Barrett et al., 1998), people have learned gender-specific associations about the expression of emotion that affects their perceptions of others. Importantly, the female-sadness association was not uniform, but varied in strength across individuals with some people even showing reverse associations. From a perception-action view, others' states are processed through our own, stored, internal representations, which produce stronger prosocial responses for targets with relatable or expected need (Hofelich & Preston, 2011; Preston & Hofelich, 2012; Preston & de Waal, 2002). Supporting this, participants with stronger female-sadness associations perceived the *distraught* patients to be sicker, and offered more help to these patients over the more positive *resilient* females.

While female and male observers responded similarly overall, females had more extreme responses on multiple measures. Females liked, empathized with, and offered more help than males, particularly to *resilient* targets. These responses may have been affected by the fact that all four targets were females, which needs to be examined. Females also scored higher on the ME measure of emotional empathy, which is consistent with prior literature often finding higher self-reported trait empathy in females (Eisenberg & Lennon, 1983). However, it is unlikely that females better intuited the need of the *resilient* females, since they did not rate them as more sick or distraught; they

actually rated them as more resilient and still offered more help than male participants. Moreover, males actually scored higher on other-oriented empathy from the PSB, which has been found previously only for the helping subscale (Penner et al., 1995), and trait differences did not explain the relationship between IAT scores and helping. Taken together, the results suggest that female-sadness associations do not simply reflect individual differences in sensitivity to others' emotions, and that people give more help to targets who express emotions in ways that are compatible with their expectations, regardless of trait empathic tendencies.

Overall, this study demonstrates that while *distraught* female targets are helped more than *resilient* female targets, the extent of this effect varies with people's internal associations between females and sadness, with those who expect females to express sadness giving disproportionately more help to sad female patients. Chapter 3 replicates these results with a different measure of priming, and investigates whether responses to male patients would be similar.



## References

- Axelrod, R., & Hamilton, W. D. D. (1981). The evolution of cooperation. *Science*, 211, 1390-1396.
- Barrett, L. F., & Bliss-moreau, E. (2009). She's emotional. He's having a bad day: Attributional explanations for emotion stereotypes. *Emotion*, 9(5), 649 - 658.  
doi:10.1037/a0016821
- Barrett, L. F., Robin, L., Pietromonaco, P. R., & Eyssell, K. M. (1998). Are women the "more emotional" sex? Evidence from emotional experiences in social context. *Cognition & Emotion*, 12(4), 555-579.
- Batson, C. D., Fultz, J., & Schoenrade, P. A. (1987). Distress and empathy: Two qualitatively distinct vicarious emotions with different motivational consequences. *Journal of Personality*, 55(1), 19-39.
- Batson, C. D., Sympson, S. C., Hindman, J. L., Decruz, P., Todd, R. M., Weeks, J. L., Jennings, G., et al. (1996). "I've Been there, Too": Effect on empathy of prior experience with a need. *Personality and Social Psychology Bulletin*, 22(5), 474-482.  
doi:10.1177/0146167296225005
- Betancourt, H. (1990). An attribution-empathy model of helping behavior: Behavioral intentions and judgments of help-giving. *Personality and Social Psychology Bulletin*, 16(3), 573-591. doi:10.1177/0146167290163015
- Birnbaum, D. W., Nosanchuk, T. A., & Croll, W. L. (1980). Children's stereotypes about sex differences in emotionality. *Sex Roles*, 6(3), 435-443.

- Brody, L. R., & Hall, J. A. (2000). Gender, emotion, and expression. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of Emotions* (pp. 338-348). New York: The Guilford Press.
- Dijksterhuis, A., Aarts, H., Bargh, J. A., & van Knippenberg, A. (2000). On the relation between associative strength and automatic behavior. *Journal of Experimental Social Psychology*, 36(5), 531-544. doi:10.1006/jesp.2000.1427
- Eibl-Eibesfeldt, I. (1972). *Love and hate: The natural history of behavior patterns* (Trans Geoffrey Strachan). Oxford. Holt, Rinehart & Winston.
- Eisenberg, N., & Lennon, R. (1983). Sex differences in empathy and related capacities. *Psychological Bulletin*, 94(1), 100-131.
- Eisenberg, N., Wentzel, M., & Harris, J. D. (1998). The role of emotionality and regulation in empathy-related responding. *The School Psychology Review*, 27(4), 506-521.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6), 1464-1480.
- Hess, U., Adams, R. B. J., & Kleck, R. E. (2008). The role of facial expression in person perception. In N. Ambady & J. J. Skowronski (Eds.), *First Impressions* (pp. 234-254). New York: Guilford Press.
- Hodges, S. D., Kiel, K. J., Kramer, A. D. I., Veach, D., & Villanueva, B. R. (2010). Giving birth to empathy: the effects of similar experience on empathic accuracy, empathic concern, and perceived empathy. *Personality & Social Psychology Bulletin*, 36(3), 398-409. doi:10.1177/0146167209350326

- Hofelich, A. J., & Preston, S. D. (2011). The meaning in empathy: distinguishing conceptual encoding from facial mimicry, trait empathy, and attention to emotion. *Cognition & Emotion*, 26(1), 119-28. doi:10.1080/02699931.2011.559192
- Marsh, A. A., Kozak, M. N., & Ambady, N. (2007). Accurate identification of fear facial expressions predicts prosocial behavior. *Emotion*, 7(2), 239-251. doi:10.1037/1528-3542.7.2.239
- Mehrabian, a, & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality*, 40(4), 525-43.
- Parmley, M., & Cunningham, J. G. (2008). Children's Gender-Emotion Stereotypes in the Relationship of Anger to Sadness and Fear. *Sex Roles*, 58, 358-370. doi:10.1007/s11199-007-9335-9
- Penner, L. A., Fritzsche, B. A., Craiger, J. P., & Freifeld, T. S. (1995). Measuring the prosocial personality. In J. Butcher & C. D. Spielberger (Eds.), *Advances in Personality Assessment* (Vol. 10). Hillsdale, NJ: LEA.
- Plant, E. A., Hyde, J. S., Keltner, D., & Devine, P. G. (2000). The gender stereotyping of emotions. *Psychology of Women Quarterly*, 24(1), 81-92. doi:10.1111/j.1471-6402.2000.tb01024.x
- Plant, E. A., Kling, K. C., & Smith, G. L. (2004). The influence of gender and social role on the interpretation of facial expressions. *Sex Roles*, 51(3/4), 187-196. doi:10.1023/B:SERS.0000037762.10349.13
- Preston, S. D., & Hofelich, A. J. (2012). The many faces of empathy: Parsing empathic phenomena through a proximate, dynamic-systems view of representing the other in the self. *Emotion Review*, 4(1), 24-35. doi:10.1177/1754073911421378

- Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *The Behavioral and Brain Sciences*, 25(1), 1-20; discussion 20-71.
- Preston, S. D., Bechara, A., Damasio, H., Grabowski, T. J., Stansfield, R. B., Mehta, S., & Damasio, A. R. (2007). The neural substrates of cognitive empathy. *Social Neuroscience*, 2(3-4), 254-275. doi:10.1080/17470910701376902
- Preston, S. D., Hofelich, A. J., & Stansfield, R. B. (under review). The ethology of empathy: Emotions displayed and elicited by real targets of need.
- Rath, A. (2012, March 5). The meaning of Putin's tears. *PRI's The Word*.
- Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion knowledge: further exploration of a prototype approach. *Journal of Personality and Social Psychology*, 52(6), 1061-86.
- Sriram, N., & Greenwald, A. G. (2009). The brief implicit association test. *Experimental Psychology*, 56(4), 283-294.
- Weiner, B. (1980). A cognitive (attributional)-emotion-action model of motivated behavior: An analysis of judgments of help-giving. *Journal of Personality and Social Psychology*, 39(2), 186-200.
- Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., & Chapman, M. (1992). Development of concern for others. *Developmental Psychology*, 28(1), 126-136.  
doi:10.1037/0012-1649.28.1.126

## Chapter 3: Priming Associations Affect the Prosocial Response

### Introduction

Chapter 2 demonstrated that observers who expect females to express sadness give disproportionately more help to sad female patients over positive *resilient* ones. Despite this evidence supporting the PAM, a few potential questions remain. We interpreted greater helping of *distraught* females as emanating from a specific association between females and sadness. However, it is possible that participants with higher female-sadness associations generally hold more stereotyped beliefs, which would produce stronger IAT scores and a greater motivation to help distressed females because they are regarded as more vulnerable or as prototypical targets of aid. In addition, Chapter 2 only used female targets. This was done for practical reasons because none of the existing male patients were classified as *distraught*, presumably attesting to real-world gender differences in the display of emotion. We would not expect males to receive greater help based on the strength of female-sadness associations, but such a relationship would exist if the view above were correct (that people helped crying women simply because of stereotypes about crying people needing the most help) or if the general norm of whether overt distress is permissible was stronger than the gender-specific norm. Motivations to help males in overt need or distress are important to understand,

particularly for real world charities that aim to assist predominantly male targets, such as research for testicular cancer or support for veterans.

Chapter 3 was designed to replicate and extend the results from Chapter 2 while addressing the potential issues with Chapter 2. Chapter 3 included both male and female *distraught* and *resilient* patients by using modified written transcripts of the videotaped patients that were changed to depict either males or females (see Appendix 3.A).

Additional trait scales were added to further confirm that individual differences in female-sadness associations do not reflect general differences in sensitivity to social information. Also, because the IAT measure confounds multiple different associations (i.e., with scores that include increased female-sadness and male-anger associations as well as decreased female-anger and male-sadness associations), it was replaced with a simpler priming task (modified from Dijksterhuis et al., 2000). Like the IAT, priming tasks have been shown to capture implicit associations and be predictive of behavior (Dijksterhuis et al., 2000; Fazio & Olson, 2003). This priming task can generate separate scores for implicit associations between sadness and both males and females, and can also be used to generate scores for more general nonaffective gender-based stereotypes. Because the IAT score reflects the extent to which people associate attributes like sadness differentially with males and females, high IAT scores may also reflect more general beliefs about what behaviors are expected for each gender, rather than associations with females and sadness more specifically. In order to verify that female-sadness associations were driving the relationship with preferential giving to *distraught* females, and not more general stereotypical associations about gender, Chapter 3 sought to replicate the findings from Chapter 2 using the “cleaner” or more direct measure of

female-sadness associations from the priming task. Similar results in which the female-sadness priming scores, but not the stereotype scores, predict differential giving to *distraught* over *resilient* female patients would support the prediction of the PAM.

## **Methods**

### **Participants**

Seventy students (33 women and 37 men, mean age = 19, range = 18 - 22) from the University of Michigan participated in this study for course credit. Participants were tested in groups in a computer lab and were seated at every other computer terminal. All participants gave full informed consent and all procedures were approved by the Institutional Review Board of the University of Michigan.

### **Priming Task**

The priming lexical decision task was modified from Dijksterhuis and colleagues (2000) and was administered using Eprime Version 2.0 (Psychology Software Tools, Inc., Pittsburgh, PA). Participants saw 32 words and 32 non-words that were matched on length (non-words were created with Word-Gen; Duyck, Desmet, Verbeke, & Brysbaert, 2004), and indicated as quickly and as accurately as possible whether the stimulus was a real word or not. Participants responded by pressing buttons marked “yes” or “no” on a keyboard with their right or left index fingers. Each word and non-word was preceded by one of two 17ms primes (“WOMAN” or “MAN”) followed by a string of X’s for 250ms, which served as a backwards mask for the prime (as in Dijksterhuis et al., 2000). The

primes were presented in blocks of 32 trials, with the order counterbalanced across participants. The words were from four categories, sad (taken from Chapter 2; *woeful, grieving, sorrow, depressed, miserable, hopeless, crying, sadness*), female stereotype (taken from Rudman & Kilianski, 2000; *helper, supportive, shopping, domestic, kitchen, secretary, family*), male stereotype (*supervisor, boss, competitive, executive, career, leader, authority, office*; Rudman & Kilianski, 2000), or neutral (*pamphlet, umbrella, thumb, thermometer, contents, clock, barrel, square*; McKenna & Sharma, 1995). Participants were given ten practice trials before the task.

For analysis, trials were removed if there was an error and if RTs were shorter than 200ms (reflecting anticipatory responses) or exceeded 3 standard deviations (SD) of a participant's mean RT (as done in similar lexical decision tasks; e.g., Holcomb & Neville, 1990). To control for effects of word frequency on performance in the lexical decision task, the relationship between the Kucera Francis (1967) frequency count and RT was regressed out in each participant before calculating association scores (see Appendix 3.B for an explanation of why this approach was used).

To generate the cleaner gender-association scores that were hypothesized to predict prosocial responses, we created female-sadness and a male-sadness association scores as well as separate female-role stereotyping and male-role stereotyping scores. Female-sadness (and female-stereotype) associations were created by subtracting the frequency-corrected RT for sad (or female-stereotype) words following the "WOMAN" prime from the RT for neutral words following the "WOMAN" prime. Male-sadness (or male-stereotype) association scores were calculated by subtracting the frequency-



corrected RT for sad (or male-stereotype) words following the “MAN” prime from the RT for neutral words following the “MAN” prime.

### **Prosocial Response Task**

After completing the priming task, participants read eight short video transcripts of patient interviews on Qualtrics (Qualtrics Labs Inc., Provo, UT). Each participant read two exemplar transcripts from each patient type (i.e., two each of *distraught* male, *distraught* female, *resilient* male, *resilient* female), presented in a random order. Of the four *distraught* and four *resilient* transcripts, the pairs presented as male or female were randomly selected across participants. Transcripts were selected from a total set of sixteen (each of the eight transcripts was modified to have a male and female version). The eight transcripts were comprised of the transcripts from the four original female patient videos used in Chapter 2, two transcripts from original patient videos of *resilient* males, and one from a video of a *distraught* female who was not used in Chapter 2. In addition, one female patient from our original video set that was originally classified as *wistful* was included because she was highly similar to our *distraught* patients (in previous studies she was rated as the most distressed non-*distraught* patient; Preston et al., under review). Emotional cues and information about the patient’s facial expressions (such as “[voice breaking]” or “[smiling slightly]”) were added to the transcripts to convey the emotional tone of the original interview (the emotional cues were exaggerated for the non-original *distraught* patient so that the transcript would be similar to the other *distraught* ones). The transcripts were pre-tested in a separate group of participants to

ensure that they were rated similarly to the videos in Chapter 2 and that the male patients were believable (see Appendix 3.A for pre-testing methods and results).

After reading each transcript, participants rated the patient on the same dimensions in Chapter 2 using a seven-point Likert scale from 1 (not at all) to 7 (extremely). Token donations were the same as in Chapter 2, but were out of five instead of nine tokens for each patient (due to the increased number of patients). Time donations were omitted because of the null and floor effects in Chapter 2.

### **Trait Scales**

To further test whether individual differences in associations between females and sadness could be explained by differences in sensitivity to social information, participants completed the Emotional Intelligence Scale (EI; Schutte et al., 1998). The Berkeley Expressivity Questionnaire (BEQ; Gross & John, 1995) was added to determine whether associations with sadness could be explained by individual differences in participants' own expressivity. To determine whether associations were related to explicit gender role stereotypes, participants completed the Gender and Authority Measure (GAM; Rudman & Kilianski, 2000). Finally, participants completed the BEM Sex Role Inventory (BEM; Bem, 1974) to determine whether sadness associations varied with a continuous measure of participant gender identity.

### **Statistical Analyses**

All measures were averaged between the two targets within each type (*distraught* males, *distraught* females, *resilient* males, *resilient* females) and were compared using

mixed-ANOVA with participant gender as a between-subjects factor. Analyses were the same as in Chapter 2, except that patient gender was added as a repeated variable when comparing emotional, empathic and prosocial response across patient types. Alpha was set at 0.05 for all comparisons.

## **Results**

### ***Priming Task***

Accuracy on the priming task was high, with an average error rate of 4.7%. Because of the limited number of trials in this task, participants with error rates greater than 15% (2 SD above the mean) were removed from further analysis ( $n = 4$ ). Justifying the need to control for word frequency in all subsequent analyses, RT did decrease with word frequency,  $\beta = -0.21$ ,  $p = .005$ , even after error trials were removed.

### ***Associations with sadness***

To predict differential giving from gender-based associations with emotion, we first want to determine whether participants held the expected gender-emotion associations. Female participants indeed showed the expected pattern of stronger female-sadness ( $M = 20.65$ ) than male-sadness associations ( $M = -53.05$ ), post-hoc  $z = -2.11$ ,  $p = .035$ . However, males did not show this pattern,  $z = 1.01$ ,  $p = .31$ , with means that were actually nonsignificantly higher for male-sadness ( $M = 14.05$ ) than female-sadness associations:  $M = -14.59$ , overall gender by association type interaction:  $F(1, 68) = 4.70$ ,  $p = .03$ . Thus, female and male participants may hold stronger associations between sadness and their own gender, rather than for females *per se*. This may be due to the

exact sad words employed (*woeful, grieving, sorrow, depressed, miserable, hopeless, crying, sadness*), most of which represent internal, subjective states rather than overt expressions or displays of emotion. While beliefs about the expression of sadness may not differ across participants (with both genders believing females *express* sadness more), beliefs about the subjective experience of sadness may be more strongly associated with one's own gender (as the *experience* of sadness is similar in males and females; Barrett et al., 1998). As evidence favoring this interpretation, the RT effects specific to participants' genders were most pronounced for “depressed” and “grieving”—arguably the most internal emotional terms of the eight words in the set. Because these words refer to internal, experience-dependent states (rather than more expressive states like “crying”), it follows that participants would have the strongest associations between these words and their own gender (instead of both genders associating all sad adjectives with females). However, this does not affect the key prediction of the study, as female-sadness associations still varied across participants, and the question of primary interest is whether individual differences in these associations predict differential giving to *distraught* over *resilient* patients; the fact that females have this association more does not affect our ability to test this.

Participants did also demonstrate strong gender-based stereotype associations,  $ts(69) > 2.20$ ,  $ps < .03$ , which were similar for male and female participants,  $F(1, 68) = 0.01$ ,  $p = .91$ .

### ***Perceptions of patient types***

As in the pilot study (Appendix 3A), the added *distraught* transcript was rated as less distraught and distressed than the other patients,  $F(3, 219) > 9.51, ps < .001$ ; post-hoc  $zs > 3.69, ps < .001$ . Similarly, the added *resilient* transcript was rated as less resilient than the other patients,  $F(3, 219) = 3.61, p = .01$ ; post-hoc  $zs > 2.57, ps < .05$ , although all *resilient* patients were perceived as equally happy,  $F(3, 219) = 1.39, p = .25$ .

Table 2.3 summarizes the mean responses per dependent measure, by patient type, patient gender, and observer gender. Despite the fact the new transcripts were rated less extreme than the original patients, and confirming the patient transcripts were perceived the same as the videos in Chapter 2, *distraught* targets again appeared more distressed, distraught, sicker, and in more need of emotional and practical support,  $F(1, 62) > 65.09, ps < .001$ , while *resilient* targets were perceived as more resilient and happy,  $F(1, 62) > 278.77, ps < .001$ , and somewhat more softhearted,  $F(1, 62) = 3.45, p = .07$ . Female observers again rated *resilient* targets as more resilient and *distraught* targets as less resilient compared to males, gender-by-target interaction for resilient,  $F(1, 62) = 7.05, p = .01$ ; all others  $F(1, 62) < 2.32, p < .13$ . However unlike Chapter 2, male observers thought that all targets were happier,  $F(1, 62) = 4.88, p = .03$ , slightly more resilient,  $F(1, 62) = 3.12, p = .08$ , and in need of less emotional support,  $F(1, 62) = 5.48, p = .02$ , compared to females.

Interestingly, perceptions of the patients did not differ by their gender, with the exception of female patients tending to appear more distressed than males,  $F(1, 62) = 3.23, p = .077$ , all other ratings:  $F(1, 62) < .22, ps > .65$ . *Distraught* and *resilient* patients of each gender were also perceived similarly,  $F(6, 57) = 1.30, p = .27$ .

### ***Emotional and prosocial responses to patient types***

Confirming that the responses elicited by the transcripts were the same as the videos in Chapter 2, participants again felt more empathic towards but also more horrified and distressed by *distraught* targets,  $F_s(1, 61) > 66.33, ps < .001$ , and felt more pleased by *resilient* targets,  $F(1, 62) = 202.7, p < .001$ . Male and female observers reacted similarly to the patients, all gender main effects:  $F_s(1, 62) < 3.62, ps > .06$ .

The gender of the patient also did not influence the participants' emotional responses to the patients, again with the exception of a trend of females being liked less than male patients,  $F(1, 62) = 3.00, p = .088$ ; all other ratings,  $F(1, 62) < 1.75, p > .1$ . Mirroring the findings that participants more strongly associated and perhaps better related to sadness with their own gender, female observers were more distressed and horrified by male *distraught* patients, while male observers were more distressed and horrified by female *distraught* patients, three way interaction,  $F(1, 62) = 4.34, p = .04$ . All participants responded to *resilient* patients of both genders similarly.

Prosocial responses to the targets were again similar to Chapter 2, as participants offered more emotional and practical support, and donated more money to *distraught* compared to *resilient* patients,  $F_s(1, 62) > 40.69, ps < .001$ . However, male and female observers did not differ in their overall support or monetary donations,  $F_s(1, 62) < 2.26, ps > .14$ .

Giving to male and female patients did not show the same pattern as feelings of distress and horror, in which the opposite gender's sadness was most distressing. Instead, more emotional support was given to the emotional displays more "typical" of that

**Table 2.3. Mean ratings by patient type, patient gender, and observer gender.**

		Distraught				Resilient			
		Female patient		Male patient		Female patient		Male patient	
		Female observer	Male observer	Female observer	Male observer	Female observer	Male observer	Female observer	Male observer
Target emotion	distraught	5.85 (0.98)	5.96 (0.71)	6.02 (0.85)	5.69 (0.83)	2.38 (0.88)	2.44 (0.78)	2.28 (0.73)	2.47 (0.88)
	resilient	2.30 (0.86)	2.99 (1.15)	2.35 (0.99)	3.03 (1.20)	5.33 (0.89)	5.32 (0.90)	5.42 (0.97)	5.12 (1.12)
	sick	4.87 (1.27)	4.88 (1.16)	4.88 (1.13)	4.78 (1.10)	3.22 (1.22)	3.06 (0.91)	3.10 (1.20)	3.06 (0.96)
	likable	5.07 (1.17)	4.56 (1.31)	5.13 (0.92)	4.79 (1.39)	5.03 (1.14)	4.99 (1.20)	5.35 (0.90)	5.04 (1.01)
	softhearted	5.20 (1.36)	5.19 (1.35)	5.30 (1.14)	4.84 (1.44)	5.45 (1.00)	5.37 (1.10)	5.55 (0.90)	5.34 (1.11)
	distressed	5.83 (1.01)	6.18 (0.78)	6.07 (0.86)	5.94 (0.77)	2.53 (0.99)	2.51 (0.93)	2.18 (0.81)	2.25 (0.94)
	happy	1.73 (0.50)	2.18 (0.91)	1.75 (0.58)	2.24 (0.85)	5.18 (0.94)	5.43 (0.81)	5.25 (0.94)	5.22 (0.95)
Observer emotion	empathic	4.53 (1.43)	4.44 (1.35)	4.83 (1.18)	4.25 (1.39)	2.28 (0.99)	2.21 (0.93)	2.28 (1.06)	2.31 (0.86)
	distressed	5.60 (1.28)	5.31 (1.13)	6.03 (0.94)	5.12 (1.42)	4.35 (1.05)	4.12 (1.23)	4.35 (1.20)	4.24 (1.04)
	pleased	1.47 (0.66)	1.91 (0.92)	1.43 (0.50)	1.81 (0.88)	3.72 (1.33)	3.81 (1.13)	3.85 (1.12)	4.15 (1.34)
	horrified	2.73 (1.42)	2.81 (1.36)	2.95 (1.32)	2.51 (1.26)	1.48 (0.64)	1.60 (0.89)	1.45 (0.56)	1.79 (0.93)
Prosocial response	Practical support	5.00 (1.40)	4.63 (1.63)	4.88 (1.63)	4.18 (1.63)	3.88 (1.63)	4.00 (1.51)	4.25 (1.60)	3.72 (1.57)
	Emotional Support	5.97 (1.24)	5.32 (1.41)	5.85 (1.23)	5.18 (1.47)	4.38 (1.42)	4.13 (1.59)	4.55 (1.54)	4.25 (1.49)
	Money Donation	3.77 (1.36)	3.34 (1.65)	3.92 (1.31)	3.35 (1.66)	2.72 (1.58)	2.62 (1.88)	3.02 (1.56)	2.69 (1.79)

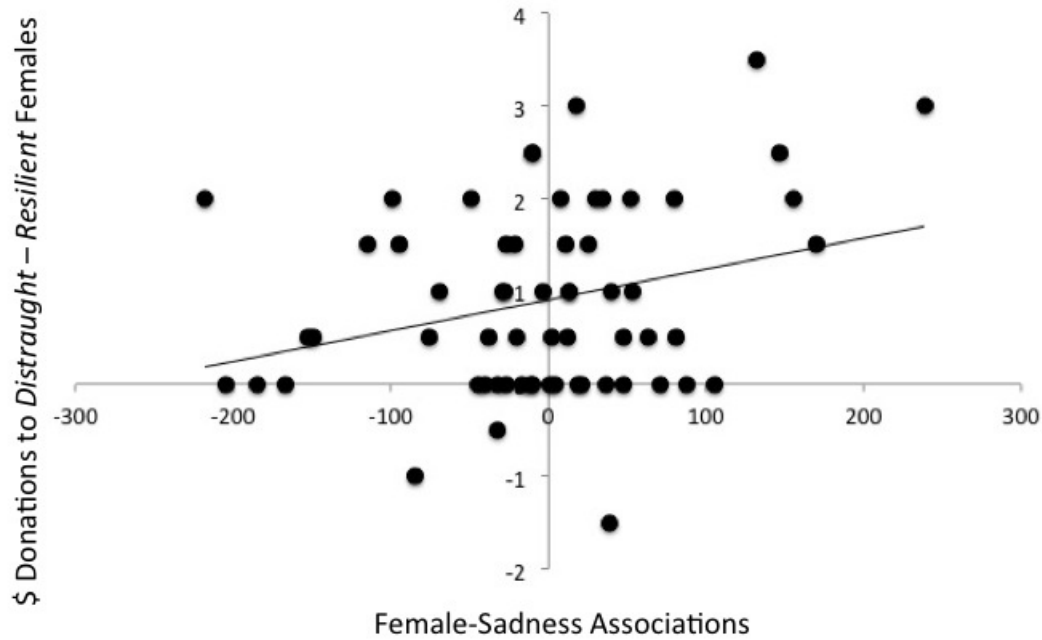
gender – *distraught* females (over *distraught* males), and *resilient* males (over *resilient* females), interaction:  $F(1, 62) = 3.81, p = .057$ . Furthermore, although more distressed by the opposite gender's sadness, observers gave more to the opposite gendered patients regardless of display type, gender by patient gender interaction,  $F(1, 62) = 7.24, p = .009$ . Collapsing across these interactions, overall giving was similar to each gender,  $F_s(1, 62) < 2.83, p > .1$ .

### ***Sadness associations and prosocial behavior***

Although not found in everyone, female-sadness associations again varied across participants and this variance was used to predict differences in giving. Replicating the results from Chapter 2, participants with strong female-sadness associations donated more money to *distraught* over *resilient* female targets,  $r(60) = .28, p = .026$  (see Figure 3.1). This relationship did not differ between male and female observers,  $\beta = -0.003, t = -0.98, p = .33$ , and was still significant after controlling for observer gender,  $\beta = 0.005, t = 2.32, p = .02$ . These results were only seen for actual money donations and were not correlated with self-reported practical or emotional support given as they were in Chapter 2,  $r(62) < .1, p > .46$ . As predicted, female-sadness associations specifically predicted preferential help for female targets, and did not predict increased giving to *distraught* over *resilient* male targets,  $r(60) = .12, p = .37$ .

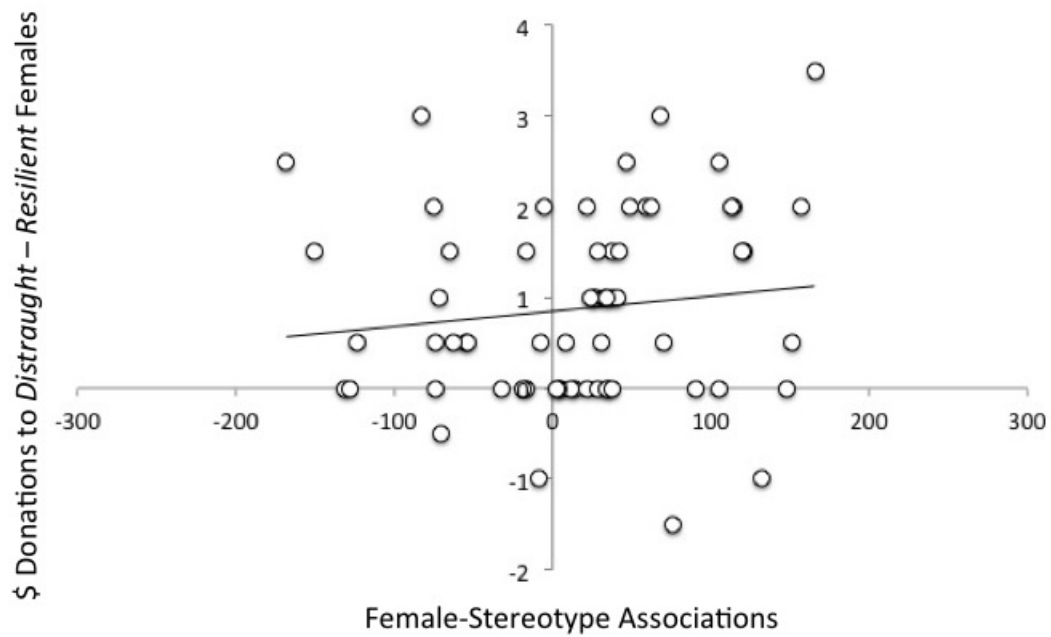


**Figure 3.1. Scatterplot of female-sadness associations and preferential giving to *distraught* over *resilient* female targets with best-fit line.**

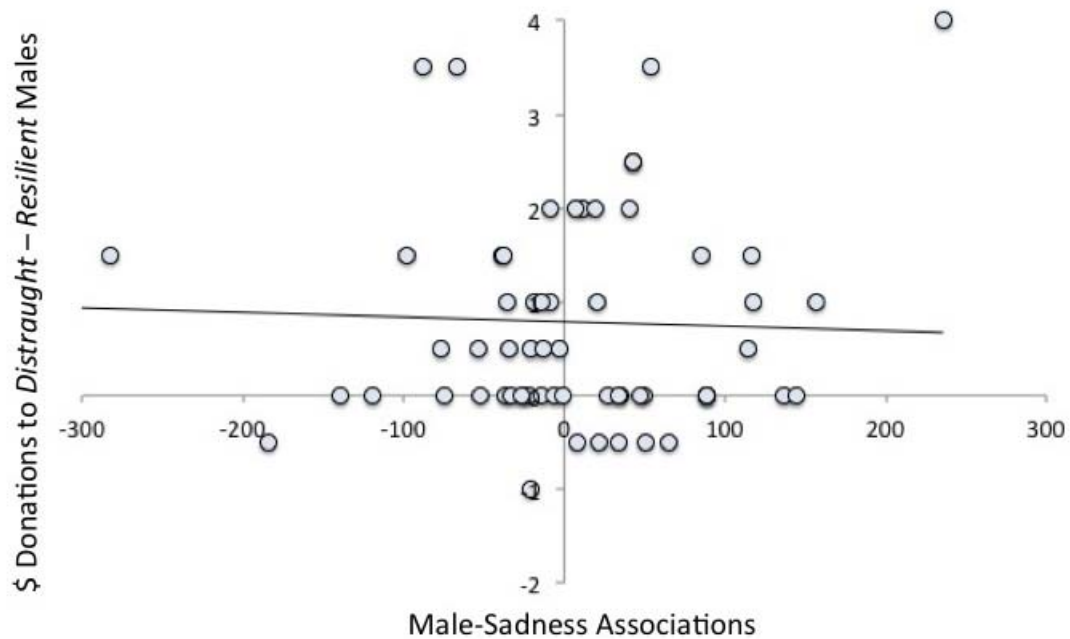


Supporting our hypothesis, the relationship between individual associations and preferential giving to *distraught* female patients was also specific to associations with sad affect, as female-stereotype scores were not associated with giving more money or self-reported help to *distraught* over *resilient* female patients,  $rs(64) < .12, ps > .33$  (see Figure 3.2), and female-sadness associations still marginally predicted preferential money donations to *distraught* over *resilient* patients even after controlling for female-stereotype associations,  $\beta = 0.003, t = 1.77, p = .08$ ; female-stereotype scores were not significant:  $\beta = 0.00001, t = 0.55, p = .59$ . This is particularly striking given the strong correlation between the two types of associations,  $r(62) = .51, p < .001$ . In contrast with the findings for female patients, neither male-sadness nor male-stereotype associations predicted preferential giving of practical support, emotional support, or money to male patients  $rs(61) < .12, ps > .40$  (see Figure 3.3).

**Figure 3.2. Scatterplot of female-stereotype associations and preferential giving to *distraught* over *resilient* female targets with best-fit line.**



**Figure 3.3. Scatterplot of male-sadness associations and preferential help to *distraught* over *resilient* male targets.**



### ***Relationship between association scores and trait measures***

Again confirming that female-sadness associations do not simply reflect a general sensitivity to social information, individual differences in these scores were not related to emotional intelligence, trait expressivity, or explicit gender role stereotypes,  $r(63) < .175$ ,  $p > .17$ . Furthermore, female-sadness associations still predicted preferential giving even when controlling for each of these trait measures,  $\beta > .003$ ,  $ts > 2.07$ ,  $ps < .04$ , while female-stereotype associations still did not,  $\beta < .0018$ ,  $ts < 1.02$ ,  $ps > .31$ . Participants with strong preferences for males in authority (GAM) had smaller male-sadness associations,  $r(63) = -.26$ ,  $p = .037$ , indicating the association scores were consistent with explicit attitudes about gender – those who preferred males to be in authority (and thus in a more “typical” role) also had more “gender-typical” implicit associations that males are not sad. The association scores also reflected consisted relationships between sadness and stereotype associations for each gender. For associations about females, sadness and stereotype measures were positively correlated,  $r(62) = .51$ ,  $p < .001$ , while male-sadness associations were not related to male-stereotype associations,  $r(61) = .15$ ,  $p = .25$ .

## Discussion

Attesting to the robustness of the findings from Chapter 2, *distraught* and *resilient* targets were characterized similarly and the giving effects were replicated, despite presenting patients in transcripts and using different measures of association.

Associations between females and sadness varied across participants, and again predicted money donations to *distraught* over *resilient* female patients, with observers giving more to targets who express their emotions in ways consistent with their associations. This finding supports predictions of the perception-action model (PAM), as it demonstrates that differences in participants' existing representations of females (specifically, the extent to which they expect females to be sad) predict differences in the amount of support given to sad female targets, confirming that understanding and empathizing with another draws on observers' representations of the target and relevant state. This effect was specific to associations about sadness, rather than more general, non-affective associations about gender roles, indicating the differences in giving were specific to individual variation in relevant representations activated by seeing the target in need (such as their emotional display). Female-sadness associations also did not predict preferential giving to male targets, supporting our predictions that this effect is not due to a more general belief that distressed targets should receive aid, but rather depends on specific affective associations for the target. Associations between females and sadness were also not related to emotional intelligence or expressivity measures and these measures also did not account for the relationship with giving, providing further evidence that these associations are not simply a reflection of sensitivity to social information.

In contrast to Chapter 2, strong female-sadness associations in the priming task were not found across all participants, as only females showed the expected association between females and sadness. This result may be due to sadness associations in this task capturing beliefs about expression as well as the experience of emotion within one's own gender, which is similar for males and females (Barrett, Robin, Pietromonaco, & Eysell, 1998). Such associations may render sad emotion in same-gendered targets more expected and less aversive. Supporting this, participants found *distraught* patients of the opposite gender more distressing and horrifying than same-gender *distraught* patients.

Associations based on experience may have been more salient to participants, but they likely still had knowledge of the more “widely held” belief that females express sadness more than males (Eagly & Koenig, 2006). Although feelings of distress and horror were more pronounced for distress of the opposite gender, participants gave more emotional support to the gender more “typical” of the emotional displays. Ratings of emotional support were higher for *distraught* female targets, while for *resilient* patients ratings were higher for males. These findings demonstrate that participants still had access to the more typical beliefs that females express sadness more than males, despite female-sadness association scores being stronger only in female participants.

While the variation in female-sadness associations again predicted preferential giving to *distraught* over *resilient* female targets, the relationship between male-sadness associations and preferential help for *distraught* over *resilient* male patients was not found, although theoretically it would also be predicted by the PAM. This may be because associations between males and displays of sadness were not typically held by participants, as the both the median (-6.08) and the most common male-sadness scores (-

13.34) were negative. However, even though the majority of participants had negative associations, it is still surprising those with negative male-sadness associations did not favor the *resilient* males, which would have also produced a positive correlation. It is possible the context of chronic or terminal illness in this study may have rendered the sadness of *distraught* male targets more permissible, making participants more likely to override any anti-empathic response they would have towards these targets. This is one limitation of this study, and future work can determine whether the predicted relationship between male-sadness associations and preferential giving would appear when the context of males' distress is manipulated to be less permissible.

## General Discussion

Together, our results demonstrate that individual differences in associations between females and sadness predict prosocial responses to female targets in need, with the most help going to targets who express their emotions in expected ways. While across participants, giving is higher for targets who are the most distressed and who act in ways consistent with widely held gender beliefs about emotion, these studies demonstrate that the extent of this effect can be predicted by continuously varying measures of internal associations. Although correlations between priming and IAT measures are typically low (Fazio & Olson, 2003), the fact that both the IAT and female-sadness priming measures predicted preferential donations to *distraught* over *resilient* female patients demonstrates the robustness of this effect. This relationship was not associated with or better explained by the extent to which individuals had more general, non-affective associations about females (stereotype associations). Additionally, the relationship between sadness associations and giving was found only within emotion associations that were more commonly held, and was not seen for male-sadness associations and giving to *distraught* over *resilient* male targets.

Much work in social psychology has demonstrated the prevalence of beliefs about behaviors, roles, and attributes common to each gender (e.g., Barrett et al., 1998; Eagly & Koenig, 2006). When others violate these expectations, they can be met with feelings of surprise or even moral disapproval, based on whether the actions violate beliefs about what people normally do or beliefs about what they “ought” to do (Cialdini & Trost, 1998). For example, passive women and aggressive men are perceived as more popular and well-adjusted than when the characteristics are reversed (Costrich, Feinstein, Kidder,

Marecek, & Pascale, 1975), affective satisfaction and compliance of patients are higher when female doctors are passive, rather than aggressive or dominant (Burgoon, Birk, & Hall, 1991; Schmid Mast, 2004), and leaders who behave in gender-congruent ways are perceived as more effective (at least in male participants; Rojahn & Willemsen, 1994).

The current work is consistent with these studies, and further extends this literature to prosocial responses to targets who conform with or violate expectations about displayed emotions. Interestingly, patients who were the most violating of widely held expectations about sadness (*distraught* males) were not liked less than more conforming targets (*distraught* females). Unfamiliar or unexpected displays based on associations instead produced feelings of distress and horror, suggesting that such violations in needy others render them more distressing, rather than less likable or competent. The context of distress in this study likely limited the extent to which participants disliked or perceived the targets unfavorably, as all patients were hospitalized and described as being chronically or terminally ill. It also may account for the lack of correlation between male-sadness associations and preferential giving to sad male targets. Future work should examine prosocial responses in contexts where unexpected displays of distress are less permissible, as we would expect differences in empathy and giving to be more extreme.

Much more research is needed on how individuals differ in the display of need, and how such displays affect observers with differing temperaments, personalities, and cultural backgrounds. Our findings suggest both perceptions of and emotional responses to targets are influenced by culturally-prescribed beliefs about emotion expression, even if associations with the experience of emotion are less pervasive across individuals.



Furthermore, the extent to which these associations differ across individuals predicts preferential help for targets who express emotions in ways that are consistent with these associations. While prior research has focused primarily on observers, largely holding targets constant through fictionalized narratives and images designed to be as sympathetic as possible, these studies demonstrate decreased giving, distress, and horror responses can be elicited by targets. In the real world, others' distress does not always engender compassion, and can even provoke aggression at times (e.g., Marsh, Adams, & Kleck, 2005; Miller, Stiff, & Ellis, 1988; Zahn-Waxler et al., 1992). Such unsympathetic responses are hardly restricted to laboratory situations and adults, as even children identify targets of bullying as others who deviate from expected "norms" in appearance or actions (Hantler, 1994). Only by investigating these important dynamics can we begin to understand our concurrent capacity for empathy and indifference in everyday life.

## References

- Barrett, L. F., Robin, L., Pietromonaco, P. R., & Eyssell, K. M. (1998). Are women the “more emotional” sex? Evidence from emotional experiences in social context. *Cognition & Emotion*, 12(4), 555-579.
- Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, 42(2), 155-162.
- Ben-Ner, A., Kong, F., & Putterman, L. (2004). Share and share alike? Gender-pairing, personality, and cognitive ability as determinants of giving. *Journal of Economic Psychology*, 25(5), 581-589. doi:10.1016/S0167-4870(03)00065-5
- Burgoon, M., Birk, T. S., & Hall, J. R. (1991). Compliance and satisfaction with physician-patient communication: An expectancy theory interpretation of gender differences. *Human Communication Research*, 18(2), 177-208.
- Cialdini, R. B., & Trost, M. R. (1998). Social influence: Social norms, conformity, and compliance. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *The Handbook of Social Psychology* (pp. 151–192). McGraw-Hill.
- Costrich, N., Feinstein, J., Kidder, L., Marecek, J., & Pascale, L. (1975). When stereotypes hurt: Three studies of penalties for sex-role reversals. *Journal of Experimental Social Psychology*, 11, 520-530.
- Dijksterhuis, A., Aarts, H., Bargh, J. A., & van Knippenberg, A. (2000). On the relation between associative strength and automatic behavior. *Journal of Experimental Social Psychology*, 36(5), 531-544. doi:10.1006/jesp.2000.1427

- Duyck, W., Desmet, T., Verbeke, L. P. C., & Brysbaert, M. (2004). WordGen: A tool for word selection and nonword generation in Dutch, English, German, and French. *Behavior Research Methods, Instruments, & Computers*, 36(3), 488-499.
- Eagly, A. H., & Koenig, A. M. (2006). Social role theory of sex differences and similarities: Implication for prosocial behavior. In K. Dindia & D. J. Canary (Eds.), *Sex differences and similarities in communication 2nd ed* (pp. 161-177). Lawrence Erlbaum Associates Publishers.
- Fazio, R. H., & Olson, M. a. (2003). Implicit measures in social cognition. research: Their meaning and use. *Annual Review of Psychology*, 54, 297-327.  
doi:10.1146/annurev.psych.54.101601.145225
- Gross, J. J., & John, O. P. (1995). Facets of emotional expressivity: Three self-report factors and their correlates. *Personality and Individual Differences*.  
doi:10.1016/0191-8869(95)00055-B
- Hantler, A. M. (1994). Children's views of bullying. *Health Education*, 5, 8-14.
- Holcomb, P. J., & Neville, H. J. (1990). Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. *Language and Cognitive Processes*, 5(4), 281-312.
- Marsh, A. A., Adams, R. B., & Kleck, R. E. (2005). Why do fear and anger look the way they do? Form and social function in facial expressions. *Personality and Social Psychology Bulletin* (Vol. 31, pp. 73-86).
- McKenna, F. P., & Sharma, D. (1995). Intrusive cognitions: An investigation of the emotional Stroop task. *Journal Of Experimental Psychology: Learning Memory And Cognition*, 21(6), 1595-1607. doi:10.1037/0278-7393.21.6.1595

- Miller, K., Stiff, J., & Ellis, B. H. (1988). Communication and empathy as precursors to burnout among human service workers. *Communication Monographs*, 55, 250-265.
- Rojahn, K., & Willemssen, T. M. (1994). The evaluation of effectiveness and likability of gender-role congruent and gender-role incongruent leaders. *Sex Roles*, 30(1-2), 109-119. doi:10.1007/BF01420743
- Rudman, L. a., & Kilianski, S. E. (2000). Implicit and Explicit Attitudes Toward Female Authority. *Personality and Social Psychology Bulletin*, 26(11), 1315-1328. doi:10.1177/0146167200263001
- Schmid Mast, M. (2004). Dominance and gender in the physician-patient interaction. *The Journal of Men's Health & Gender*, 1(4), 354-358. doi:10.1016/j.jmhg.2004.10.013
- Schutte, N. S., Malouff, J. M., Hall, L. E., Haggerty, D. J., Cooper, J. T., Golden, C. J., & Dornheim, L. (1998). Development and validation of a measure of emotional intelligence. *Personality and Individual Differences*, 25(2), 167-177. doi:10.1016/S0191-8869(98)00001-4
- Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., & Chapman, M. (1992). Development of concern for others. *Developmental Psychology*, 28(1), 126-136. doi:10.1037/0012-1649.28.1.126

## **Appendix 3.A**

### **Methods for pre-testing patient transcripts**

#### **Study 1: Mturk**

##### **Participants**

Thirty-four participants were recruited from Mturk to read and rate six patient transcripts. No identifying or demographic information was collected from these participants, but all were located in the United States.

##### **Stimuli**

The videos of the three original distraught patients (all female) and three of the resilient patients (one male and two female) were transcribed. Emotional cues were added to the transcripts to capture the non-verbal emotions patients displayed, and any content related to their age or health background was removed. The gender cues for the male patient were changed to female, so that pre-testing participants read all six transcripts as female patients. The word counts (including the emotion cues and interviewer questions) were similar for the two types of patients, *distraught*  $M = 258.3$ , *resilient*  $M = 232.3$ , although the transcripts for the *distraught* patients were slightly longer.

##### **Procedure**

Participants read all six female patient transcripts online and after each transcript answered questions about the patient and their emotional responses on a Likert scale from 1 (not at all) to 7 (extremely). These questions were taken from Chapter 2 and included their perceptions of the patient (how distraught, resilient, distressed, and happy the patient seemed) as well as their emotional responses to the patient (the extent to which they felt empathic, distressed, pleased, and horrified). They also rated how sick the patient seemed and how much they liked the patient.

## **Analysis**

Ratings were first visually assessed across all of the transcripts to confirm that the three patients within each group were rated similarly. Then, ratings were averaged across transcripts within the same patient type and compared using t-tests. Finally, these ratings were compared to the ratings from Chapter 2 for each patient type using independent Welch t-tests. Bonferroni-corrected p value cut-offs were used for each set of t-test comparisons.

## **Results**

### **Patient transcripts**

Visual inspection of the data indicated that all three *distraught* patients appeared more distressed and distraught than the three *resilient* patients, while the *resilient* patients seemed more resilient and happy (see Table 3.A.1 for means). T-tests confirmed these observations differed across the patient types,  $t_s(33) > 8.04$ ,  $p_s < .001$ . For emotional responses, the three *distraught* patients seemed to evoke more distress and slightly more

empathy and feelings of horror, while participants felt more pleased in response to the three *resilient* patients. T-tests between *resilient* and *distraught* patients confirmed these differences,  $t(33) > 4.59, p < .001$ . Although the three *distraught* patients had the highest ratings of perceived sickness, the ratings for the *resilient* patients were more variable. However, t-tests collapsed across patient types revealed that *distraught* patients were indeed rated as sicker than *resilient*,  $t(33) = 11.78, p < .001$ . Participants also seemed to like all six patients equally well, and t-tests confirmed that the *resilient* patients were only liked marginally more than the *distraught* patients,  $t(33) = -1.78, p = .084$ . All tests but the difference in liking remained significant with Bonferroni corrections for multiple comparisons.

**Table 3.A.1. Mean ratings by display type and individual patient for the transcripts.**

Perceived Patient Emotion						
	Patient ID	Distressed	Happy	Distraught	Resilient	Sick
Distraught	1	5.77	1.77	5.57	3.46	5.8
	11	5.79	1.56	5.62	2.62	5.44
	5	5.91	2.2	5.57	3.57	5.63
Resilient	14	2.91	4.47	2.38	5.79	4.06
	4	2.56	4.29	2.26	5.12	3.62
	9	2.56	4.59	2.21	5.56	4.03
Responses to Patient						
	Patient ID	Empathic	Distressed	Pleased	Horrorified	Like
Distraught	1	5.89	4.77	1.63	3.66	5.6
	11	5.74	4.85	1.41	3.18	5.24
	5	5.83	4.91	1.69	3.34	5.54
Resilient	14	5.32	2.85	3.21	2.03	5.91
	4	4.68	3.12	2.85	1.91	5.44
	9	5.15	2.91	3.21	1.97	5.74

### Comparing transcripts to video ratings

Although the patients types were rated similarly in the transcripts as they were in the videos, it is important to assess whether the magnitude of the ratings were similar across modalities, as watching the videos may produce stronger responses than just reading transcripts of the interview. To test this, ratings from the *distraught* and *resilient* patient videos used in Chapter 2 were compared to the transcript ratings (see Table 3.A.2 for mean ratings by video from Chapter 2).

**Table 3.A.2. Mean ratings from Chapter 2 by display type and individual patient for the videos.**

		Perceived Patient Emotion				
	Patient ID	Distressed	Happy	Distraught	Resilient	Sick
Distraught	1	5.61	1.58	5.48	2.77	4.89
	5	6.19	2.22	5.69	2.93	5.07
Resilient	4	1.67	5.2	1.81	5.1	2.31
	9	2.07	5.14	2.16	5.56	3.75
		Responses to Patient				
	Patient ID	Empathic	Distressed	Pleased	Horrificed	Like
Distraught	1	4.98	4.18	1.49	2.6	4.35
	5	5.53	4.78	1.51	3.17	4.78
Resilient	4	3.95	1.67	3.8	1.25	5.32
	9	4.79	2.21	3.7	1.58	5.67

The *distraught* patients were perceived as similarly distressed, distraught, and happy in the videos and transcripts,  $ts(124.79) < 0.45$ ,  $ps > .65$ , but were perceived as more resilient in the transcripts than in the videos,  $t(133.89) = 3.26$ ,  $p = .0014$ .



Compared to the videos, *resilient* patients were seen as similarly resilient and distraught in the transcripts,  $ts(116.04) < 1.35$ ,  $ps > .18$ , but were rated as more distressed,  $t(113.42) = 3.79$ ,  $p < .001$ , and less happy,  $t(126.13) = -3.50$ ,  $p < .001$ , in the transcripts than in the videos. Emotional responses to both types of patients were similar in videos and transcripts,  $ts(97.89) < 3.07$ ,  $p > .003$  (Bonferroni-correct  $p = .0025$ ), with the exception of *resilient* patients eliciting more distress in the transcripts compared to the videos,  $t(103.21) = 4.41$ ,  $p < .001$ . Both types of patients were perceived as sicker in the transcripts than in the videos,  $ts(128.93) > 3.76$ ,  $ps < .001$ , and *distraught* patients were liked more in transcripts than in the videos,  $t(153.15) = 5.01$ ,  $p < .001$ .

## Study 2: Subject Pool

The Mturk testing demonstrated that the six transcripts were rated consistently as *distraught* and *resilient*. However, in order to evenly present the patient types as males or females within each participant, two more transcripts needed to be added. The fourth original *resilient* patient was transcribed (male), but there were only three original *distraught* patients. In order to maintain the integrity of the stimuli (and thus not construct a fake transcript), a fourth patient video was selected who was also highly distressed and distraught based on previous ratings (Preston et al., under review). This female patient was originally *wistful* and was transcribed with extra emotional cues to better match the transcripts of the other *distraught* patients. For this study, the main goal was to confirm that the extra two patients were rated similarly to the other *distraught* and *resilient* targets. To do this, the female transcripts of these new patients were compared to the other *distraught* and *resilient* female targets (rated in Study 1). A second goal of this

study was to demonstrate that the male targets were seen as believable. To do this, ratings of the *distraught* male targets were compared to the *distraught* female targets to determine if the male targets were seen as sick, in need, distressed as female targets, or if they evoked less empathy from participants.

## **Participants**

Twenty-nine students (13 males) from the Introductory Psychology Subject pool at the University of Michigan were tested as a group in a computer lab, with individuals seated at every other terminal. Participants completed a short one block of the priming task (Chapter 3) and then read and responded to the eight patient transcripts.

## **Results: Distraught Patient**

Ratings of the four female *distraught* transcripts were compared using a repeated measures (RM) ANOVA to determine whether ratings differed across patients. Patients were perceived as equally distressed and distraught,  $F_s(3, 26) < 1.49, p_s > .24$ , and did not differ in the amount of distress they evoked in participants,  $F(3, 26) = 0.26, p = .85$ , indicating that the new *distraught* patient was similar to the original *distraught* targets. However, the added patient was perceived as the most happy out of the *distraught* targets,  $F(3, 25) = 3.59, p = .03$ , and although perceptions of resilience tended to differ among these targets,  $F(3, 26) = 2.94, p = .052$ , the differences were among the original patients, not the added one.

## **Results: Resilient Patient**

Ratings of the four female *resilient* transcripts were also compared using RM ANOVA. While these patients were perceived as equally happy,  $F(3, 27) = 0.48, p = .7$ , they differed in how resilient they appeared,  $F(3, 27) = 4.45, p = .01$ , with the added patient perceived as the least resilient. The patients were perceived as equally distraught,  $F(3, 27) = 0.26, p = .85$  and elicited the same amounts of distress in observers,  $F(3, 27) = 0.26, p = .86$ , but the new patient was seen as the most distressed,  $F(3, 27) = 4.48, p = .01$ .

### **Comparing the two new patients**

Although the two added transcripts were similar to the original patients, it is important to test that these new patients, as the happiest *distraught* patient and the least resilient and most distressed *resilient* patient, still differed from one another. T-tests confirmed that the *distraught* patient was perceived as more distressed and distraught, and less resilient than the *resilient* patient,  $ts(29.79) > 3.95, ps < .001$ . These results suggest the added patients were still rated differently from one another, even if they were not exactly the same as the other patients in their display type.

### **Results: Male versus Female *Distraught* targets**

To determine whether male *distraught* targets were seen as believable, ratings of felt empathy, perceived sickness, need, and distress for *distraught* targets were compared by gender. Male patients evoked similar amounts of empathy,  $t(28) = -1.29, p = .21$ , and were perceived as similarly sick,  $t(28) = -1.42, p = .16$ , distressed,  $t(28) = 0.46, p = .65$ , and in need of emotional support as female targets,  $t(28) = -0.29, p = .77$ . Although need

for practical support was similar, males ( $M = 4.93$ ) actually tended to be perceived as in more need than females ( $M = 4.48$ ),  $t(28) = -1.71$ ,  $p = .098$ . Because male patients were not seen as less sick or in less need, it suggests the male *distraught* patients were perceived believably.

## Conclusions

Overall, the perceptions of and emotional responses to *distraught* compared to *resilient* patients were the same using the transcripts as were found in Chapter 2 using videos. Furthermore, many responses were similar in magnitude between transcripts and videos. *Distraught* patients appeared more resilient and were liked more in the transcripts compared to the videos, but these ratings were still lower than for *resilient* patients (although liking *resilient* patients more did not survive the Bonferroni correction). *Resilient* patients were perceived as more distressed and less happy, and elicited more distress in transcripts compared to videos, but again, these values were still lower than those associated with *distraught* patients. The added patients in Study 2 were also representative of their distinct patient types, and male *distraught* patients appeared to be equally believable as females. Thus, it appears that the emotions expressed and elicited by the different patient types are preserved in the transcripts, and can be used as patient stimuli comparable to the original videos.

### **Appendix 3.B**

#### **Alternate ways of addressing word frequency**

The words used in the priming task were selected based on their representativeness of the constructs the study was measuring (sadness, female-stereotypes, male-stereotypes). Therefore, although the subset of words were adjusted to match in general length, they were not matched on frequency, as the female and male stereotype words were much more frequent than the sad adjectives. Regressing out the effects of word frequency on RT for each participant ended up being the optimum way to deal with this confound. An alternate strategy for controlling the effects of word frequency and the reasons why it was not preferred are presented below.

#### **Alternate method**

The method that was used to calculate association scores was to subtract the RT to the words of interest (sad, female stereotype, male stereotype) from the RT to the neutral words within a given block with the same prime (MAN or WOMAN). Because the frequency of words varied significantly across categories ( $F(3, 207) = 173.96, p < .001$ ), without correcting for word frequency, these scores would be biased towards differences in frequency, rather than the effects of priming. An alternate strategy is to subtract RT to the words of interest after the relevant prime from the RT to the same category words after the irrelevant prime (for example, subtracting the RT to sad words after the

WOMAN prime from the RT to sad words after the MALE prime). This score would provide an indication of how much better the relevant prime is at activating the concept and facilitating recognition compared to the other, presumably insufficient prime. One potential benefit of this strategy is to have the RT subtractions within categories of words; although frequency varied within word categories, it was not as extreme as across-category variations. Words were not repeated in the task, resulting in different words presented with each prime. However, it is likely the random selection of words resulted in the frequencies of words being equal across blocks (and therefore primes).

Comparing the frequency of words presented after each prime revealed that, at least across participants, frequency was equivalent across the two blocks,  $t(279) = -0.52$ ,  $p = .60$ . Additionally, within each word category, the frequency across blocks was also equivalent across participants,  $ts(69) < 1.47$ ,  $ps > .15$ . However, because the primary interest was in individual differences in these association scores, the fact frequency did not vary between blocks across all participants is not particularly relevant if individual differences in word frequency between the blocks reliably varied with the association scores. Indeed, these sadness scores were positively correlated with the difference in word frequency in each participant's block,  $r(68) = .35$ ,  $p = .003$ . Participants who saw more frequent sad words during the WOMAN block compared to the MAN block were even faster to sad words after the WOMAN prime compared to the MAN prime, suggesting that these scores were still confounded by difference in word frequency. Similar correlations were found for female-stereotype scores,  $r(68) = .28$ ,  $p = .018$ , but not for male-stereotype scores,  $r(68) = 0.004$ ,  $p = .97$ .

Because these scores did not effectively remove the confound of word frequency, this method was not used as an alternative strategy to regressing out word frequency.

## **Chapter 4: The Effects of Remitted and Current Depression on Empathy for Another's Distress**

### **Abstract**

People vary widely in the ways they express need, which can elicit a range of responses in observers, from distress to empathy and helping (Preston, Hofelich, & Stansfield, under review). However, little is known about what causes these divergent responses to another's need. Perception-action models of empathy (Preston & de Waal, 2002) predict these differences are driven by observers' specific internal representations, which are formed through their own past experiences with emotional situations. Therefore, empathy for distressed, upset patients should be affected by differences in observers' past experience of negative affect. The current study compared psychophysiological, empathic, and prosocial responses to targets experiencing hospitalization in participants who varied in past and current experience of depression. While past experience of depression was associated with distinct perceptions of negative emotion and greater helping when patients were perceived as similar, current depression experience lead to decreased empathy and helping specifically for sad, distressed patients. The current results highlight the importance of studying observer-target interactions in the prosocial response.



## **Introduction**

Seeing another person in distress often elicits feelings of empathy in an observer, but it can also elicit more negative responses, like personal distress (Batson, Fultz, & Schoenrade, 1987), anger and horror (Preston et al., under review) and even aggression (Zahn-Waxler, Radke-Yarrow, Wagner, & Chapman, 1992). Such responses can inhibit prosocial action towards the person in need (target) and are influenced by the nature of the situation (e.g., Eisenberg & Fabes, 1990), trait tendencies of the observer (Davis, 1983; Eisenberg & Fabes, 1990; Eisenberg, Wentzel, & Harris, 1998), and instructions to take the target's perspective (Lamm, Batson, & Decety, 2007). However, the vast majority of empathy paradigms have characterized these responses using a single prototypical target (e.g., Batson, Eklund, Chermok, Hoyt, & Ortiz, 2007; Batson, Quin, Fultz, Vanderplas, & Isen, 1983; Batson et al., 1996; Coke, Batson, & McDavis, 1978; Graziano, Habashi, Sheese, & Tobin, 2007; Kogut & Ritov, 2005; Maner et al., 2002; Marsh, Kozak, & Ambady, 2007) or a few possible scenarios where the content is manipulated to increase or decrease observers' empathic response (e.g. Batson et al., 2007; Eisenberg & Fabes, 1990; Jackson, Brunet, Meltzoff, & Decety, 2006). As a result, little is known about how these divergent responses vary across individuals in everyday empathic encounters.

In real-life situations, and in contrast to prototypical empathy scenarios, targets cope with their need differently, even when their circumstances are similar. For example, in a sample of 14 chronically or terminally ill hospital patients, people expressed their emotions in various ways when describing their struggles and quality of life with their illness (Preston et al., under review). Even in this relatively small sample of patients,

distinct “types” of patients were found using participants’ ratings of the patients’ expressed emotions, ranging from upset and *distraught*, to more reserved and *reticent*, to upbeat and *resilient*<sup>3</sup>. Despite being in similar life situations, these patients displayed different emotions and elicited divergent responses from participant observers, with individuals differing in their preferences to help one type of patient over the others. For example, while many participants offered *distraught* patients the most help, an equivalent number of people preferred to help the more positive *resilient* patients, even though they were perceived as less sick (Preston et al., under review). The less expressive, *reticent* patient was disliked more than the other patients and was offered less help than all other types (Preston et al., under review).

The perception-action model of empathy (PAM; Preston & de Waal, 2002) assumes that understanding another’s emotion partially relies on the same neural substrates that give rise to the first-hand experience of that emotion. Individual differences in responses to a target are thus driven by the observer’s internal representations for the target’s emotional state and situation, which are formed through the observers’ own prior experience. Because of this, observers who have experienced events or emotions similar to that of the target should activate richer personal representations for that state, promoting understanding, empathy, and tailored responses to the target’s need (Preston & Hofelich, 2012; Preston & de Waal, 2002).

---

3 *Distraught* patients (n=3) were upset and distressed, often breaking into tears. *Resilient* patients (n=4) were positive and upbeat, but also talked about their struggles. *Sanguine* patients (n=3) were less emotional than *distraught* or *resilient* targets, did not discuss major problems, and tended to make jokes. A single patient was *reticent*, giving only brief responses (e.g., single words such as “fine” or “none”), and not expressing emotion overtly. *Wistful* patients (n=3) appeared slowed by their illnesses, and talked about fears of dying, but were not as overtly distressed as the *distraught* targets.

While similar life experiences between observer and target often increase empathy and helping (e.g., Barnett, 1984; Barnett, Tetreault, Esper, & Bristow, 1986; Batson et al., 1996), shared experience does not always make an observer better able to accurately understand what the target is feeling. For example, observers who have just given birth felt the most empathic concern for a target who was a new mother, but were no better at intuiting what the target was feeling than observers who had never been pregnant (Hodges, Kiel, Kramer, Veach, & Villanueva, 2010). Furthermore, even effects of similar experience on empathy are modulated by the emotional display of the target, as observers who had themselves experienced hospitalization for illness did not feel equally empathic towards all of the patient targets (Preston et al., under review). According to the PAM, these findings arise because the specific content of the mental representation matters, which varies with the context of the situation and how one appraises their experience (Preston & de Waal, 2002; Preston & Hofelich, 2012). For example, a childless young woman who ends a short marriage amicably has a very different experience of divorce than that of a stay-at-home mother who leaves a husband of 15 years after discovering an extramarital affair. These diverse contexts would lead the two women to have quite different mental representations of divorce, even though they both experienced it.

Because even “similar” life experiences can vary widely across individuals, an alternate approach is to compare responses across participants with similar *emotional* experiences. The displays of extreme distress and sadness in *distraught* targets produced the most divergent responses across participants, therefore, an observer’s past experience with intense negative affect may be relevant for predicting responses to these patients.

Depression is a salient example of such an emotional experience, and although the presenting symptoms differ across individuals (such as undereating/overeating, not sleeping/sleeping too much), intense negative mood is a consistent and core experience in all cases of depression.

Despite being one of the most common mental illnesses (Ault et al., 2007), the literature on how depression affects empathic responding is relatively sparse and presents mixed results. Studies on the effects of maternal depression suggest it may decrease empathy, as children of depressed mothers are more likely to be depressed in their first year of life (Field, 1995), and are less likely to respond empathically to another distressed infant or to the distress of their mother than children with non-depressed mothers (Jones, Field, & Davalos, 2000). However, other studies suggest empathy may not be impaired, as adults with depression score higher than non-depressed controls only on the personal distress subscale of the Interpersonal Reactivity Index (Davis, 1980), while their scores on empathic concern and perspective taking are similar to non-depressed individuals (O'Connor, Berry, Weiss, & Gilbert, 2002). Furthermore, the negative self-perception associated with depression does not seem to generalize to perceptions of others. For example, although people with depression show an automatic, negative bias when engaging in self-referential thought or self-judging, they do not show such bias when judging characteristics or performances of others (Bargh & Tota, 1988; Sweeney, Shaeffer, & Golin, 1982). Some evidence even points to a positive relationship between empathy and depression tendencies, as high trait empathy predicts an individual's tendency to develop depressive symptoms to distressing life events (Gawronski & Privette, 1997; O'Connor, Berry, Lewis, Mulherin, & Crisostomo, 2007).

According to the PAM, depression experience should not affect empathic responding across all targets, but rather should be dependent on the target's emotional display. Supporting this, pilot data from a previous study found that women with depression responded with more empathic concern to *distraught* targets and less to the unexpressive *reticent* target than non-depressed controls (see Appendix 4.A). Depressed participants also had more differentiated responses to the different patients within the *distraught* type, while controls responded more similarly across the three *distraught* patients. This was demonstrated in the cluster analysis; instead of having one group of distressed, upset patients (the original *distraught*), the grouping solution for depressed participants divided the *distraught* patients into two groups. This suggests depressed and non-depressed participants responded to the negative affect expressed by these patients differently, with participants who had experienced depression differentiating between displays of distress in *distraught* targets. Because experience provides a richer representation of the specific situation or state experienced (Preston & Hofelich, 2012; Preston & de Waal, 2002), this increased differentiation likely reflects elaborated representations of sadness or distress that have formed as result of depressed women's extensive experience with negative affect.

Due to time limitations, the pilot study did not collect data on how participants perceived the patients, only their emotional responses to them. Also, all depressed participants had been clinically diagnosed and were medicated for depression, while all controls were non-depressed and had no history of depression. Therefore, it is impossible to tell from the pilot data whether the increased differentiation between displays of *distraught* targets reflected differences in how negative affect was represented (and thus

expected in all participants with past depression experience, regardless of current depression status), or whether it was due to other, state-specific components of currently felt depression (and thus expected only in participants currently experiencing depression). It is also important to replicate these findings, as recent work on emotional granularity suggests people currently experiencing depression actually report *less* differentiation in the negative emotions they feel compared to non-depressed controls (Demiralp et al., in press). The decreased differentiation (granularity) is thought to contribute to their inability to adaptively respond to their negative emotional state and regulate their emotion, as these tasks are easier to do when an emotional state is more specific and has a causal object (e.g., Barrett et al., 2001). However, it is unclear whether this decreased granularity is specific to their felt emotional states across experiences (i.e., their *experiences* of emotion are just less differentiated), rather than how emotion is represented and accessed when perceiving it in others (i.e., their actual *representations* of emotional concepts are less differentiated). Because the clustering methods used in the current study probe differentiation among participants' perceptions of the targets' emotions, it is likely this task better reflects their representations of emotional concepts, rather than the qualities of their emotional experiences. If the findings from the pilot study replicate in a larger sample and again suggest depressed participants differentiate between the negative states perceived in others, these seemingly inconsistent results may also reflect another distinction in the ways depressed individuals judge or represent their own versus others' states.

In order to understand how depression experience affects empathic and prosocial responses to various targets, the current study examined emotional, psychophysiological,

empathic, and helping responses to the different patient types in participants with and without a past history of depression. To examine differences between past and current depression experiences, we specifically recruited participants with and without a prior history of depression, and measured the current experience of depressive symptoms in all participants using a depression scale (Beck Depression Inventory, BDI-II; Beck, Steer, & Carbin, 1988). Although the BDI is a valid measure of depression in college studies (Bumberry, Oliver, & McClure, 1978), the current study used this scale to assess individual differences in the current experience of depression symptoms, rather than as a criteria for depression.

Because depression can affect several aspects of cognition, such as attention and memory (e.g., Coyne & Gotlib, 1986), differences in the perceptions of or emotional responses to targets between participants may be confounded with more general differences in attention to negative information. While some studies have shown that attentional biases in depression are limited to current, rather than remitted, experience (McCabe & Gotlib, 1993; McCabe, Gotlib, & Martin, 2000), others have found lasting attention and executive impairments that are similar in current and remitted depression (Fritzsche et al., 2010; Joormann & Gotlib, 2007; Paelecke-Habermann, Pohl, & Lepow, 2005). To test and control for this, participants also completed an Intrusive Cognitions task (“emotional Stroop”; McKenna & Sharma, 1995) that assessed differences in the tendency to be distracted by emotional information.

To determine whether current or remitted depression experience affects the way emotion is perceived in others, participants rated the emotions they perceived in each patient, along with their own emotional and prosocial responses. Clustering methods

were used to group patients with similar emotional displays based on ratings of perceived emotion, allowing us to assess whether participants with different depression histories similarly differentiate between the negative displays of *distraught* targets. Based on the PAM and pilot data, we predicted that participants with a depression history would make greater distinctions between the emotions expressed by the *distraught* targets, resulting in clustering solutions that separated these patients into distinct groups, while participants without a depression history would not. To further determine whether people will spontaneously generate the categories of patients previously found using cluster analysis, participants also sorted the patients into groups based on their affect and communication style. The numbers and types of groups generated were compared across participants with and without depression histories. The card-sorting task was expected to mirror the clustering solution, with depression history participants creating more groups that distinguished between the sad displays of targets. In terms of empathic responses, both past and current depression experience was expected to increase empathic responses to the most similar targets: the *distraught* patients.

## **Methods**

### **Participants**

Sixty-three females (age  $M = 18.8$ , range = 18-21) participated in this study for course credit. Thirty-one participants had been previously diagnosed with depression (12 were on medication for depression) and 32 had no history of depression. Participants were tested individually in the lab and gave full informed consent. All procedures were approved by the University of Michigan Institutional Review Board.



## **Overview**

Participants were seated at a Dell desktop PC and were attached to psychophysiological electrodes (including heart rate, skin conductance, and two respiration belts). They completed an Intrusive Cognitions task to measure individual differences in attentional capture (McKenna & Sharma, 1995), and then watched the 14 patient interviews (Preston et al., under review; Appendix 4.A). After each video, participants rated the emotions perceived in the patient and felt in themselves, as well as their empathic responses to the patient. Participants also had the opportunity to donate money, in the form of tokens, to each patient. After watching all the videos, participants were asked to group photo cards of the patients according to the patients' communication and affective styles. Afterwards, participants completed the Beck Depression Inventory (BDI; Beck et al., 1988) and answered post-experiment questions about their experiences with illness and depression.

## **Intrusive Cognitions Task**

Participants completed an Intrusive Cognitions task to determine whether attentional capture to emotional stimuli differed with depression history or current depressive symptoms, and whether these differences affected perceptions of targets. To ensure that participants were not primed by the emotions of the patients, this task was always completed before viewing the patient videos. Five positive, negative and neutral words (matched on frequency and length; McKenna & Sharma, 1995) appeared in blue, red, green, or yellow font and were presented in separate blocks. Each of the five words

was displayed once per color, producing 20 trials per block. Participants responded to the color of the word with their index, middle, ring, and pinky finger on their dominant hand.

After removing error trials and reaction times (RT) greater than 3 standard deviations (SD) of each participant's mean (as in Hofelich & Preston, 2011), attentional capture scores were calculated. Positive and negative bias scores were created for each participant by subtracting median RT to neutral stimuli from the median RT the emotional stimuli (positive and negative words, respectively).

## **Stimuli**

The 14 videos contained interviews with patients who had a variety of serious chronic or terminal conditions. These videos have been used previously to study variation in the empathic response (see Preston et al., under review). Each video showed only the patient facing the camera in their hospital room. Videos were edited to include the same four questions and answers in the same order with a brief fade between each: 1) What has been the impact of your illness on your quality of life?; 2) What are your health-related worries?; 3) What in your life are you the most proud of?; and 4) What has been the hardest thing for you to cope with related to your illness? The specific illness was not mentioned in the interviews, and subjects were unaware of patients' prognoses.

## **Questionnaire Data**

After each video, participants recorded their emotional responses to the videos using a seven-point Likert scale from 1 (very little) to 7 (extremely). They first rated how similar they felt to the patient (both in terms of the patient's situation and their displayed

affect), how much practical support (e.g. getting prescriptions, changing sheets, watering plants, grocery shopping) and emotional support (e.g. talking to them, giving advice, soothing, spending time with them) they think the patient needed and were willing to give, and how much they liked the patient. They then rated 29 emotional adjectives describing the emotions they felt (taken from Batson et al., 1997, with the addition of *engaging*, *likeable*, and *entertaining* to better sample positive emotions), followed by rating the same adjectives describing the emotions the patient felt (order of “self” and “other” ratings were counterbalanced across participants). Finally, participants were given five tokens after each patient and had the opportunity to donate any number of these tokens to the patient. Participants were told the tokens would be converted into cash at the end of the study (at an unknown exchange rate), and that they would keep the cash from tokens they did not donate.

After the experiment, participants answered questions about their experiences with depression, including when they were diagnosed, if they were currently experiencing depression, and if not, how long they had been in remission and how easily they could remember what it was like to be depressed. They then filled out the BDI scale.

### **Physiological Data**

Physiological responses were recorded during the videos, and compared to a 30s baseline measurement before the start of each video. Heart rate was collected using lead II EKG, with one electrode attached inferior to the costal margin and the other anterior to the sternocleidomastoid muscle. Galvanic skin response (GSR) was measured using two electrodes attached to the hypothenar area on the palm of the non-dominant hand.

Respiration was measured as the summed signal across two belts, one attached around the participant's torso and another around the upper chest. Data were sampled at a rate of 500 Hz using a Biopac MP150 system (Biopac Systems, Santa Barbara, California) and was recorded and preprocessed with AcqKnowledge III software for Mac (Biopac Systems). Responses were calculated as changes from baseline and were standardized within each participant, across videos, before analysis.

### **Sorting Task**

After watching the videos, participants were given 14 cards, one with a picture of each patient, and were asked to sort the patients into groups based on their communication and affective styles, placing similar patients in the same group. This was done to determine the extent to which groups found with statistical clustering analysis, both in this study and previously (Preston et al., under review), matched explicitly self-generated categories formed by participants. The number and types of categories created were also compared between participants with and without depression histories.

### **Statistical Analysis**

#### *Clustering Analysis*

The cluster analysis served several purposes. First, it allows us to reduce comparisons of emotional and prosocial responses to a few different “types” of patients who are marked by similar emotional displays, rather than the entire sample. Secondly, the clustering results can provide a data-driven, qualitative description of how participants with and without a depression history perceived the emotional displays of the

patients. The final clustering solution provides information about the specific patients that are grouped into a “type,” as well as how these types are characterized, as it provides mean values for the emotions expressed in each cluster. The cluster results were also compared to the common card-sorting categories made by participants to provide information about the extent to which the clustering solution mirrors self-generated categories.

All 29 “other” rated adjectives were used to cluster the patients into groups with similar emotional displays. This differed from prior studies (Preston et al., under review), in which patients were clustered using the three factor scores that resulted from factor analysis of the other rated adjectives. In this study, preliminary factor analysis of these adjectives produced three factors scores that captured only about half of the variance in the ratings<sup>4</sup> (56.9%). Because these scores would not provide a very complete picture of how the patients’ emotions were perceived by the participants, all of the other rated adjectives were used. Clustering was done separately for controls and participants with a depression history.

The clustering was performed in two steps in order to best characterize 1) the number of clusters that exist in the data and 2) which patients best fit into each cluster. First, a hierarchical clustering method (Ward, 1963) was used to determine the optimum number of target groups given the pattern of other ratings across the videos. This clustering method begins with each patient in a separate cluster, and successively joins clusters that are next “most similar” (based on sums-of-squared deviations), until all patients are the in the same cluster. The agglomerative coefficient was used to determine

---

<sup>4</sup> The factor scores from Preston et al. (under review) explained about 70% of the variance.

the point at which joining additional clusters resulted in a large jump in the sum-of-squared deviations. This point was used to identify the optimum number of clusters.

Next, k-means clustering was used to group the patients into the number of clusters identified by the hierarchical procedure. This relocation procedure determines which patient fit best into each cluster, but requires that the number of clusters are known (which is why the hierarchical method was completed first). This method first partitions the data into the specified number of clusters, and then iteratively modifies the membership of each cluster until both the homogeneity within each cluster and the differences between clusters are maximized. For each group, k-means clustering was run an additional time to observe which patient clusters would divide if an extra cluster was allowed into the solution.

Once common target “types” were found with the clustering methods, they were first characterized by their emotional displays, and then the emotion and prosocial responses were compared across each type.

### *Factor Analysis*

Although all the other rated adjectives were used to cluster the patients into “types,” data reduction measures were still performed in order to reduce the number of comparisons needed to characterize perceptions of and emotional responses to the patient types. Factor analysis was used to condense the 29 adjectives rated for other and self emotion into other and self emotion factor scores, respectively. Principle components analysis (PCA) was used on the correlation matrix across all adjectives and videos. The Scree plot was used to select the number of factors that were extracted and Varimax-rotated. The saved PCA coefficients were used for each factor score. PCA was first

performed across all participants, and then was done separately for participants with and without a depression history to confirm that the partitioning was similar.

### *Data Analysis*

All variables were averaged across patients in each cluster (target type) before analysis. The other emotion factors were used to characterize the target types, along with ratings how sick the patients seemed, and how much practical and emotional support participants thought they needed. The self scores, ratings of liking, and psychophysiological measures were used to assess emotional responses to the patient types. Prosocial responses were assessed using ratings of practical support given, emotional support given, and actual token donations. All measures were compared using mixed Analysis of Variance (ANOVA) to determine whether responses differed across target type, between participants with and without depression history, and with current experience of depression symptoms (BDI score). Multiple regression was also used to determine whether prosocial responses to the different clusters could be predicted by similar life and emotional experiences. An alpha level of .05 was used in all tests.

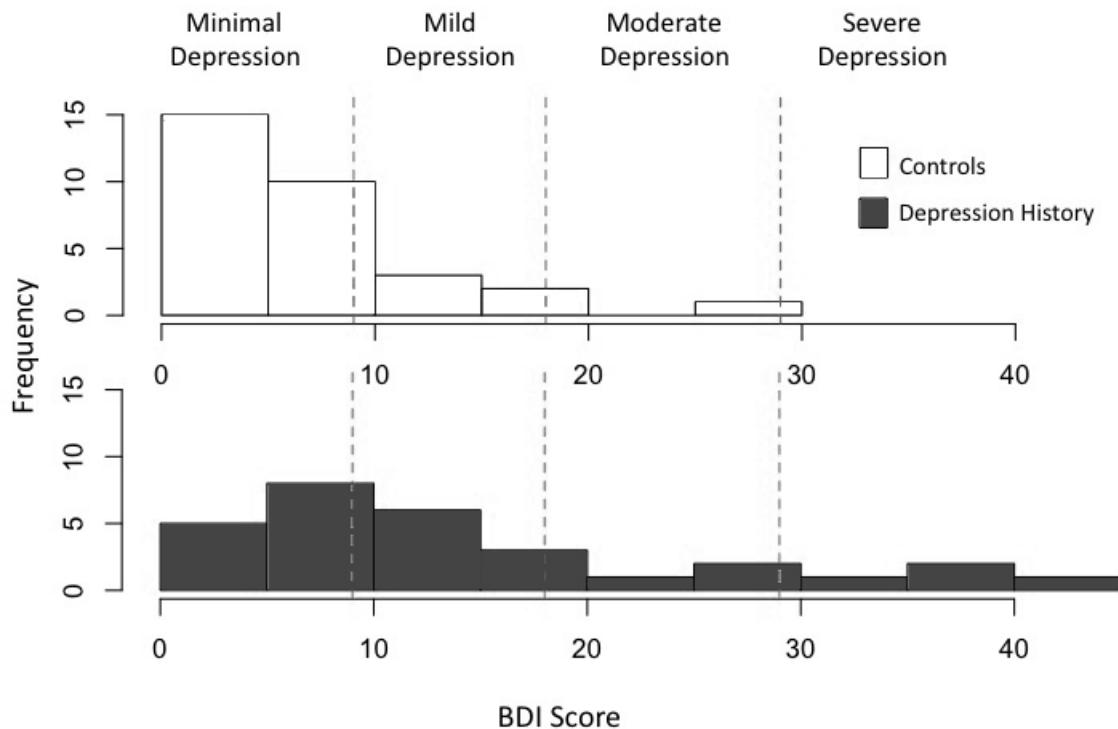
## **Results**

### ***Group Differences***

Participants with a history of depression had higher scores on the BDI ( $M = 15.12$ ) than participants without depression history ( $M = 6.79$ ),  $t(58) = -3.49$ ,  $p = .001$ . However, the range of scores was wide in both groups, Depression History: 1 – 43, Controls: 0 – 26 (see Figure 4.1), suggesting the current experience of depressive symptoms was not confined to participants with a depression history. Among the

depression history group, the average time since depression diagnosis was 3.2 years (range 0-10), and all but nine participants reported that they were not currently depressed. Corroborating self-reported depression status, the BDI scores for currently depressed participants ( $M = 28.11$ ; range: 12 – 43) were higher than for depression history participants who were not currently depressed ( $M = 9.28$ ; range: 1 – 21),  $t(9.76) = 4.91$ ,  $p < .001$  (scores among participants who did not report being currently depressed did not differ by their depression history,  $t(43.65) = -1.44$ ,  $p = .16$ ). The length of time participants had been remitted ranged from 2 months to 6 years, with the average being 1.2 years; participants reported they could still remember the experience ( $M = 5.43$  on 7-point scale).

**Figure 4.1. Histograms of BDI scores by group with standard cut-offs for depression severity**





History groups did not differ in their tendencies to be distracted by positive,  $t(61) = 0.90, p = .37$ , or negative,  $t(61) = 0.22, p = .83$ , emotional stimuli in the Intrusive Cognitions task. Current experience of depression symptoms (BDI score) also did not predict tendencies for positive or negative attentional capture,  $rs(59) < .06, ps > .64$ . Therefore, attentional capture was also not included as a covariate in further analysis. Because there were no differences between depressed participants who were and were not taking medication (BDI,  $t(15.27) = -1.48, p = .16$ ), medication status was also not included in the analysis<sup>5</sup>.

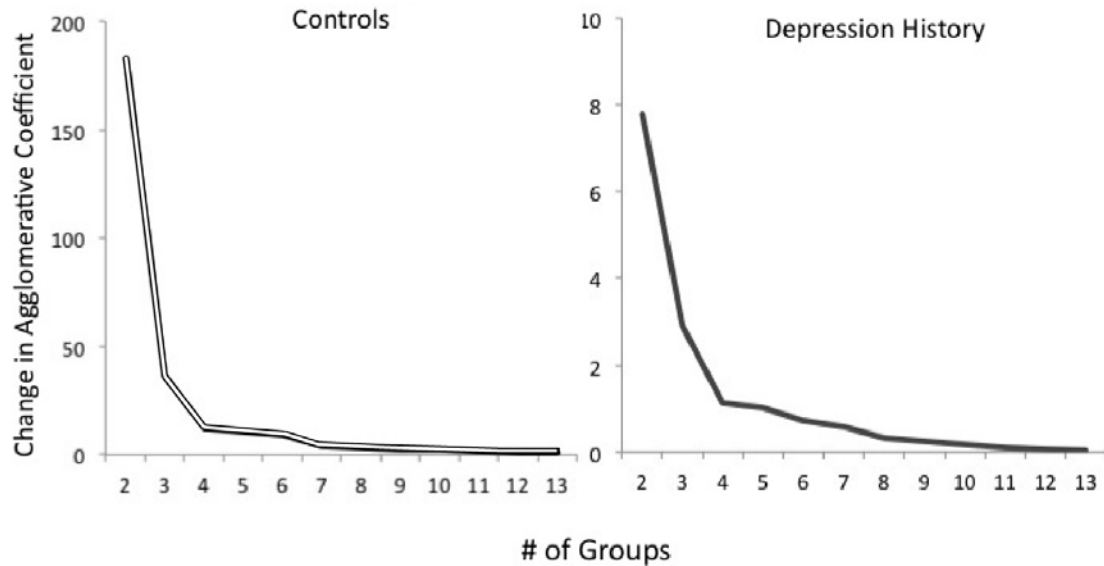
### ***Within-Group Cluster Analysis - Controls***

The agglomerative coefficient from the hierarchical clustering analysis indicated a three-cluster solution best fit the data, as the steepest drop off occurred when three groups were formed (see Figure 4.2; right). From the k-means clustering solution, the first cluster included the previously classified *distraught* patients (Preston et al., under review), who were characterized similarly (see Table 4.1). The top six highest rated adjectives for the *distraught* cluster were *sad, upset, worried, concerned, troubled, and distressed*. The second cluster was similar to the original *resilient* group of patients, but included two other patients previously classified as *sanguine* (a separate, more positive group of patients in the prior study). These patients were rated as highly *likeable, warm, engaging, softhearted, happy, and compassionate*. The last cluster combined the previously classified *wistful* patients, the *reticent* patient, and the remaining *sanguine* patient. These patients (referred to here as *reserved*) were not as distressed as the *distraught* patients, but were less engaging and happy than the *resilient* patients.

---

<sup>5</sup> If medication status is included in the analysis, its effects are also non-significant, all  $p > .13$ .

**Figure 4.2. Change in agglomerative coefficient for Ward's clustering when additional groups are added. Displayed by depression history group.**



**Table 4.1. Comparison of videos in clustering solutions by depression history and current depression.**

Three Cluster Solution				
	1	2	3	
	Distraught	Resilient	Reserved	
Control	1, 5, 11	2, 3, 4, 9, 10, 14	6, 7, 8, 12, 13	
Depression History	1, 5, 11	2, 3, 4, 7, 9, 10, 14	6, 8, 12, 13	
Four Cluster Solution				
	1	2	3	4
	Distraught	Resilient	Reserved	Extra Cluster
Control	1, 5, 11	2, 3, 4, 9, 10, 14	7, 8, 12, 13	6
Depression History	1, 5	2, 3, 4, 7, 9, 10, 14	6, 8, 12, 13	11
Currently Depressed	1, 5, 11	2, 3, 4, 7, 9, 10, 14	8, 13	6, 12

*Reserved* patients were characterized by high scores on *concerned*, *likeable*, *softhearted*, *worried*, *compassionate*, and *warm*. Because this analysis included fewer

groups than the previous study (which had five clusters; Preston et al., under review), it was not possible to reproduce the results exactly. However, many similarities emerged, with the greatest distinctions again between upset, *distraught* patients, more positive *resilient* patients, and less expressive *reserved* patients.

To further examine how participants perceived the patients' emotions, k-means analysis was run a second time to examine how the clusters change if a four-cluster solution was forced. Based on the agglomerative plots (Figure 4.2; right), a four-cluster solution was also a reasonable fit for the data. In control participants, the *distraught* and modified *resilient* clusters stayed the same. The original *reticent* patient broke apart from the *reserved* patients into a separate cluster, as in the 5-cluster solution found previously (Preston et al., under review). This patient was rated similarly, although slightly more negative, than the *reserved* patients: *troubled*, *sad*, *softhearted*, *concerned*, *upset*, and *warm*.

### ***Within-Group Cluster Analysis – Depression History***

The agglomerative coefficient again suggested a three-cluster solution was the best fit for the depression history group (Figure 4.2; left). Although the change in the agglomerative coefficient is much larger in controls than in the depression history group (indicating a larger change in the sum of squares when the data is divided into additional groups), the same pattern is present, which suggests the same number of groups are an ideal fit for this data. K-means clustering produced a *distraught* cluster that was identical to the one found in controls and in previous studies (Preston et al., under review), and was characterized as *sad*, *upset*, *troubled*, *grieved*, *distressed*, and *sorrowful*. The other

two clusters were only slightly different, with one patient switched between the clusters found in controls. The *resilient* cluster also included a patient that was in the controls' *reserved* cluster, and who was originally *wistful*. This group was again rated as highly *likeable, engaging, warm, happy, softhearted, and compassionate*. The final cluster of patients included the rest of the *reserved* patients. These patients again seemed to be perceived in between *distraught* and *resilient* patients, and were rated as *likeable, concerned, troubled, softhearted, worried, and engaging*. Other than the one patient in the controls' *reserved* and the history groups' *resilient* cluster, the clustering solutions were identical between controls and participants with a depression history. Clusters were formed with that patient in the *reserved* group, as the initial hierarchical solution for depression history participants suggested the patient initially fit with that cluster (although the k-means clustering indicated a better fit in *resilient* group).

Again, k-means clustering was used to determine the qualities of a four-cluster solution. The agglomerative coefficient indicated this was also a reasonable fit, and importantly, the four-cluster solution appeared to fit both the data in both groups equally well, as the patterns of decrease in the agglomerative coefficient were similar. For participants with a depression history, the most distinct four clusters did not involve the original *reticent* patient splitting off from the larger group of *reserved* patients, as it did in the controls. Rather, one of the *distraught* patients became a unique cluster. This patient was rated differently than the other two *distraught* patients, although just as negative, being perceived as *sad, grieved, upset, sorrowful, troubled, and distressed*. Because the k-means algorithm works to create clusters that are most different from one another, the fact the *distraught* clusters divided suggests participants with a depression history rated

the most negative group of patients more distinctly than the other, more positive patients. It also replicates the pilot data (Appendix 4.A), despite the differences in the measures (all 29 other rated adjectives versus ratings of felt negative emotion, positive emotion, empathic concern and personal distress in the pilot data) and clustering methods (k-means versus hierarchical), indicating consistency in the finding that depression experience is related to heightened differentiation between the negative displays of others.

***Are the differences in depression history clusters driven by current depression?***

To assess whether the increased differentiation between perceived negative affect found in depression history participants would also be found in participants currently experiencing depression symptoms, the k-clustering analysis was repeated in participants with high BDI scores (using a median split, scores > 10; this divides people with minimal depression from those with moderate, mild, and severe depression using standard cut-offs; see Figure 4.1 above). The three-cluster solution was identical to that found for participants with a depression history, with both original *sanguine* and the same *wistful* patient included with the *resilient* cluster.

However, when a four-cluster solution was examined, the results did not resemble the solution found in the depression history participants. Instead, the group of *distraught* patients remained intact, with the larger *reserved* cluster breaking up as it did in the controls. Instead of the *reticent* patient breaking off, two of the original *wistful* patients broke away to form the extra cluster. These patients were rated as *likeable*, *engaging*, *softhearted*, *concerned*, *troubled*, and *worried*. This suggests the increased differentiation between negative affect seen in depression history participants was specific to past

experience of diagnosed depression, rather than the extent of participants' current depressive state (which could be below the threshold or duration necessary to classify as an episode of clinical depression). These findings provide additional evidence that increased differentiation for negative states perceived in others is likely due to representational changes as a result of past experience, rather than state-related components of current depression.

### ***Card Sorting***

On average, participants split the patients into four groups, which is slightly higher than the number of groups found by the three-cluster hierarchical solution. The number of groups created by participants with and without a depression history did not differ statistically,  $t(57.12) = -1.24, p = .22$ , but on average, controls made fewer categories ( $M = 3.84$ ) than participants with a depression history ( $M = 4.17$ ). Similarly, while the numbers of categories that involved descriptions of sadness did not differ between the two groups statistically,  $t(45.12) = -1.07, p = .29$ , controls again had numerically fewer sad groups ( $M = 1.06$ ) than participants with a depression history ( $M = 1.28$ ). However, mirroring the clustering solution, participants with a depression history divided the *distraught* patients into more groups on average ( $M = 1.89$ ) than the controls ( $M = 1.56$ ),  $t(57.95) = -2.07, p = .04$ .

Because participants tended to make more than 3 groups, very few exactly produced the ones found through clustering. Out of the 63 participants, 12 participants had the exact *distraught* group, one participant produced the exact *resilient* group, and no one had the entire *reserved* group together. However, many seemed to group patients into

those emotional “types”, with 36 participants creating a group described like the *distraught* patients (e.g., “seemed extremely sad and depressed” or “very emotionally distressed, very sad”), 12 had a *resilient*-like group (e.g., “worried about their illness, however, were very upbeat and happy”), and 12 had a *reserved* group (e.g., “they seemed to have barely any emotion” or “seems more reserved and private than any of the others”).

### ***Characterizing the patient types***

Responses to the three patient target types found in the initial clustering solutions – *distraught*, *resilient*, and *reserved* – were compared across participants. Principle components analysis (PCA) of the 29 adjectives rated for the patients’ emotions revealed three “other” factors that accounted for 56.9% of the variance (named for the strongest loading factor, listed with the top five loading adjectives  $> .5$ ): *Panicked* (panicked, horrified, perturbed, afraid, distressed; 37.1% of variance), *Sad* (sad, grieved, troubled, upset, sorrowful; 15.1% of variance), and *Softhearted* (softhearted, tender, compassionate, warm, likable; 4.7% of variance). These factors were the same when ratings from depression history and control groups were analyzed separately.

Confirming components of the grouping from Preston et al. (under review), *distraught* patients appeared more panicked, sad, sick and in more need of emotional and practical support than the other patient types,  $F_s(2, 124) > 17.17, p < .001$  (see Table 4.2).

*Reserved* patients were perceived as sadder than *resilient* patients,  $p < .001$ , and were seen as less softhearted than either *distraught* or *resilient* patients,  $F(2, 114) = 9.07, p < .001$ . Participants with and without depression histories did not differ in their overall

ratings,  $F(1, 57) < 1.16, p > .29$ , nor in their perceptions of the different clusters,  $F(2, 114) < 0.76, p > .48$ . However, current depression scores tended to interact with perceptions of how softhearted the patients were and their perceived need of practical support,  $F_s(2, 114) > 3.01, p_s < .05$ . High depression scores were associated with seeing the *resilient* patients as more softhearted,  $\beta = .013, p = .02$ , and the *distraught* patients as less softhearted,  $\beta = -0.16, p = .10$ , and believing *distraught* patients needed less practical support,  $\beta = -0.034, t = -2.29, p = .026$ . Depression scores were not associated with perceiving the patients as any less sick, sad, or panicked,  $F_s(2, 114) < 1.40, p_s > .25$ .

### ***Emotional responses to the patient types***

PCA analysis of the adjectives rated for the participants' emotional responses revealed four "self" factors that accounted for 54.2% of the variance. These factors were similar to those found in previous studies with these adjectives (Preston et al., under review), and included a factor similar to Batson's *Personal Distress* (Batson et al., 1997; troubled, distressed, worried, upset, concerned; 29.78% of variance), a factor similar to Batson's *Empathic Concern* (softhearted, compassionate, tender, sympathetic, warm; 14.5% of variance), and two new factors that were not found by Batson but were found in the prior study: *Amused* (amused, funny, happy, pleased, likeable; 6.2% of variance), and *Horried* (horrified, perturbed, panicked; 3.6% of variance).



**Table 4.2. Means (SD) for variables by depression history group and cluster and beta values for the relationship with BDI (bold indicates significant cluster x current depression interactions)**

		Controls			Depression History			Current Depression ( $\beta$ )		
		1	2	3	1	2	3	1	2	3
		Distraught	Resilient	Reserved	Distraught	Resilient	Reserved	Distraught	Resilient	Reserved
Perceptions of Patients	F1 Panicked	0.95 (0.78)	-0.25 (0.29)	-0.21 (0.38)	1.08 (0.6)	-0.33 (0.26)	-0.23 (0.34)	-0.009	0.001	0.006
	F2 Sad	1.01 (0.69)	-0.76 (0.33)	0.21 (0.28)	0.93 (0.37)	-0.76 (0.29)	0.28 (0.22)	-0.001	0.004	-0.003
	F3 Softhearted	0.36 (0.79)	0.13 (0.41)	-0.31 (0.32)	0.06 (0.49)	0.25 (0.34)	-0.29 (0.39)	<b>-0.016</b>	<b>0.013</b>	<b>0</b>
	Sick	4.87 (1.16)	3.50 (1.03)	3.74 (1.12)	4.7 (1.25)	3.44 (1.08)	3.79 (1.00)	-0.015	0.005	0.001
	Practical Need	5.32 (1.14)	3.96 (0.99)	4.03 (1.26)	5.18 (1.04)	4.01 (1.20)	4.00 (1.17)	<b>-0.034</b>	<b>-0.011</b>	<b>0.004</b>
	Emotional Need	6.25 (0.74)	3.54 (1.33)	3.99 (1.21)	6.18 (1.08)	3.71 (1.38)	4.23 (1.09)	-0.037	-0.002	-0.011
Emotional Responses	F1 EC	3.66 (1.40)	3.33 (1.26)	3.07 (1.08)	3.41 (1.15)	3.23 (1.11)	2.98 (0.91)	<b>-0.012</b>	<b>0.022</b>	<b>0</b>
	F2 PD	3.49 (1.27)	1.92 (0.83)	2.29 (0.99)	3.33 (1.24)	1.81 (0.82)	2.15 (0.93)	0.009	0.02	0.005
	F3 Amused	1.78 (0.70)	2.66 (1.12)	1.96 (0.81)	1.56 (0.65)	2.49 (1.09)	1.84 (0.79)	0.002	0	0
	F4 Horrified	1.57 (0.89)	1.13 (0.36)	1.18 (0.51)	1.37 (0.49)	1.04 (0.16)	1.09 (0.28)	-0.004	0	0.002
	F5 Angry	1.34 (0.65)	1.07 (0.23)	1.11 (0.33)	1.45 (0.76)	1.04 (0.12)	1.10 (0.22)	0.004	0	0.001
	Life Similarity	1.26 (0.58)	1.30 (0.49)	1.22 (0.40)	1.46 (0.71)	1.61 (0.64)	1.60 (0.74)	0.006	-0.004	0
	Emotion Similarity Like	1.78 (0.93)	1.91 (0.99)	1.61 (0.63)	2.2 (1.16)	2.45 (1.02)	2.18 (0.95)	0.028	0.018	0.02
Physiological Responses	GSR (mean) (Z)	0.11 (0.40)	-0.07 (0.37)	0.01 (0.36)	-0.12 (0.63)	-0.05 (0.31)	0.11 (0.42)	-0.001	-0.001	0.003
	GSR (PP) (Z)	0.12 (0.56)	0.02 (0.27)	-0.11 (0.36)	0.39 (0.50)	-0.08 (0.30)	-0.13 (0.41)	-0.002	0.004	-0.003
	Respiration (Z)	0.09 (0.50)	0.09 (0.28)	-0.16 (0.36)	-0.20 (0.56)	-0.03 (0.38)	0.13 (0.36)	0.007	-0.003	-0.002
	Heart Rate (Z)	-0.21 (0.54)	0.11 (0.36)	0.00 (0.37)	-0.20 (0.53)	0.13 (0.35)	-0.03 (0.37)	0.006	-0.004	0
Prosocial Responses	Practical Give	5.95 (1.06)	4.46 (1.44)	4.52 (1.29)	5.57 (1.58)	4.57 (1.70)	4.54 (1.42)	<b>-0.053</b>	<b>-0.012</b>	<b>-0.014</b>
	Emotional Give	5.28 (1.34)	4.47 (1.42)	4.28 (1.30)	5.11 (1.73)	4.52 (1.70)	4.36 (1.72)	<b>-0.045</b>	<b>0.01</b>	<b>-0.005</b>
	Tokens	3.94 (1.29)	3.23 (1.47)	3.05 (1.62)	3.56 (1.28)	3.21 (1.43)	3.12 (1.38)	-0.029	-0.006	-0.026

While control participants had these exact factor scores, PCA of the ratings from the depression history participants alone were quite different. Four factors again best fit their self ratings, but the first factor was a combination of *Personal Distress and Empathic Concern* (sympathetic, grieved, sad, upset, sorrowful; 31.4% of variance), then *Likeable* (likeable, warm, engaging, happy, pleased; 14.4% of variance), *Panicked/Horrified* (panicked, horrified, perturbed, alarmed; 6.2% of variance), and *Angry*, which just consisted of the adjective angry (3.96% of variance). Consistent with the clustering results and pilot data findings, responses in participants with a depression history seemed to be best captured by negative, rather than positive, components. In order to best capture the emotional responses across participants, five scores were calculated from the adjectives that loaded on each factor ( $> .5$ ), rather than using the factor scores themselves. This resulted in “self” scores for personal distress, empathic concern, amused/likable, horrified, and anger.

*Distraught* patients elicited more feelings of personal distress, empathic concern, horror, and anger than the other patient types,  $F_s(2, 124) > 17.17, p < .001$ , replicating prior studies (Preston et al., under review). Resilient patients were liked more and evoked more amusement than *distraught* or *reserved* patients,  $F(1, 124) = 21.04, p < .001$ . *Resilient* patients also elicited more empathic concern and less personal distress than *reserved* patients (post-hoc  $p < .001$ ). Participants felt more emotionally similar to *distraught* and *resilient* patients compared to *reserved* patients,  $F(2, 124) = 4.74, p = .01$ , but reported similar life experiences for all patient types,  $F(2, 124) = 1.27, p = .28$ . Participants with a history of depression tended to report more life similarity overall across patient types,  $F(1, 57) = 3.21, p = .079$ , but responded no differently than controls

to the patient types, cluster by history interactions:  $F_s(2, 114) < 2.27, p_s < .11$ . Current experience of depression affected feelings of empathic concern, cluster by BDI interaction:  $F_s(2, 114) = 5.34, p = .006$ . Parameter estimates revealed that higher depression scores were associated with less empathic concern for *distraught* patients ( $\beta = -0.012$ ) and more empathy for *resilient* patients ( $\beta = 0.022$ ), and were unrelated to empathic concern for *reserved* patients ( $\beta = .00$ ), although none of the predictors were significant on their own,  $t_s < 1.23, p > .22$ . Current experience of depression was not related to differences in personal distress, amusement, horror, or anger responses to the different patient types,  $F(2, 114) < 0.87, p > .42$ .

*Distraught* patients surprising evoked the largest decrease in heart rate compared to the other clusters,  $F(2, 110) = 4.47, p = .035$ . Although they elicited the most personal distress, anger and horror (usually associated with *increased* heart rate), they also elicited the most empathic concern, which is an other-oriented state associated with decreased heart rate (Eisenberg & Fabes, 1990). The other psychophysiological measures did not differ by cluster,  $F_s(2, 110) < 0.997, p_s > .37$ . Participants with and without a history of depression differed in their respiration rates,  $F(2, 110) = 6.43, p = .002$ . Controls had faster respiration to *reserved* patients ( $M = -0.18$ ) than *distraught* ( $M = 0.14$ ), post-hoc  $p = .03$ , while participants with a depression history had faster respiration to *distraught* ( $M = -0.26$ ) than to *reserved* ( $M = .14$ ), post-hoc  $p = .007$ . This suggests participants with a depression history may have found the *distraught* patients more arousing, despite the overall heart rate decrease in response to these patients across participants. Depression history and BDI did not interact with any other measure,  $F_s(2, 110) < 1.5, p_s > .23$ .

### ***Prosocial responses***

Participants gave more practical support, emotional support, and more token donations to *distraught* compared to the other patient types,  $F_s(2, 124) > 23.11, p < .001$ . Giving did not differ by depression history,  $F_s(2, 114) < 1.76, ps > .18$ , but current experience of depression affected amounts of practical and emotional support given to the different clusters, cluster by BDI score interactions:  $F_s(2, 114) > 4.54, p > .013$ . Parameter estimates indicated participants with higher depression scores gave less practical support,  $\beta = -0.05, t = -2.5, p = .015$ , and less emotional support,  $\beta = -0.05, t = -2.42, p = .019$ , to *distraught* patients. BDI was not related to giving for any other patient type,  $ts < 0.71, ps > .48$ .

Other variables were examined to determine what might be driving the surprising finding that currently depressed participants gave less to *distraught* patients. Because perceived need of practical support showed the same pattern of results as the giving measures (with BDI negatively predictive only for *distraught* targets), this variable was explored as a potential candidate for explaining the relationship between giving and BDI. Step-wise multivariate regression was used to determine whether perceived need would account for the negative relationship between current depression and giving to *distraught* patients. While this analysis does not provide information about actual causality, if perceived need could statistically account for the relationship between BDI and giving measures, it would be a good candidate for explaining this relationship.

When entered alone, BDI was negatively associated with emotional and practical support given to *distraught* patients,  $\beta_s < -0.043, ts < -2.25, p < .028$ . When perceived need of practical support was added to the model, the  $\beta$  coefficients became less negative

and the effects of BDI on practical support given became non-significant,  $\beta = -0.015$ ,  $t = -1.03$ ,  $p = .31$  and were a marginal trend for emotional support given,  $\beta = -0.028$ ,  $t = -1.73$ ,  $p = .09$ . These findings provide support for perceptions of need accounting for the decreased giving in participants who are currently depressed. Although perceptions of how sick the patient was, how distressed they seemed, and how much distress they evoked were highly correlated with perceptions of needed practical support,  $rs(64) > .29$ ,  $ps < .02$ , none of these variables accounted for the relationship between BDI and need,  $\beta s < -0.021$ ,  $ts < -1.78$ ,  $ps < .08$ , suggesting it was unlikely these factors could explain the lower ratings of need among depressed participants.

### ***Life and emotional similarity***

Another surprising result was that giving did not differ among participants with and without a history of depression. However, the extent to which participants with a depression history perceived their experience as similar to the patients' is an important factor in whether their shared experience translates to increased understanding and giving. Supporting this, ratings of similar life experiences were associated with giving more practical support, emotional support, and tokens,  $F_s(1, 152.18) > 3.36$ ,  $ps < .069$ , but only for participants with a past experience of depression, interaction:  $F_s(1, 152.18) > 5.58$ ,  $ps < .02$ ; depression history:  $\beta's > 0.44$ ,  $ts > 1.76$ ,  $p < .08$ ; controls:  $\beta's > -0.21$ ,  $ts < 0.79$ ,  $ps > .43$ . In controls, similar life experiences even tended to be associated with decreased token donations,  $\beta = -0.56$ ,  $t = -1.74$ ,  $p = .08$ . Similar emotional experiences also predicted increased giving of practical and emotional support,  $F_s(1, 129.33) > 7.57$ ,  $ps < .007$ , but again only among participants with a depression history, interaction:  $F_s(1,$

135.16) > 6.09,  $ps < .015$ ; depression history:  $\beta$ 's > 0.56,  $ts > 3.49$ ,  $ps < .001$ ; controls:  $\beta$ 's < 0.16,  $ts < 0.96$ ,  $ps > .34$ .

## Discussion

Major depression disorder has a lifetime prevalence of about 17% (Kessler et al., 2005), and nearly every person in the population is likely to be affected by their own or a close other's depression (Henderson & Rickwood, 2000). However, little is known about how one's past and current experience of depression affects perceptions of and responses to the emotions and needs of others. In the current study, participants with a past experience with depression perceived greater distinctions between the upset, *distraught* targets than controls, indicating that depression experience was associated with richer, more differentiated representations of others' distress. Moreover, giving in participants with a past depression experience was strongly tied to the extent to which they perceived their prior emotional and life experience as similar to the patients'. This supports predictions from the perception-action model (PAM; Preston & de Waal, 2002), as the PAM asserts that past experience provides observers richer representations of related states and situations, and these representations are recruited and enable empathic responding when attention is paid to targets whose specific emotional and life situations are similar.

Despite finding expected results for those with a past history of depression, participants who were currently depressed actually gave less to patients who displayed intense negative affect (*distraught*), even though these patients were hypothesized to be

the most emotionally similar to them and to evoke the most empathy from them. Possible explanations for these findings are discussed below.

Replicating previous studies (Preston et al., under review) and the pilot study, consistent display types again emerged from this sample of 14 patients. The *distraught* patients were replicated exactly, and a similar group of *resilient* patients were found. The other types of patients, *wistful*, *sanguine*, and *reticent* were either grouped with the *resilient* patients or clustered together in bigger *reserved* group of patients. A noticeable difference between this and prior studies is that fewer types of patients were found, which may have been due to slight differences in clustering methods (i.e., using all the other rated adjectives versus other factor scores, or using k-means versus hierarchical methods alone), or because the current participants in this study did not make equally large distinctions between patients whose emotional displays were somewhere between the extreme distress of *distraught* and the more positive emotion of *resilient* patients, resulting in these patients ending up in a single cluster (*reserved*) rather than several clusters (*wistful*, *sanguine*, *reticent*). The card-sorting results also confirmed this, as many participants created groups similar to these patient types, even though they did not recreate the groups exactly. Perceptions of and responses to these patient types were also consistent with other studies (Chapters 2 and 3; Preston et al., under review). *Distraught* patients were perceived as most sick, sad, panicked, and in need of support than the others, and elicited the most distress, empathy, and help in participants. *Resilient* patients were again seen as most happy and softhearted, were liked the most, and elicited the most amusement from participants. *Reserved* patients elicited the least empathic concern of

any patient type similar to the prior *reticent* patient who was included in this group in the current study.

The effects of depression on emotional perception and empathy differed depending on whether the observers' depression was a past or current experience. These two things were not independent across participants, as BDI scores were higher in depression history participants than controls. However each was associated with distinct effects. When history of depression was examined alone, participants with depression experience rated the *distraught* patients more distinctly than controls. Instead of the original *reticent* patient breaking off from the larger *reserved* group as in controls, the most distinct four groups of patients for participants with depression history included two "types" of *distraught* patients. This replicates the pilot data, and suggests the past experience of diagnosable depression is associated with perceiving greater differentiation within negative affect expressed by others compared to those without such experience. This is consistent with the idea that such experience provides a richer representation of negative affect in general, and that these representations influence perceptions of similar emotion in another.

Notably, this increased differentiation between negative affect perceived in targets was not seen when participants were divided by current experience of depression symptoms, regardless of history (the median split for high BDI contained 18 participants with depression history and 10 controls, and included those with depression scores above what would be clinically classified as minimal;  $> 10$ ), suggesting it was specific to having a history of depression severe enough to be diagnosed. Similar findings were seen in the self-rated adjectives, as the emotional responses of participants with depression history



were better captured by negative factor components than the responses of controls. While the consistency of these findings across multiple studies and populations attests to the robustness of these effects, it is possible that the distinctions made among the *distraught* patient types were specific to the context of the patients' situations. For example, the *distraught* patient that separated from the other two in the depression history group was distressed over the loss of her husband, while the other two were distressed about their illness. Perhaps people with a depression history made finer distinctions across sources of distress, rather than the type of distress displayed. Future work should attempt to tease apart these possibilities, which are inherently combined in these naturalistic stimuli. For example, testing whether similar differentiation effects would be found if targets were manipulated to vary in the source or displays of their distressed affect.

Surprisingly, participants with a depression history did not give more help to the *distraught* targets, as initially predicted, but this is because they did not actually feel that their experience was similar to the *distraught* patients as we had assumed. They seemed to be the most aroused by *distraught* targets, rather than orienting to them, as their respiration rate was much higher for these targets compared to controls. This could perhaps be a result of increased feelings of personal distress, or arousal more akin to shared, empathic distress. However, self-report ratings of felt distress did not corroborate the psychophysiological findings. Across all targets, the more participants with a depression history felt they shared similar life or emotional experiences with the patients, the more help they gave. Although not in line with the initial predictions, these findings support the assertion of the PAM that ability to resonate with another, and thus help, critically depends on how similar the specific experience of the other is, as individual

appraisals and contexts can be very different between two people who experience the same event or emotional state. Indeed, the participants' perceived similarity between their own and the patients' experiences was a better predictor of helping than the patients' specific emotional displays. This is true even within participants who have all experienced depression – an emotional state that appears to manifest across people in quite similar ways, at least in terms of felt negative affect, and especially compared to the more heterogeneous experiences often used in empathy paradigms, such as divorce or childbirth.

In contrast to the past experience of depression and also counter to our predictions, the extent to which participants were currently depressed was associated with feeling less empathic concern, and giving less emotional and practical support to *distraught* patients. These effects were specific to *distraught* patients and did not reflect an overall decrease in empathy or help, as current depression was also associated with more empathic concern for *resilient* patients. Such results were surprising, as similar current experience was predicted to enhance, rather than decrease, empathy for targets who expressed similar emotion. It seemed that this decreased giving was not due to currently depressed participants feeling more personally distressed in response to *distraught* targets, liking them less, or perceiving them as less distressed. Intrusive cognition scores for emotional information were also not related to the extent of current depression symptoms in this sample, ruling out differences associated with currently depressed participants being more prone to attentional capture or distraction by negative emotional information as a possible explanation. This decreased giving also was not due to currently depressed participants feeling less emotionally similar to *distraught* targets,

as depression scores were related to feeling more similar to *distraught* patients (post hoc tests confirm BDI was positively correlated with emotional similarity,  $r(60) = .30, p = .019$ ).

Instead, what seemed to best explain this decreased giving was that depressed participants thought *distraught* patients needed less practical support. Perceptions of need partially mediated<sup>6</sup> the relationship between depression score and giving. However, as with any associative scores, temporal causality can not be assumed, as the opposite could also be true; perhaps participants rated *distraught* targets as needing less help as justification for giving less help, rather than because they genuinely perceived them to be less in need.

There was not a clear answer in the data as to why currently depressed participants would perceive the *distraught* targets as less in need of practical support. Related variables, such as perceived sickness, distress of the patients, and felt distress in response to the patients did not account for this relationship, and therefore are unlikely explanations. According to the PAM, such anti-empathic responses may occur when the observer has a conflicting goal state, which can down-regulate more automatic resonance through cognitive mechanisms (Preston & de Waal, 2002). Perhaps the negative self-focus associated with the current experience of depression induces a competing goal state – attention and focus on their own feelings and their own needs – in observers, preventing them from resonating with the *distraught* targets. Depressed participants may also have found the *distraught* patients annoying, perhaps viewing the patients' overt displays of

---

<sup>6</sup> Mediation is used here not in the strict sense (Baron & Kenny, 1986), but in the sense that perceived need reduced the beta value associated with current depression to non-significance when it was added to the model. Because multi-collinearity issues can arise in mediation analysis, the reduction of significance is not usually enough to qualify for mediation.

distress as an over-reaction to their need. Supporting this, the most depressed participants also tended to like *distraught* patients less. However, our pilot data found that currently depressed women felt more empathy for *distraught* targets. These differences may result from measuring depression symptoms here as an individual difference, varying in participants who both had and had not received a depression diagnosis, rather than using the score as an additional assessment of depression only in participants who were already clinically interviewed and diagnosed with depression. Some controls may have been pre-diagnosis or had avoided being diagnosed, but it is also possible current depression scores reflected sub-clinical depression tendencies in some of the college students in this sample.

A second possible explanation for the decreased ratings of need is that participants with many depression symptoms were simply more accurate at inferring the actual need of the patients compared to non-depressed participants. While it is impossible to empirically test this hypothesis in the current study, both depression (e.g., Alloy & Abramson, 1979) and negative mood (e.g., Sinclair & Mark, 1995) have been associated with increased accuracy in judgments that have objective criteria. Although depression is often associated with impaired recognition of others' facial expressions (Cooley & Nowiki, 1989; Surguladze et al., 2004), current depression experience in this study was not associated with differences in the perception of distressed emotion in targets. Future work may benefit from examining whether current depression experience is associated with accurate judgment about the need of others in situations where the extent of their need or specific type of need is known.

Together, the results of this study suggest past and current depression experience can affect responses to others in opposing ways. Having a past experience of depression was associated with more differentiated perceptions of the negative emotions displayed by targets and increased prosocial responses when the situations and emotions of patients were perceived as similar. Contrastingly, current depression symptoms were not associated with the same distinctions within negative emotion. Surprisingly, depressed participants gave less to *distraught* targets, despite feeling similar to them, likely because they perceived these targets as less in need. Such differences have important implications for social interactions, as the experience of depression is increasingly common in the population.

One question still outstanding is whether the decreased empathy for *distraught* targets among currently depressed participants could be explained simply by their current negative state, or whether these responses were specific to other aspects of current depression. To test this, a follow-up study was run. Participants were first induced to feel sadness or neutral emotion, and then watched and responded to one *resilient* and one *distraught* patient. If these effects could be explained by the negative mood of the currently depressed participants, the induction of sadness should result in similar reductions in empathy for the *distraught* target.

## Follow-up Study

### Methods

#### Participants

Fifty adults (28 men; age  $M = 33.6$ , range 18-65) were recruited through Mturk to participate in this study. Participants were paid \$0.50 to complete the 15-minute online task. One participant was removed due to failure to correctly answer the check questions, and two participants were removed due to outlying scores in the opposite direction of their induction group, leaving 47 participants (26 men) in the final analysis. All participants consented to participate and all procedures were approved by the Institutional Review Board of the University of Michigan.

#### Procedure

Participants first filled out the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988) as a baseline measure of their current mood state. Then, participants read a story designed to either induce sad emotion or neutral emotion (counterbalanced across participants; stories taken from Raghunathan & Pham, 1999), and filled out the PANAS a second time in order to assess change in emotion. Participants then watched two videos of hospital patients, one *distraught* and one *resilient* (videos 11 and 9, respectively; Chapter 4). The patients were selected as the highest rated *distraught* patient on “distressed” and “distraught” and the highest rated *resilient* patient on “happy” and “resilient” from the study in Chapter 3. After each video, participants rated their responses using the same adjectives and variables as previously

(Chapter 4). Instead of donating a portion of 5 tokens to each of the patients, participants divided 11 tokens between the two patients after watching and responding to both videos. Finally, they filled out demographic information.

### **Statistical Analysis**

PANAS scores were combined into the negative and positive emotion factors (as in Watson et al., 1988; positive: *interested, excited, strong, enthusiastic, proud, alert, inspired, determined, attentive, active*; negative: *upset, distressed, guilty, scared, hostile, irritable, ashamed, nervous, jittery, afraid*). Scores on the initial PANAS were subtracted from scores on the post-induction PANAS to produce change scores for positive and negative emotions. As a manipulation check, scores were compared between the sad and neutral group using independent t-tests. PANAS scores were also compared between male and female participants to determine whether the manipulation differed in effectiveness across gender.

The 29 adjectives rated for observed and felt emotion were again factor analyzed using PCA to reduce and identify “other” and “self” component scores. The “other” scores along with ratings of practical support needed, emotional support needed, and sick were used to confirm patients were rated consistent with previous studies. The “self” scores along with ratings of liking were used to assess emotional responses to the patients, and practical support given, emotional support given and tokens allocated to each patient were compared to assess prosocial behaviors. All variables were analyzed using mixed ANOVAs with induction group as a between subjects factor and change in negative emotion (mirroring the continuous BDI score) included as a covariate to

determine whether responses differed between groups and as a function of how negative participants were feeling.

## **Results**

### **Manipulation check**

Confirming the manipulation check was successful, participants induced to feel sadness felt more negative ( $M = 0.53$ ) and less positive ( $M = -0.85$ ) than the participants in the neutral group (negative:  $M = 0.06$ ; positive:  $M = -0.30$ ),  $ts(34.09) > 2.48$ ,  $ps < .018$ . Male and female participants did not differ in their overall emotional responses  $ts(46) < 0.72$ ,  $ps > .47$ , nor in their responses to the specific emotion inductions,  $ts(23) < 1.63$ ,  $ps > .12$ . Because the manipulation seemed to affect male and female participants similarly, gender was not included as a covariate in the analysis.

### **Perceptions of Patients**

PCA analysis of the “other” factor scores reveal four factors that explained 73.7% of the variance (named for the most similar component from the original study or strongest loading adjective and listed with the top 5): *Sad* (39.10% of variance; grieved, distressed, sorrowful, sad, afraid), *Panicked* (20.41% of variance; horrified, perturbed, panicked, angry, disconcerted), *Softhearted* (10.15% of variance; warm, engaging, softhearted, tender, likable), and *Funny* (4.08% of variance; funny, pleased, amused, happy; only four adjectives with  $> .5$  loading). The first three components were similar to those found in Chapter 4, with the addition of Funny.



Consistent with previous work, the *distraught* patient was rated as more sad and in need of more emotional support than the *resilient* patient,  $F_s(1, 46) > 26.60, p < .001$ , while the *resilient* patient was rated as more softhearted and funnier,  $F_s(1, 46) > 5.47, p < .02$ . Perceptions of sadness tended to be affected by the emotion induction,  $F(1, 46) = 3.70, p = .06$ . While both groups rate the *resilient* patient similarly, sad participants perceived the *distraught* patient to be less sad than the neutral participants. Perceptions of the patients were not associated with the extent to which participants felt negative,  $F_s(1, 46) < 2.24, p > .14$ .

### **Emotional responses to the patients**

PCA analysis of the adjectives rated for felt emotion revealed four “self” factors that explained 75.56% of the variance: *Horried* (43.89% of variance; perturbed, horrified, panicked, angry, disconcerted); *Empathic Concern* (14.99% of variance; softhearted, warm, moved, compassionate, tender); *Personal Distress* (13.07% of variance; grieved, sorrowful, alarmed, sad, distressed); and *Amused* (3.62% of variance; funny, pleased, happy, amused). Despite the adjectives loading slightly differently on each, these factors were identical to the ones found in the initial study.

The *distraught* patient elicited more personal distress,  $F(1, 45) = 15.52, p < .001$ , but the *resilient* patient elicited more empathic concern, amusement, and was liked more,  $F_s(1, 45) > 4.40, p_s < .04$ . Empathic concern for the two patients was affected by the emotion induction,  $F(1, 45) = 6.03, p = .018$ . Participants felt similar for the *distraught* patient, but neutral participants felt even more empathic concern for the *resilient* patient compared to the sad participants. Sad participants felt more amused after both patients

compared to neutral participants,  $F(1, 45) = 4.63, p = .037$ , and tended to feel less empathic concern,  $F(1, 45) = 3.57, p = .065$ . Negative emotion again did not affect ratings of the patients,  $F_s(1, 46) < 2.08, p_s > .16$ .

### **Prosocial Responses**

Unlike the prior study, practical support, emotional support, and token donations did not differ between the two patients,  $F_s(1, 45) < 0.83, p_s > .37$ . Both groups helped the patients similarly,  $F_s(1, 45) < 0.75, p_s > .39$  and negative emotion did not affect helping,  $F_s(1, 45) < 0.70, p_s > .41$ . Because there were specific predictions that participants induced to feel sadness would give less to distraught patients compared to control participants, planned comparisons were performed to test this. Unfortunately, this was not found,  $t_s(46) < 1.27, p > .21$ , although numerically, the means for sad participants, particularly for emotional support to *distraught* targets ( $M = 5.04$ ), were lower than for neutral participants ( $M = 5.61$ ). However, this pattern was not exclusive to the *distraught* patient, as sad participants tended to give less emotional support to the *resilient* patient as well (sad  $M = 4.88$ ; neutral  $M = 5.57$ ),  $t(46) = -1.79, p = .079$ .

### **Conclusions**

The perceptions of and responses to the two patients in this study were similar to the ratings for their overall clusters, suggesting they were representative of the entire clusters used in Chapter 4. While the manipulation check demonstrated the sad story reliably elicited negative emotion in participants, sad participants did not show the same decreases in giving to *distraught* targets as depressed subjects did in the previous study.

Sad participants seemed to feel less empathy overall, and if anything, gave slightly less to the *resilient* patient. Differences in giving and empathy were also not predicated by the amount of negative affect participants felt in response to the story. The results of this study suggest that the decreased giving to *distraught* targets associated with current depression are not simply due to the current experience of negative emotion, and instead may be due to other aspects of depression.

## References

- Ault, K., Bishop, E., Carpenter, L., Chen, P., Chromy, J. R., Copello, E., Cunningham, D. B., et al. (2007). Results from the 2007 National Survey on Drug Use and Health: National Findings. Network. Substance Abuse and Mental Health Services Administration.
- Bargh, J. A., & Tota, M. E. (1988). Context-dependent automatic processing in depression: accessibility of negative constructs with regard to self but not others. *Journal of Personality and Social Psychology*, 54(6), 925-39.
- Barnett, M. A. (1984). Similarity of experience and empathy in preschoolers. *Journal of Genetic Psychology*, 145(2), 241-250.
- Barnett, M. A., Tetreault, P. A., Esper, J. A., & Bristow, A. R. (1986). Similarity and empathy: The experience of rape. *Journal of Social Psychology*, 126, 47-50.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173-82.
- Batson, C. D., Eklund, J. H., Chermok, V. L., Hoyt, J. L., & Ortiz, B. G. (2007). An additional antecedent of empathic concern: valuing the welfare of the person in need. *Journal of Personality and Social Psychology*, 93(1), 65-74.  
doi:10.1037/0022-3514.93.1.65
- Batson, C. D., Fultz, J., & Schoenrade, P. A. (1987). Distress and empathy: two qualitatively distinct vicarious emotions with different motivational consequences. *Journal of Personality*, 55(1), 19-39.

- Batson, C. D., Quin, K. O., Fultz, J., Vanderplas, M., & Isen, A. M. (1983). Influence of self-reported distress and empathy on egoistic versus altruistic motivation to help. *Journal of Personality and Social Psychology*, 45(3), 706-718.
- Batson, C. D., Sager, K., Garst, E., Kang, M., Rubchinsky, K., & Dawson, K. (1997). Is empathy-induced helping due to self-other merging? *Journal of Personality and Social Psychology*, 73(3), 495-509.
- Batson, C. D., Sympson, S. C., Hindman, J. L., Decruz, P., Todd, R. M., Weeks, J. L., Jennings, G., et al. (1996). "I've Been there, Too": Effect on empathy of prior experience with a need. *Personality and Social Psychology Bulletin*, 22(5), 474-482. doi:10.1177/0146167296225005
- Beck, A. T., Steer, R. A., & Carbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, 8(1), 77-100. doi:10.1016/0272-7358(88)90050-5
- Bumberry, W., Oliver, J. M., & McClure, J. N. (1978). Validation of the Beck Depression Inventory in a university population using psychiatric estimate as the criterion. *Journal of Consulting and Clinical Psychology*, 46(1), 150-155. doi:10.1037//0022-006X.46.1.150
- Coke, J. S., Batson, C. D., & McDavis, K. (1978). Empathic mediation of helping: A two-stage model. *Journal of Personality and Social Psychology*, 36(7), 752-766. doi:10.1037//0022-3514.36.7.752
- Coyne, J. C., & Gotlib, I. (1986). Studying the role of cognition in depression: Well-trodden paths and cul-de-sacs. *Therapy*, 6, 695-705.

- Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, 10, 85.
- Davis, M. H. (1983). The effects of dispositional empathy on emotional reactions and helping: A multidimensional approach. *Journal of Personality*, 51(2), 167-184.
- Demiralp, E., Thompson, R., Mata, J., Jaeggi, S., Buschkeuhl, M., Barrett, L. F., Ellsworth, P., et al. (in press). Feeling blue or turquoise? Emotional differentiation in major depressive disorder. *Psychological Science*.
- Eisenberg, N., & Fabes, R. A. (1990). Empathy: Conceptualization, measurement, and relation to prosocial behavior. *Motivation and Emotion*, 14(2), 131-149.
- Eisenberg, N., Wentzel, M., & Harris, J. D. (1998). The role of emotionality and regulation in empathy-related responding. *The School Psychology Review*, 27(4), 506-521.
- Field, Tiffany. (1995). Infants of depressed mothers. *Infant Behavior and Development*, 18, 1-13. doi:10.1080/10673220902899714
- Fritzsche, A., Dahme, B., Gotlib, I. H., Joorman, J., Magnussen, H., Watz, H., Nutzinger, D. O., et al. (2010). Specificity of cognitive biases in patients with current depression and remitted depression and in patients with asthma. *Psychological Medicine*, 40(5), 815-826. doi:10.1017/S0033291709990948.
- Gawronski, I., & Privette, G. (1997). Empathy and reactive depression. *Psychological Reports*, 80, 1043-1049.
- Graziano, W. G., Habashi, M. M., Sheese, B. E., & Tobin, R. M. (2007). Agreeableness, empathy, and helping: A person x situation perspective. *Journal of Personality and Social Psychology*, 93(4), 583-599.

- Henderson, A. S. & Rickwood, D. J. (2000). Mental health as a national healthy priority area: Focus on depression. *The Medical Journal of Australia*, 172(3), 100-101.
- Hodges, S. D., Kiel, K. J., Kramer, A. D. I., Veach, D., & Villanueva, B. R. (2010). Giving birth to empathy: the effects of similar experience on empathic accuracy, empathic concern, and perceived empathy. *Personality & Social Psychology Bulletin*, 36(3), 398-409. doi:10.1177/0146167209350326
- Hofelich, A. J., & Preston, S. D. (2011). The meaning in empathy: distinguishing conceptual encoding from facial mimicry, trait empathy, and attention to emotion. *Cognition & Emotion*, 26(1), 119-28. doi:10.1080/02699931.2011.559192
- Jackson, P. L., Brunet, E., Meltzoff, A. N., & Decety, J. (2006). Empathy examined through the neural mechanisms involved in imagining how I feel versus how you feel pain. *Neuropsychologia*, 44(5), 752-61. doi:10.1016/j.neuropsychologia.2005.07.015
- Jones, N. A., Field, T., & Davalos, M. (2000). Right frontal EEG asymmetry and lack of empathy in preschool children of depressed mothers. *Child Psychiatry and Human Development*, 30(3), 189-204.
- Joormann, J., & Gotlib, I. H. (2007). Selective attention to emotional faces following recovery from depression. *Journal of Abnormal Psychology*, 116(1), 80-5. doi:10.1037/0021-843X.116.1.80
- Kogut, T., & Ritov, I. (2005). The “identified victim” effect: an identified group, or just a single individual? *Journal of Behavioral Decision Making*, 18(3), 157-167. doi:10.1002/bdm.492

- Lamm, C., Batson, C. D., & Decety, J. (2007). The neural substrate of human empathy: Effects of perspective-taking and cognitive appraisal. *Journal of Cognitive Neuroscience, 19(1)*, 42-58. doi:10.1162/jocn.2007.19.1.42
- Maner, J. K., Luce, C. L., Neuberg, S. L., Cialdini, R. B., Brown, S., & Sagarin, B. J. (2002). The effects of perspective taking on motivations for helping: Still no evidence for altruism. *Personality and Social Psychology Bulletin, 28(11)*, 1601-1610. doi:10.1177/014616702237586
- Marsh, A. A., Kozak, M. N., & Ambady, N. (2007). Accurate identification of fear facial expressions predicts prosocial behavior. *Emotion, 7(2)*, 239-251. doi:10.1037/1528-3542.7.2.239
- McKenna, F. P., & Sharma, D. (1995). Intrusive cognitions: An investigation of the emotional Stroop task. *Journal Of Experimental Psychology: Learning Memory And Cognition, 21(6)*, 1595-1607. doi:10.1037/0278-7393.21.6.1595
- Mccabe, S. B., & Gotlib, I. H. (1993). Attentional processing in clinically depressed subjects : A longitudinal investigation. *Cognitive Therapy and Research, 17(4)*, 359-377.
- Mccabe, S. B., Gotlib, I. H., & Martin, R. A. (2000). Cognitive vulnerability for depression: Deployment of attention as a function of history of depression and current mood state. *Cognitive Therapy and Research, 24(4)*, 427-444.
- O'Connor, L. E., Berry, J. W., Lewis, T., Mulherin, K., & Crisostomo, P. S. (2007). Empathy and depression: The moral system on overdrive. In T. F. D. Farrow & P. W. R. Woodruff (Eds.), *Empathy in Mental Illness* (pp. 49-75). Cambridge University Press.



- O'Connor, L. E., Berry, J. W., Weiss, J., & Gilbert, P. (2002). Guilt, fear, submission, and empathy in depression. *Journal of Affective Disorders*, 71(1-3), 19-27.
- Paelecke-Habermann, Y., Pohl, J., & Lepow, B. (2005). Attention and executive functions in remitted major depression patients. *Journal of Affective Disorders*, 89(1-3), 125-35. doi:10.1016/j.jad.2005.09.006
- Preston, S. D., & Hofelich, A. J. (2012). The many faces of empathy: Parsing empathic phenomena through a proximate, dynamic-systems view of representing the other in the self. *Emotion Review*, 4(1), 24-35. doi:10.1177/1754073911421378
- Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *The Behavioral and Brain Sciences*, 25(1), 1-20; discussion 20-71.
- Preston, S. D., Hofelich, A. J., & Stansfield, R. B. (under review). The ethology of empathy: Emotions displayed and elicited by real targets of need.
- Raghunathan, R., & Pham, M. (1999). All negative moods are not equal: Motivational influences of anxiety and sadness on decision making. *Organizational Behavior and Human Decision Processes*, 79(1), 56-77. doi:10.1006/obhd.1999.2838
- Rook, K. S., Pietromonaco, P. R., & Lewis, M. A. (1994). When are dysphoric individuals distressing to others and vice versa? Effects of friendship, similarity, and interaction task. *Journal of Personality and Social Psychology*, 67(3), 548-59.
- Rosenblatt, a, & Greenberg, J. (1991). Examining the world of the depressed: do depressed people prefer others who are depressed? *Journal of Personality and Social Psychology*, 60(4), 620-629.
- Segrin, C., & Dillard, J. P. (1992). The interactional theory of depression: A meta-analysis of the research literature. *Journal of Social and Clinical Psychology*, 11(1), 43-70.

- Sweeney, P. D., Shaeffer, D., & Golin, S. (1982). Attributions about Self and Others in Depression. *Personality and Social Psychology Bulletin*, 8(1), 37-42.  
doi:10.1177/014616728281006
- Ward, J. H. J. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, 58(301), 236-244.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect : The PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063-1070.
- Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., & Chapman, M. (1992). Development of concern for others. *Developmental Psychology*, 28(1), 126-136.  
doi:10.1037/0012-1649.28.1.126

## **Appendix 4.A**

### **Pilot Depression Study**

#### **Methods**

Twenty-nine adult females participated. Control and depressed women were screened through advertisements in the University of Iowa Hospitals and Clinics newsletter for a neuroimaging study of emotion (the study included the collection of pilot neuroimaging data that are not presented here). Eleven women met the BDI-II criteria for clinical depression (mean age = 44, range: 27-57), and 18 were non-depressed controls (mean age = 27.4, range = 19-45). Participants in the depressed group all had been previously diagnosed with depression (self-reported history; n = 3 also had a comorbid anxiety diagnosis) and all were taking medication for depression (SSRI n = 9, dopamine reuptake inhibitor n = 2).

The general procedure was similar to that described above, but each participant watched only 13 of the patient interviews while lying in a 3T fMRI scanner (the 14<sup>th</sup> video depicted a *sanguine* male and was used as a practice video). After watching each video, participants rated their responses on the level of overall feelings of positive and negative emotion, empathic concern, personal distress, how severe they thought the patient's illness was, and how much help they would offer. They also rated the intensity of their emotional responses to each target. In this pilot study, empathic concern and personal distress were not derived from factor analysis of individual adjectives, as was

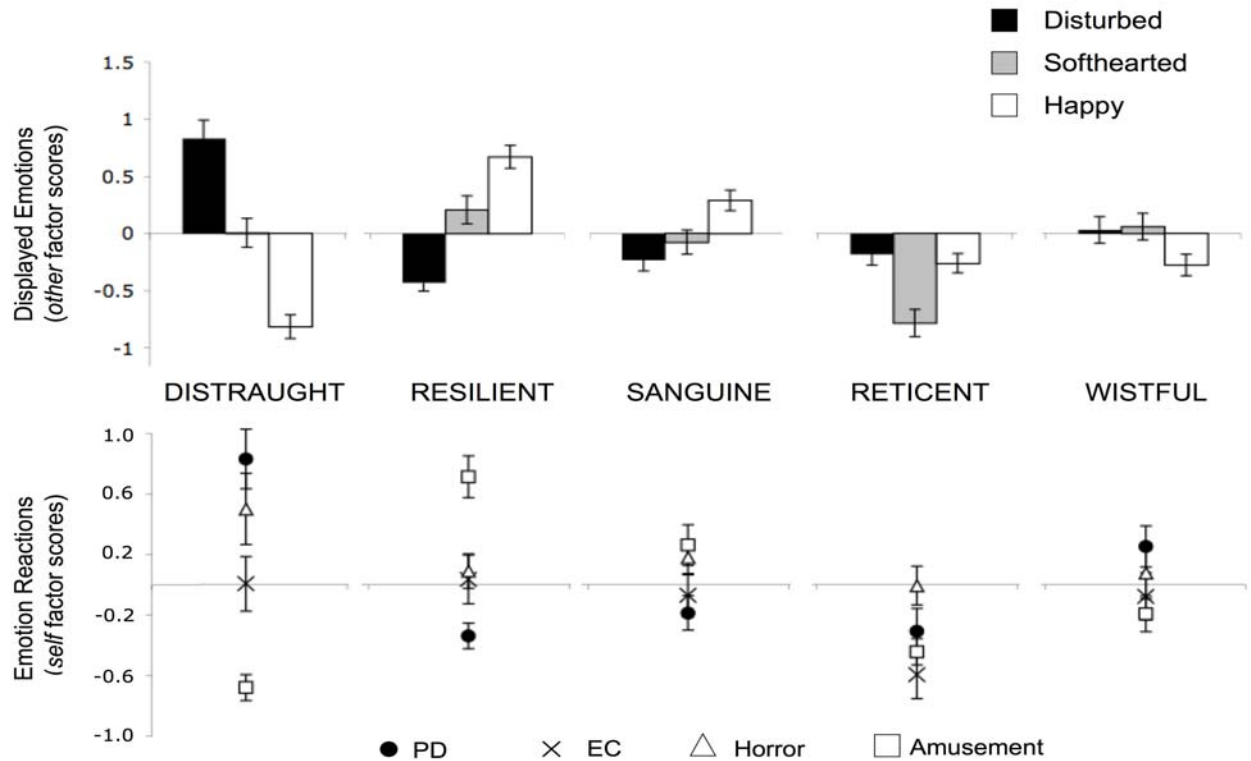
done in Chapter 4, but were rated directly after providing participants with all of the adjectives that typically load onto each term (from Batson, 1987). After the neuroimaging portion of the experiment, participants completed the Interpersonal Reactivity Index (IRI; Davis, 1980) and Mehrabian and Epstein Scale of Emotional Empathy (ME; Mehrabian & Epstein, 1972). All ratings were standardized across videos within each individual. Statistical analyses used an alpha level of .05, and all pairwise comparisons not reported are nonsignificant ( $p > .05$ ).

## Results

### *Confirming patient clustering from Preston et al., under review.*

Observers did not rate the 29 adjectives for the patient's emotion in the pilot study due to time constraints, so targets were classified using ratings of the participants' felt emotions. Control observer ratings of empathic concern, personal distress, positive emotion, and negative emotion were averaged across observers for each target and used to cluster targets into types with the Ward method (as in Preston et al., under review). Despite using a slightly different self-report method, and attesting to the robustness of the groups, the control data from this study produced identical target types previously found, with the same patient targets clustering into the same five groups – *distraught*, *resilient*, *sanguine*, *reticent*, *wistful* (Preston et al., under review; see Figure 4A.1).

**Figure 4.A.1. Displayed emotions and emotional reactions to the patients in Preston et al., under review.**



From Preston et al., under review. Significant emotion factors from PCA analysis of observer ratings of both the targets' displayed emotion (other) and their own response during the video (self) to each target type. The five target types are listed in the center. The three "other" emotion factors are displayed as bars for each cluster in the top graph. The four "self" emotion factors are displayed using symbols in the bottom graph (Personal Distress (PD); Empathic Concern (EC)).

### ***Do depressed women see the patients the same way?***

Repeating the target classification procedure with the data from depressed participants yielded both similarities and differences. Three display types were largely affirmed, with the depressed group having an identical *distraught* group, a *resilient* group that included three of the four original targets, and a cluster that combined the *reticent*

patient from above with a low-expressive patient originally classified as *wistful*. Unlike the non-depressed women above, the best-fit clustering solution for depressed women had only four instead of five groups (using the agglomerative coefficient jump). One would surmise that a cluster could be dropped if two previous groups were agglomerated into one and, supporting this, there was actually one much larger target cluster in the depressed solution. However, this larger group was not derived from two preexisting smaller groups, but rather intermixed individuals from across the three more positive types (*resilient*, *sanguine*, and *wistful*). Moreover, forcing a five-factor solution did not divide this larger group into two, but instead divided the *distraught* group into two even smaller groups. These findings suggest that the depressed observers perceived calm or positive targets as less distinct and highly distressed targets as more distinct.

Confirming this assignment of affective labels to each of the display types identified by depressed women, the *distraught* targets evoked more personal distress (main effect of type,  $F(3, 30) = 11.09, p < .001$ ; see Table 4A.1) and negative emotion compared to the other types (main effect of target type,  $F(3, 30) = 15.94, p < .001$ ). The *distraught* targets were also perceived as being more sick than any other target type ( $F(3,30) = 11.47, p < .001$ ) and were offered more help than every target type but *resilient* (main effect of target type,  $F(3, 30) = 6.96, p = .001$ ), with help to *resilient*, *reticent* and the positive targets not differing ( $p > .311$ ). The *resilient* targets evoked the least negative emotion than all other types ( $m = 1.88; p < .005$ ) except for *reticent* targets, who evoked similar low levels of negative emotion ( $m = 3.0; p = .13$ ). *Resilient* targets also elicited the most positive emotion ( $m = 4.94$ ; main effect of target type,  $F(3, 30) = 14.62, p < .001$ ; post hoc comparisons  $p < .02$ ) and, like previously found (Preston et al.,

under review), the *reticent* targets engendered less empathic concern than all other types ( $m = 3.31$ ; main effect of target type,  $F(3, 30) = 18.0$ ,  $p < .001$ ; post hoc comparisons  $p < .02$ ).

**Table 4.A.1. Mean (standard error) emotional reactions to the common target types in control and depressed participants in pilot study.**

		Common Target Display Types		
		1	2	3
		Distraught	Resilient	Reticent
Controls	Positive Emotion	-0.65 (.14)	0.75 (.09)	-0.87 (.12)
	Negative Emotion	1.08 (.12)	-0.53 (0.12)	-0.023 (0.25)
	Personal Distress	1.08 (0.11)	-0.46 (0.09)	-0.18 (0.21)
	Empathic Concern	0.42 (0.15)	0.13 (0.11)	-0.47 (0.25)
	Perceived severity	0.84 (0.11)	-0.40 (0.14)	-0.38 (0.22)
	Emotion Intensity	0.62 (0.13)	0.11 (0.11)	-0.80 (0.22)
	Help offered	0.19 (0.16)	0.03 (0.12)	-0.59 (0.25)
Depressed	Positive Emotion	-0.75 (0.18)	0.89 (0.12)	-0.63 (0.16)
	Negative Emotion	0.85 (0.16)	-0.71 (0.16)	0.04 (0.33)
	Personal Distress	1.03 (0.15)	-0.46 (0.12)	-0.74 (0.28)
	Empathic Concern	0.83 (0.20)	0.11 (0.15)	-1.2 (0.33)
	Perceived severity	0.91 (0.15)	-0.28 (0.19)	-0.70 (0.29)
	Emotion Intensity	0.8 (0.17)	0.26 (0.14)	-1.42 (0.29)
	Help offered	0.68 (0.22)	0.03 (0.16)	-0.74 (0.33)

***Group differences in responses to patients.***

Because the clustering solutions differed for depressed and non-depressed

observers, we compared their responses only across targets that were assigned to the same group in both populations (including all three *distraught* targets, three of the four *resilient* targets, and the original *reticent* target). Comparisons were done on standardized scores within participant, which were averaged together for all common patients within these target types, creating a single score per measure for each type. RM ANOVAs were then used to compare responses across target types with group as a between-subjects variable. Only the interactions between types and group will be reported, as the main effects of target type were captured in the above analyses and main effects of group are uninformative with standardized responses.

There were significant interaction effects, reflecting differing responses for control and depressed women to specific display types. Depressed observers had more empathic concern than nondepressed for *distraught* displays (depressed  $m = .83$ ; control  $m = .42$ ) and less empathic concern for *reticent* types (depressed  $m = -1.20$ ; control  $m = -.47$ ; interaction of type and group:  $F(2, 46) = 3.39, p = .04$ ). Depressed participants also tended to have slightly higher ratings of emotion intensity for the *distraught* (depressed  $m = .80$ ; control  $m = .62$ ) and *resilient* display types (depressed  $m = .26$ ; control  $m = .11$ ), but less for the *reticent* type (depressed  $m = -1.42$ ; control  $m = -.8$ ; interaction of type by group:  $F(2, 46) = 2.64, p = .08$ ).

Importantly, although the depressed group was older than the control group, ratings of empathic concern and emotion intensity for each of the common display types did not correlate with age (all  $r(27) < -.22, p > .28$ ). On the other hand, BDI-II score was positively correlated with empathic concern for *distraught* and *resilient* targets ( $r(26) > .39, p < .05$ ), and emotion intensity for *resilient* targets ( $r(26) = .61, p = .001$ ).



## References

- Batson, C. D. (1987). Prosocial motivation: Is it ever truly altruistic? *Advances in Experimental Social Psychology*, 20, 65-122.
- Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, 10, 85.
- Mehrabian, a, & Epstein, N. (1972). A measure of emotional empathy. *Journal of personality*, 40(4), 525-43.

## **Chapter 5: General Discussion**

Despite an abundance of need in the world, failures of empathy are not uncommon. From more benign acts like withholding help to more insidious acts of abuse, bullying, or violence, the consequences of failed empathy highlight the importance of understanding what contributes to its presence and absence. This dissertation presented three studies that examined individual differences in the empathic response. These studies specifically focused on understudied interactions between experience-dependent representations of observers and the emotional displays of targets. Together, this work demonstrates that differing representations in observers can produce predictable differences in the empathic response.

What differentiates this work from much of the empathy literature is the use of multiple targets of need combined with the assessment of observer representations. In the studies presented, all of the targets were real hospital patients displaying natural, unstaged responses to questions about their experience with illness. These ecological stimuli can sometimes be limiting, as encountered in Chapter 3 when the gender of targets needed to be manipulated, but they allow the examination of empathic and prosocial responses to more realistic and commonly encountered need than what is typically studied. Historically, empathy has been assessed in response to a single target of need (e.g., Batson, Eklund, Chermok, Hoyt, & Ortiz, 2007; Batson, Quin, Fultz, Vanderplas, &

Isen, 1983; Batson et al., 1996; Coke, Batson, & McDavis, 1978; Davis, 1983; Graziano, Habashi, Sheese, & Tobin, 2007; Maner et al., 2002; Marsh, Kozak, & Ambady, 2007) or to a few prototypical targets designed to elicit specific reactions, such as other-oriented concern or personal distress (e.g., Eisenberg & Fabes, 1990; Eisenberg et al., 1989). More recent work has compared responses to multiple targets designed to vary on a few dimensions, such as the congruence of their situation and emotional response (e.g., Szczurek, 2012; Zaki, Hennigan, Weber, & Ochsner, 2010), the type of pain they are experiencing (e.g., Bruneau, Dufour, & Saxe, 2012), or their general expressivity (Zaki, Bolger, & Ochsner, 2008). In the majority of these studies, the targets are fictional and simply consist of a vignette or short description (with the exception of Zaki et al., 2008; Zaki, Weber, Bolger, & Ochsner, 2009). Although the targets presented in this dissertation certainly do not capture the entire spectrum of need displays commonly encountered in daily life, such a diverse sample both extends and improves upon the methods currently used.

Furthermore, nearly all of the work comparing responses to various targets is done across participants, and, with the exception of modeling variations in trait empathy or expressivity (e.g., Zaki et al., 2008), differences between observers are treated as noise. However, it is clear observers vary widely in their experiences, associations, and representations about the world. As demonstrated in Chapters 2 and 3, even associations believed to be “commonly held” within a society or culture can vary across people, and this variance is not simply noise, as it predicts an individual’s willingness to help one target over another. Therefore, the work presented here demonstrates that in order to achieve a complete understanding of the empathic response, it is important to consider

the correspondence between the representations of observers and the specific emotional displays of targets.

The importance of this correspondence is predicted by the Perception Action Model (PAM; Preston & de Waal, 2002), as the unique personal representations that give rise to an observer's first hand experience of an emotion also underlie their understanding of that emotion in others. The greater the correspondence between the representations of an observer and the emotional displays or situations of a target, the greater the observer's capacity for empathy and understanding (Preston & Hofelich, 2012; Preston & de Waal, 2002). Supporting this, Chapter 2 demonstrated that the extent to which participants held associations between females and sadness predicted their preferential helping of sad, upset female targets, with participants who had the strongest associations giving the most. Chapter 3 replicated this effect, and demonstrated that it was specific to individual differences in associations with sadness, rather than differences in related but non-affective associations about stereotypical roles of females. Furthermore, this effect was not due to general beliefs about which displays are most deserving of aid, as female-sadness associations did not predict preferential helping of *distraught* targets when they were male. Chapter 3 also showed that perceptions of and responses to targets were influenced by commonly held associations about the targets' gender. For example, participants perceived the same targets as more distressed when read as a female rather than a male, and they gave more help to the gender more typical of each emotional display, helping sad females (over sad males) and more positive males (over positive females).

Instead of looking at continuously varying associations between emotional responses and a specific target, Chapter 4 examined how both past and current experiences with negative emotion influenced perceptions of and responses to various targets. Because past experiences influence and shape an individual's representations of specific situations, concepts, states, and emotions, the past experience of depression should impact an observer's representations of negative affect, perhaps providing them a richer understanding of distress (e.g., Preston & Hofelich, 2012; Appendix 4.A). Supporting this, participants with a depression history gave more help to patients they felt shared similar experiences and emotional histories, even though these were not the *distraught* targets as initially predicted. Participants with a depression history also viewed negative affect in targets with greater differentiation than controls without such experience. Their ratings distinguished between two groups of negative *distraught* patients, while controls made distinctions among patients in the larger, more emotionally mixed *reserved* cluster. This perhaps reflects representational changes in participants who have experienced extreme negative affect, resulting in a more differentiated view of negative emotions displayed by the patients or, alternatively, the reasons why the patients were distressed.

Notably, participants who were currently depressed did not seem to have the rich representations of distress that those with a past depression experience had, as they did not similarly differentiate between the emotional displays of *distraught* patients. Although they were currently experiencing a depressed state, the lack of similar results perhaps suggests that these representations of distress become more accessible when the current experience of depression subsides. It is also possible that the current experience

of depression involves not only a depressed mood but also a more self-focused cognitive state, in which participants are acutely aware of, and focus their attention on, their own negative state and needs. This is consistent with research linking people's symptoms of depression to tendencies to ruminate about felt negative states, to the extent to which this rumination interferes with their daily activities (e.g., Nolen-Hoeksema, 2000). The negative self-focus in currently depressed participants thus could be thought of as a goal state in competition with the goal of empathizing or helping the patient targets, and the cognitive processes engaged in maintaining the focus on their own state could have limited the extent to which they spontaneously resonated with the targets. The PAM predicts this can happen much in the same way cognitive processes can be recruited to effortfully take the perspective of another and empathize with them despite having little shared experience (Chapter 1; Preston & de Waal, 2002). Even if currently depressed participants had similarly differentiated representations of negative emotion as individuals with remitted depression, such a self-focused state may mask the extent to which these representations influenced perceptions of, or emotional responses to, the targets in the current study. Future work should disentangle these possibilities, perhaps by explicitly manipulating the current state of participants (for example, inducing a completing, self focused state by directing non-depressed participants to ruminate on a negative event before viewing the targets), or by assessing the differentiation of conceptual representations in other ways that may be less sensitive to differences in current competing state (such as tasks that involve categorizing emotions, rather than responding to emotions in others).

These findings are particularly relevant to work on emotion granularity, a concept that refers to the specificity with which people describe their own emotional experiences (Barrett, Gross, Christensen, & Benvenuto, 2001). Individuals vary in the extent to which they make distinctions between specific felt emotional states (versus distinctions only in general valence or arousal), and increased granularity particularly for negative emotions is associated with better emotional regulation (e.g., Barrett et al., 2001). Recently, it has been shown that people currently experiencing depression have *less* differentiation of their felt negative emotional states and report their emotions at a lower granularity compared to non-depressed controls (Demiralp et al., in press). It is possible that emotional granularity increases with depression remission or is perhaps a characteristic of those who are able to successfully overcome depression. It is also unclear whether a decreased granularity of one's own emotional experiences would necessarily predict a decreased ability to differentiate between emotional states perceived in others, as some evidence suggests that one's conceptual representations of emotions are only modestly associated with the labels one uses for their emotional responses (Suvak et al., 2011). This may also be expected from the depression literature; while many studies find negative biases for self attributions in depression, similar biases are not seen when people with depression judge the attributes of others (e.g., Sweeney, Shaeffer, & Golin, 1982; Weary, Elbin, & Hill, 1987)

Participants who were currently experiencing depression symptoms also did not show the same patterns of giving as participants with a past history of depression. Unexpectedly, these participants responded with less empathy for *distraught* targets who expressed the most negative emotion, while giving to other types of targets was not

affected. It seemed that currently depressed participants also perceived the *distraught* patients to be less in need of help, which statistically explained the negative relationship between current depression and helping. A follow-up study demonstrated that this behavior did not seem to just be a product of a negative mood state, as participants induced to feel sadness did not give less to the *distraught* target. Instead their empathy and, to an extent, their giving behavior, was reduced for both the *distraught* and *resilient* patients. These preliminary results suggest that something specific about the current experience of depression, beyond negative mood state, is driving the effects on giving.

This assertion is not surprising, given that current experience of depression is associated with more than just a negative or sad mood, as people with depression also suffer from loss of pleasure, trouble sleeping, feelings of worthlessness, along with impairments in episodic memory and encoding (e.g., Austin, 2001) that could all contribute to their decreased willingness to help *distraught* patients. These other aspects of depression may manifest as competing goal states according to the PAM, and prevent them from resonating with the *distraught* patients. Notably, these findings are not a result of depression attenuating participants' ability to resonate with others in general, as the decrease in empathy and help was specific to *distraught* targets, and current depression was associated with increased empathy for *resilient* patients.

Although *distraught* patients shared similar emotional experiences with the currently depressed participants (i.e., both were feeling intense negative affect), it is also possible that the similar emotional state was not salient to currently depressed participants, and this preventing them from resonating with these patients. While there is evidence that perceptions of depression differ between people who have and have not



been depressed (Kirk, Haaga, Solomon, & Brody, 2000; Vollmann et al., 2010), it is unknown whether the current experience of depression affects the perception of “ordinary” sadness or distress in others. Perhaps people who are currently depressed perceive such emotions as less severe than never-depressed controls, especially as compared to their own experiences. Such an explanation fits nicely with the finding that depressed participants perceived the need of the *distraught* patients as less, and although depression symptoms were not related to perceiving the patients as less distressed, there may be a disconnect between the emotions perceived in another and the assessment of the severity of the other’s need.

To summarize, the predictions that observers would feel empathy for and help targets whose emotional expressions were compatible with their representations was supported by Chapters 2 and 3, and by participants with a past history of depression in Chapter 4. Supporting the PAM, these effects depended on specific correspondence between the representations of the observers and the emotions displayed by the targets. Participants who were currently depressed showed the opposite of what was expected, and these results were again specific to one type of target – the upset, *distraught* patients – and current depression was not associated with overall changes in empathy or prosocial behavior. According to the PAM, such responses may be a result of different goal states, and although most research on this aspect of the model has focused on explicitly manipulated goal states (e.g., Lanzetta & Englis, 1989), it is interesting to speculate about how other states, such as depression, may also produce conflicting goals states in observers (such as self focus or self blame) that can attenuate empathic responses to overtly distressed others.

It is also possible that representational differences exist between people with past versus current experiences of an emotional state like depression. While the effects attributed to past experience of depression seem to be consistent with the effects of continuously varying conceptual associations in Chapters 2 and 3, the effects associated with current experience of depression were different. If not effects of conflicting goal states associated with current depression, these results may also reflect differences in storage of or access to conceptual knowledge formed through association or past experience, such as the implicit belief that females are sad or memories about relevant but somewhat distant emotional experiences, compared to knowledge associated with currently experienced states. Regardless, each of these effects highlight the importance of considering the interactions between representations of observers and displays targets, as many of these interesting and novel findings are masked when such differences are treated as noise. Examining these interactions is critical to achieve a better understanding of how responses can vary across individuals and situations, and how such differences can give rise to both extraordinary acts and startling absences of empathy.

## References

- Austin, M. P. (2001). Cognitive deficits in depression: Possible implications for functional neuropathology. *The British Journal of Psychiatry*, 178(3), 200-206.  
doi:10.1192/bjp.178.3.200
- Barrett, L. F., Gross, J., Christensen, T. C., & Benvenuto, M. (2001). Knowing what you're feeling and knowing what to do about it: Mapping the relation between emotion differentiation and emotion regulation. *Cognition & Emotion*, 15(6), 713-724. doi:10.1080/02699930143000239
- Batson, C. D., Eklund, J. H., Chermok, V. L., Hoyt, J. L., & Ortiz, B. G. (2007). An additional antecedent of empathic concern: valuing the welfare of the person in need. *Journal of Personality and Social Psychology*, 93(1), 65-74.  
doi:10.1037/0022-3514.93.1.65
- Batson, C. D., Quin, K. O., Fultz, J., Vanderplas, M., & Isen, A. M. (1983). Influence of self-reported distress and empathy on egoistic versus altruistic motivation to help. *Journal of Personality and Social Psychology*, 45(3), 706-718.
- Batson, C. D., Sympson, S. C., Hindman, J. L., Decruz, P., Todd, R. M., Weeks, J. L., Jennings, G., et al. (1996). "I've Been there, Too": Effect on empathy of prior experience with a need. *Personality and Social Psychology Bulletin*, 22(5), 474-482.  
doi:10.1177/0146167296225005
- Bruneau, E. G., Dufour, N., & Saxe, R. (2012). Social cognition in members of conflict groups: behavioural and neural responses in Arabs, Israelis and South Americans to

- each other's misfortunes. *Philosophical Transactions of the Royal Society of London. Biological sciences*, 367(1589), 717-30. doi:10.1098/rstb.2011.0293
- Coke, J. S., Batson, C. D., & McDavis, K. (1978). Empathic mediation of helping: A two-stage model. *Journal of Personality and Social Psychology*, 36(7), 752-766. doi:10.1037//0022-3514.36.7.752
- Davis, M. H. (1983). The effects of dispositional empathy on emotional reactions and helping: A multidimensional approach. *Journal of Personality*, 51(2), 167-184.
- Demiralp, E., Thompson, R., Mata, J., Jaeggi, S., Buschkeuhl, M., Barrett, L. F., Ellsworth, P., et al. (in press). Feeling blue or turquoise? Emotional differentiation in major depressive disorder. *Psychological Science*.
- Eisenberg, N., & Fabes, R. A. (1990). Empathy: Conceptualization, measurement, and relation to prosocial behavior. *Motivation and Emotion*, 14(2), 131-149.
- Eisenberg, N., Fabes, R. A., Miller, P. A., Fultz, J., Shell, R., Mathy, R. M., & Reno, R. R. (1989). Relation of sympathy and personal distress to prosocial behavior: A multimethod study. *Journal of Personality and Social Psychology*, 57(1), 55-66.
- Graziano, W. G., Habashi, M. M., Sheese, B. E., & Tobin, R. M. (2007). Agreeableness, empathy, and helping: a person x situation perspective. *Journal of Personality and Social Psychology*, 93(4), 583-99. doi:10.1037/0022-3514.93.4.583
- Henderson, A. S. & Rickwood, D. J. (2000). Mental health as a national healthy priority area: Focus on depression. *The Medical Journal of Australia*, 172(3), 100-101.
- Kirk, L., Haaga, D. A. F., Solomon, A., & Brody, C. (2000). Perceptions of depression among never-depressed and recovered-depressed People. *Cognitive Therapy and Research*, 24(5), 585-594.

- Lanzetta, J. T. & Englis, B. G. (1989). Expectations of cooperation and competition and their effects on observer's vicarious emotional responses. *Journal of Personality and Social Psychology*, 56(4), 543-554.
- Maner, J. K., Luce, C. L., Neuberg, S. L., Cialdini, R. B., Brown, S., & Sagarin, B. J. (2002). The effects of perspective taking on motivations for helping: Still no evidence for altruism. *Personality and Social Psychology Bulletin*, 28(11), 1601-1610. doi:10.1177/014616702237586
- Marsh, A. A., Kozak, M. N., & Ambady, N. (2007). Accurate identification of fear facial expressions predicts prosocial behavior. *Emotion*, 7(2), 239-251. doi:10.1037/1528-3542.7.2.239
- Nolen-Hoeksema, S. (2000). The role of rumination in depressive disorders and mixed anxiety/depressive symptoms. *Journal of Abnormal Psychology*, 109, 504–511.
- Preston, S. D., & Hofelich, A. J. (2012). The many faces of empathy: Parsing empathic phenomena through a proximate, dynamic-systems view of representing the other in the self. *Emotion Review*, 4(1), 24-35. doi:10.1177/1754073911421378
- Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *The Behavioral and Brain Sciences*, 25(1), 1-20; discussion 20-71.
- Suvak, M. K., Litz, B. T., Sloan, D. M., Zanarini, M. C., Barrett, L. F., & Hofmann, S. G. (2011). Emotional granularity and borderline personality disorder. *Journal of Abnormal Psychology*, 120(2), 414-426. doi:10.1037/a0021808.
- Sweeney, P. D., Shaeffer, D., & Golin, S. (1982). Attributions about self and others in depression. *Personality and Social Psychology Bulletin*, 8(1), 37-42. doi:10.1177/014616728281006

- Szczurek, L. (2012). The stranger effect: Reactions to violators of emotional norms. *Society for Personality and Social Psychology*. San Diego, CA.
- Vollmann, M., Scharloo, M., Salewski, C., Dienst, A., Schonauer, K., & Renner, B. (2010). Illness representations of depression and perceptions of the helpfulness of social support: comparing depressed and never-depressed persons. *Journal of Affective Disorders*, 125(1-3), 213-20. doi:10.1016/j.jad.2010.01.075
- Weary, G., Elbin, S., & Hill, M. G. (1987). Attributional and social comparison processes in depression. *Journal of Personality and Social Psychology*, 52(3), 605-10.
- Zaki, J., Bolger, N., & Ochsner, K. (2008). The Interpersonal Nature of Empathic Accuracy. *Psychological Science*, 19(4), 399-404.
- Zaki, J., Hennigan, K., Weber, J., & Ochsner, K. N. (2010). Social cognitive conflict resolution: contributions of domain-general and domain-specific neural systems. *The Journal of Neuroscience*, 30(25), 8481-8. doi:10.1523/JNEUROSCI.0382-10.2010
- Zaki, J., Weber, J., Bolger, N., & Ochsner, K. (2009). The neural bases of empathic accuracy. *Proceedings of the National Academy of Sciences*, 106(27), 11382-7. doi:10.1073/pnas.0902666106