

Georgia State University

ScholarWorks @ Georgia State University

Psychology Dissertations

Department of Psychology

8-8-2018

Functions of Guilt and Reparative Behaviors in Middle Childhood

Meghan R. Donohue
Georgia State University

Follow this and additional works at: https://scholarworks.gsu.edu/psych_diss

Recommended Citation

Donohue, Meghan R., "Functions of Guilt and Reparative Behaviors in Middle Childhood." Dissertation, Georgia State University, 2018.
https://scholarworks.gsu.edu/psych_diss/179

This Dissertation is brought to you for free and open access by the Department of Psychology at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Psychology Dissertations by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.

FUNCTIONS OF GUILT AND REPARATIVE BEHAVIORS IN MIDDLE CHILDHOOD

by

MEGHAN ROSE DONOHUE

Under the Direction of Erin C. Tully, PhD

ABSTRACT

Guilt theoretically functions to motivate reparative behaviors, which, in turn, theoretically function to alleviate guilt and prevent psychopathology. Although several empirical studies in adults have demonstrated that guilt and reparative behaviors function as theorized, research has not investigated causal relations between guilt and reparative behaviors in children. Thus, this study examined whether guilt motivates children's reparative behaviors, and whether their reparative behaviors successfully alleviate guilty feelings. Six-to ten-year olds ($N = 97$) were randomly assigned to one of three conditions. In the experimental condition, children were led to believe that they had transgressed to cause a peer's distress. Children self-reported their guilt feelings following the pretend transgression, and then had the opportunity to repair the transgression by giving stickers and writing a note to the victimized peer. Following the repair

opportunity, children self-reported their guilt a second time. Children in the experimental condition (i.e., children who felt guilty) engaged in greater reparative behavior relative to children in a no-guilt condition who were led to believe that they had caused a peer's positive emotions. Further, children in the experimental condition reported reduced guilt at the second measurement, whereas children in the no-repair condition (who were led to believe that they had transgressed but were not be given a repair opportunity) did not report reduced guilt over time. Results demonstrate that guilt and reparative behaviors function as theorized in children and may begin to inform reparative interventions aimed at preventing unalleviated, maladaptive guilt and psychopathology.

INDEX WORDS: Guilt, Reparative behavior, Transgressions, Prosocial behavior, Moral emotions, Self-conscious emotions

FUNCTIONS OF GUILT AND REPARATIVE BEHAVIORS IN MIDDLE CHILDHOOD

by

MEGHAN ROSE DONOHUE

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2017

Copyright by
Meghan Rose Donohue
2017

FUNCTIONS OF GUILT AND REPARATIVE BEHAVIORS IN MIDDLE CHILDHOOD

by

MEGHAN ROSE DONOHUE

Committee Chair: Erin Tully

Committee: Lindsey Cohen

Laura McKee

Erin Tone

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

May 2018

ACKNOWLEDGEMENTS

I would like to thank Alley Kyle-Vitonis for her work in rating and coding variables. I am grateful to my committee members, Dr. Lindsey Cohen, Dr. Laura McKee, and Dr. Erin Tone, for their valuable time and thoughtful suggestions. I am forever indebted to Dr. Rebecca Williamson for allowing me to collect data at the Fernbank Museum of Natural History with her lab group for this project. I am also so thankful to my chair and mentor, Dr. Erin Tully, not only for her continued support of my ideas and this project, but for providing me with truly stellar research training throughout my time at Georgia State University. Personally, I thank my best friend, Nada, for her support throughout graduate school, and my father and mother for their support in making this dream a reality. Finally, thank you to the families who generously offered their time in support of child development research.

This work was funded by a Health Resources and Services Administration (HRSA) Graduate Psychology Education Fellowship awarded to Lindsey Cohen and a Georgia State University University Services and Administration (URSA) Dissertation Grant awarded to Meghan Rose Donohue.

TABLE OF CONTENTS

| | |
|--------------------------------------------------------------------------------|-----------|
| ACKNOWLEDGEMENTS | V |
| LIST OF TABLES | X |
| LIST OF FIGURES | XI |
| 1 INTRODUCTION..... | 1 |
| 1.1 Functionalist Perspectives of Guilt | 1 |
| 1.2 The Development of Guilt and Reparative Behaviors in Childhood..... | 7 |
| 1.3 The Measurement of Guilt and Reparative Behaviors in Childhood..... | 11 |
| 1.4 The Development of Guilt and Reparative Behaviors in Childhood..... | 15 |
| <i>1.4.1 Guilt motivates reparative behaviors.....</i> | <i>15</i> |
| <i>1.4.2 Reparative behaviors alleviate guilt.</i> | <i>20</i> |
| 1.5 Overview of the Current Study and Hypotheses..... | 22 |
| 2 METHOD | 23 |
| 2.1 Participants | 23 |
| <i>2.1.1 Recruitment.....</i> | <i>23</i> |
| 2.2 Experimental Design..... | 24 |
| 2.3 Procedures..... | 25 |
| <i>2.3.1 Experimental Conditions</i> | <i>27</i> |
| 2.3.1.1 Distress Feedback/Repair (Experimental) Condition..... | 27 |
| 2.3.1.2 Distress Feedback/No Repair (No-Repair) Condition. | 27 |

| | | |
|--------------|-------------------------------------------------------------------------------|-----------|
| 2.3.1.3 | Satisfied feedback/Repair (No-Guilt) Condition..... | 28 |
| 2.3.2 | <i>Paradigm phases.</i> | 28 |
| 2.3.2.1 | Choice phase. | 28 |
| 2.3.2.2 | Induction phase. | 28 |
| 2.3.2.2.1 | Distress feedback..... | 29 |
| 2.3.2.2.2 | Satisfied feedback. | 29 |
| 2.3.2.3 | Guilt rating phase 1. | 30 |
| 2.3.2.3.1 | Children’s fault question (manipulation check)..... | 30 |
| 2.3.2.3.2 | Children’s free-response report of their emotions (manipulation check). 30 | |
| 2.3.2.3.3 | Guilt scale time 1 (outcome measure)..... | 31 |
| 2.3.2.4 | Repair phase. | 31 |
| 2.3.2.5 | Guilt rating phase 2. | 31 |
| 2.3.2.5.1 | Guilt scale time 2 (outcome measure)..... | 31 |
| 2.3.2.6 | Follow-up phase. | 31 |
| 2.3.2.6.1 | Children’s report of reasons for giving stickers..... | 31 |
| 2.4 | Measures | 32 |
| 2.4.1 | <i>Demographic information.</i> | 32 |
| 2.4.2 | <i>Guilt scale.</i> | 32 |
| 2.4.3 | <i>Repair task.</i> | 36 |
| 2.4.4 | <i>Sorting task.</i> | 38 |
| 2.5 | Materials | 39 |

| | | |
|---------|-----------------------------------------------------------------------|----|
| 2.5.1 | <i>Gift options.</i> | 39 |
| 2.5.1.1 | Attractive toy (A). | 40 |
| 2.5.1.2 | Unattractive toy (B). | 40 |
| 2.5.2 | <i>Feedback images.</i> | 40 |
| 2.5.2.1 | Distress feedback image. | 41 |
| 2.5.2.2 | Satisfied feedback image. | 41 |
| 2.5.3 | <i>Stickers.</i> | 42 |
| 2.6 | Coding and Rating of Data. | 42 |
| 2.6.1 | <i>Children's free-response report of their emotion.</i> | 42 |
| 2.6.2 | <i>Note ratings.</i> | 43 |
| 2.6.3 | <i>Children's apology.</i> | 44 |
| 2.6.4 | <i>Children's report of reasons for giving stickers.</i> | 44 |
| 2.6.5 | <i>Training and reliability of coders/raters</i> | 45 |
| 2.7 | Manipulation checks and predictions. | 46 |
| 2.8 | Piloting participants. | 47 |
| 3 | RESULTS | 48 |
| 3.1 | Descriptive Statistics | 48 |
| 3.1.1 | <i>Correlations among variables.</i> | 50 |
| 3.1.2 | <i>Group differences among variables.</i> | 50 |
| 3.2 | Manipulation Checks | 51 |
| 3.2.1 | <i>Condition differences in children's guilt scale time 1 scores.</i> | 51 |

| | | |
|---------|---------------------------------------------------------------------------------------------|----|
| 3.2.2 | <i>Children’s report of their perceived fault</i> | 51 |
| 3.2.2.1 | Children’s fault question. | 51 |
| 3.2.2.2 | Children’s apology. | 52 |
| 3.2.3 | <i>Children’s free-response report of their emotions</i> | 53 |
| 3.2.4 | <i>Children’s report of reasons for giving stickers</i> | 55 |
| 3.3 | Tests of Study Hypotheses | 56 |
| 3.3.1 | <i>Guilt motivates children’s greater reparative behaviors</i> | 56 |
| 3.3.1.1 | Number of stickers given. | 56 |
| 3.3.1.2 | Note ratings. | 57 |
| 3.3.2 | <i>Children’s reparative behaviors alleviate guilt</i> | 59 |
| 4 | DISCUSSION | 62 |
| 4.1 | Validity of the transgression paradigm | 63 |
| 4.2 | Guilt motivates children’s reparative behaviors | 65 |
| 4.3 | Reparative behaviors alleviate children’s guilt | 68 |
| 4.4 | Toward an empirical model of guilt, reparative behaviors, and psychopathology | 70 |
| 4.5 | Limitations | 71 |
| 4.6 | Summary | 72 |
| | REFERENCES | 74 |
| | APPENDIX | 90 |

LIST OF TABLES

| | |
|----------------------------------------------------------------------------------------------------------------------------|----|
| Table 1. Components of experimental conditions. | 27 |
| Table 2. Reliabilities for Scored or Rated Variables | 43 |
| Table 3. Categories of themes included in children’s report of reasons for giving stickers..... | 45 |
| Table 4. Descriptive Statistics of Primary Study Variables: Means, Standard Deviations, and Ranges or Percentages | 49 |
| Table 5. Correlations among Continuous Variables | 49 |
| Table 6. Emotion Words Contained in Children’s Free Responses by Experimental Condition | 54 |
| Table 7. Themes Endorsed in Children’s Report of Reasons for Giving Stickers by Experimental Condition..... | 56 |

LIST OF FIGURES

| | |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 1. Functionalist theoretical model depicting pathways through which reparative behaviors help determine guilt’s relation to psychological outcomes (Donohue, 2016; unpublished general exam manuscript)..... | 6 |
| Figure 2. Flow chart of study design. Pink boxes indicate control elements of the paradigm. ... | 25 |
| Figure 3. Scree plot for the factor analysis of the guilt scale at time 1 (A) and time 2 (B). | 34 |
| Figure 4. Visual aids for the first binary response. | 36 |
| Figure 5. Visual aids for the second binary response (if the child chose “yes” for the first response). | 36 |
| Figure 6. Still photos of the gift options: A) attractive toy and B) unattractive toy. | 39 |
| Figure 7. Still photos of the distress and slightly satisfied feedback images..... | 41 |
| Figure 8. Mean (\pm SE) number of stickers (A) and mean (\pm SE) note ratings of prosocial themes (B) as a function of experimental condition. * = $p < .05$ | 58 |
| Figure 9. Plot of the interaction of condition and time on children’s rated guilt..... | 62 |

1 INTRODUCTION

Interpersonal transgressions and the painful guilt feelings that often follow are a common human experience. Theorists have extensively posited that guilt motivates reparative behaviors aimed at mending transgressions; in turn, reparative behaviors theoretically alleviate guilt, preventing it from becoming maladaptive and associated with psychopathology (e.g., Quiles & Bybee, 1997). Although several studies in adults have demonstrated that guilt indeed motivates reparative behaviors and that reparative behaviors successfully alleviate guilt (e.g., de Hooge, 2012), only one study has demonstrated that guilt motivates reparative behaviors in children (Colasante, Zuffiano, Bae, & Malti, 2014), and no study has investigated whether reparative behaviors alleviate children's guilty feelings. Investigating reparative behaviors as a means of guilt alleviation in childhood is critical, as children who display greater unalleviated, maladaptive guilt exhibit greater psychopathology, particularly depression (e.g., Tilghman-Osborne, Cole, & Felton, 2012). The purpose of the current study is to examine whether guilt and reparative behaviors function as theorized in middle childhood, the developmental stage that precedes increased rates of depression in adolescence (Thapar, Collishaw, Pine, & Thapar, 2012).

1.1 Functionalist Perspectives of Guilt

Guilt is a cognitive and emotional experience triggered by a moral transgression that involves awareness of fault as well as painful negative emotions (Tilghman-Osborne, Cole, & Felton, 2010). Reparative behaviors¹, a specific type of prosocial behaviors (i.e., voluntary actions aimed at benefitting another person; Eisenberg & Fabes, 1998) are actions transgressors direct toward victims of their own wrongdoings that include making amends (i.e., undoing the

¹ Reparative behaviors will also at times be referred to as reparation behaviors, reparation and repair.

transgression's consequences), confessing and apologizing (Tangney, Stuewig, & Mashek, 2007) as well as other comforting behaviors and verbalizations. Children may experience a range of emotions following transgressions, including fear of punishment, empathic sadness, and guilt. Guilt is an empathy-based emotion that motivates reparative behaviors when a transgressor experiences empathy *and* a sense of personal responsibility for the distress (Zahn-Waxler & Radke-Yarrow, 1990).

Guilt is related to yet distinct from shame. Although the two self-conscious emotions are often at least moderately correlated (Tangney, 1996), they theoretically lead to differing motivational states and behaviors. Most critically for the current investigation, guilt elicits focus on negative aspects of a wrongdoing (“I did that bad *thing*”), motivating reparative behaviors to undo the transgression, whereas shame elicits focus on a bad self (“*I* did that bad thing), failing to motivate reparative behaviors, since a wholly defective self cannot be mended (Barrett, 1995; Lewis, 1971; Tracy & Robins, 2004). Indeed, empirical studies across children and adults have demonstrated that shame is associated with withdrawal, avoidance, denial, and less prosocial behavior (e.g., Barrett, Zahn-Waxler, & Cole, 1993; Tangney, 1998). Because of the unique theoretical relationship between guilt and reparative behaviors, the current study focuses on guilt, largely to the exclusion of shame.

Whether guilt functions adaptively or maladaptively has been widely debated (Tilghman-Osborne et al., 2010). Perhaps surprisingly, many developmental researchers have presented guilt an adaptive emotion, particularly during childhood (Tilghman-Osborne et al., 2010). Indeed, guilt theoretically serves adaptive functions, for example through deterring future transgressions and antisocial trajectories; when children experience guilt after transgressing, wrongdoing becomes “somatically marked” with intense discomfort such that in future

situations, this somatic state is re-enacted, and transgressive behavior is inhibited (Damasio, 1994). Guilt also theoretically facilitates adaptive social behavior when transgressions do occur. Importantly, guilt's primary function is to channel a transgressor's negative feelings into reparative behaviors that mend relationships and prevent negative psychological consequences for the transgressor (Tangney et al., 2007). Processes through which guilt motivates reparative behavior have been proposed; in Tracy and Robins' (2004) process model, guilt feelings and cognitions occur when an individual appraises a transgression event as incongruent with social goals (i.e., the desire to make and maintain relationships) as well as identity goals (i.e., the desire to retain the individual's view of her actual or ideal self), leading to reparative attempts to alleviate negative emotions and resolve discrepancies between the transgression and the individual's goals.

Some studies have demonstrated the adaptive potential of guilt; the trait-like tendency to experience guilt following transgressions (i.e., guilt-proneness) has been associated with lower levels of anxiety, depression, and externalizing problems, and higher levels of self-esteem across children, adolescents, and adults (e.g., Dearing, Stuewig, & Tangney, 2005; Luyten, Fontaine, & Corveleyn, 2002; June Price Tangney, 1991). For example, one study found that guilt-proneness in 8-to 14-year-old children was associated with lower levels of aggressive behavior (Tangney, Wagner, Hill-Barlow, Marschall, & Gramzow, 1996). However, the literature regarding guilt-proneness and internalizing problems is less consistent, with other studies finding no association between guilt-proneness and depression and anxiety (Quiles & Bybee, 1997; Tangney, Wagner, & Gramzow, 1992). Developmental theorists have frequently argued that guilt is only maladaptive when fused with shame (i.e., it is *only* the shame component that is maladaptive),

and that studies that have found associations between guilt and poorer outcomes have not sufficiently distinguished guilt from shame (Tangney et al., 2007).

Yet, others, particularly clinical researchers, emphasize the destructive potential of guilt and recognize maladaptive guilt as a distinct guilt variant. Maladaptive guilt has been defined as a destructive type of guilt that is inappropriate (i.e., inaccurate) and/or excessive (e.g., Tilghman-Osborne et al., 2012). The idea that guilt has the capacity to be maladaptive is central to some clinical research; indeed, maladaptive guilt is a symptom of certain psychological disorders, such as depression (American Psychiatric Association, 2013). Research supports maladaptive guilt as a construct distinct from shame and related to psychopathology. In one study, maladaptive guilt remained significantly associated with a range of psychopathology after shame was partialled from the maladaptive guilt measure (Quiles & Bybee, 1997). Moreover, a recent meta-analysis found that depressive symptoms were significantly associated with maladaptive guilt, $r = .42$ as well as shame, $r = .43$; (Kim, Thibodeau, & Jorgensen, 2011).

These conflicting accounts of guilt can be resolved through a functionalist perspective of the role of emotions in to the development and maintenance of psychopathology. Functionalist theorists posit that emotions cannot be inherently adaptive or maladaptive; rather, their functionality or dysfunctionality depends on how effectively they are managed (Barrett, 1995). Whereas experiences of acute state guilt in response to a transgression that motivate reparative behavior are theorized to prevent negative psychological outcomes, guilt that is unaccompanied by reparative action is theorized to lead to forms of guilt that have been associated with psychopathology (Figure 1; Quiles & Bybee, 1997; Bybee, Zigler, & Berliner, 1996). When individuals cannot or do not engage in reparative behaviors to alleviate guilty feelings,

maladaptive guilt is theorized to develop². In a vicious cycle, maladaptive guilt further blocks reparative action, leading to chronically unalleviated guilt and, over time, psychopathology. Indeed, studies have demonstrated that adults' lesser past use of reparative behaviors in response to a particular transgression was associated with their greater current maladaptive guilt over the transgression (Bassett, Bassett, Kelly, Lloyd, & Johnson, 2006; Exline, Deshea, & Holeman, 2007; Riek, Luna, & Schnabelrauch, 2014; Silfver, 2007), and maladaptive guilt has been associated with higher levels of depression and anxiety, externalizing problems, and poorer social functioning across children, adolescents, and adults (Harder, Cutler, & Rockart, 1992; Jones & Kugler, 1993; Kim et al., 2011; Luby, Belden, Pautsch, Si, & Spitznagel, 2009; Quiles & Bybee, 1997).

² Reparative difficulty is likely one of several processes through which maladaptive guilt develops. For example, it has been extensively posited that children exposed to certain environmental risks, such as being parented by a depressed mother, may develop maladaptive guilt (e.g., inappropriate attributions of personal responsibility for their depressed mother's emotions; Zahn-Waxler & Kochanska, 1990).

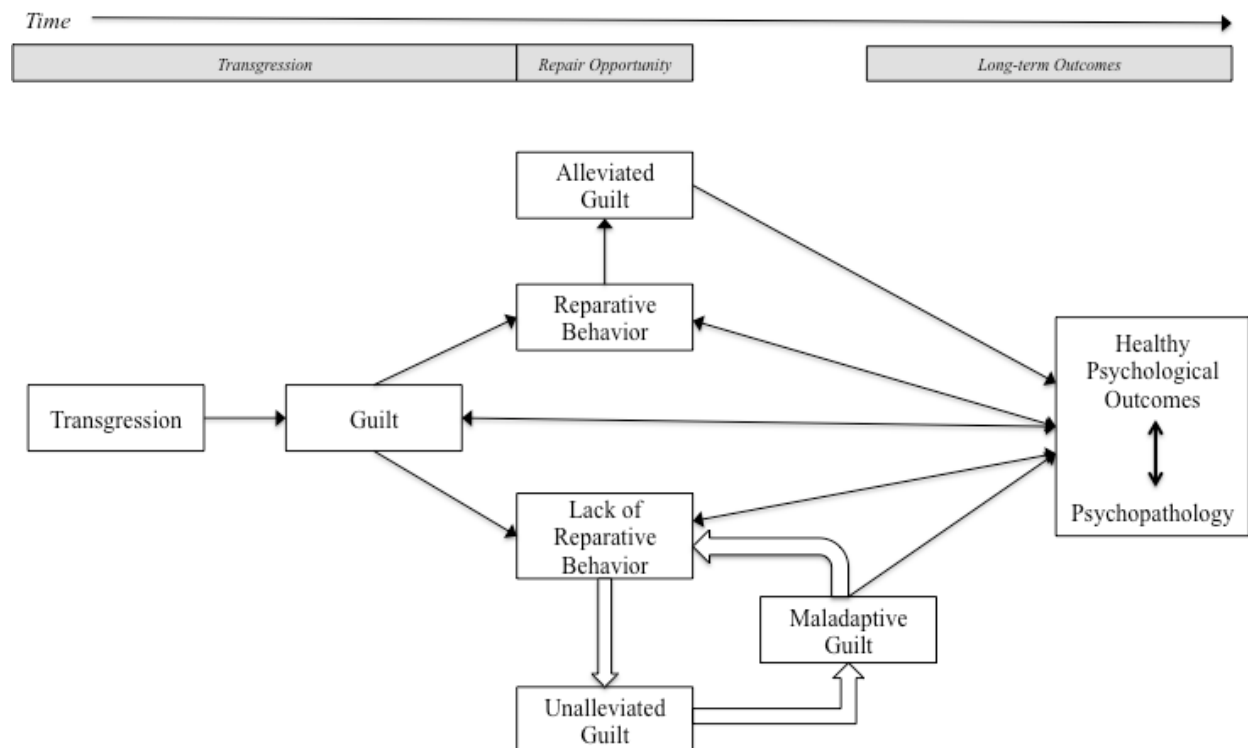


Figure 1. Functionalist theoretical model depicting pathways through which reparative behaviors help determine guilt's relation to psychological outcomes (Donohue, 2016; unpublished general exam manuscript)

Indeed, children experience maladaptive guilt, which likely places them at risk for poorer psychological outcomes, particularly depression. Whereas early childhood is a period in which maladaptive guilt (particularly inappropriate guilt) is typical (Leitenberg, Yost, & Carroll-Wilson, 1986), maladaptive guilt in middle childhood is less normative and more associated with psychopathology. For instance, in one study, 7-to 16-year-old children's greater self-reported maladaptive guilt was associated with greater self-reported depression at all ages, and maladaptive guilt was less normative and more depressotypic with older age (Tilghman-Osborne et al., 2012). Moreover, in another study, clinically depressed children were characterized by both low levels of reparative behaviors and high levels of maladaptive guilt (Luby, Belden, Sullivan, et al., 2009). The current study seeks to empirically test the theoretical claim that

reparative behaviors effectively alleviate guilt in middle childhood; such examinations are particularly important given research demonstrating the potential for children's unalleviated, maladaptive guilt to result in psychopathology, particularly depression.

1.2 The Development of Guilt and Reparative Behaviors in Childhood

Although research is limited, existing studies have demonstrated that guilt and reparative behaviors increase in frequency and complexity throughout childhood (Muris & Meesters, 2014). The current study examined 6-to 10-year-old children's guilt and reparative behavior in response to a transgression against a peer victim. By middle childhood, children have acquired the requisite cognitive skills necessary to experience guilt and engage in reparative behaviors (Muris & Meesters, 2014), they can provide autobiographical narratives of their own guilt experiences (Gavazzi, Ornaghi, & Antoniotti, 2011; Williams & Bybee, 1994), and they demonstrate understanding of what the word "guilt" means (Berti, Garattoni, & Venturini, 2000; Harris, Olthof, Terwogt, & Hardman, 1987), aiding their ability to self-report their own guilt emotions. Moreover, middle childhood is an ideal period in which to investigate children's emotional and behavioral responses to transgressing against another child, as peer relationships become increasingly related to children's psychological adjustment and children gain social independence during this developmental period (Hay, Payne, & Chadwick, 2004; Hymel, Rubin, Rowden, & LeMare, 1990; Lancy & Grove, 2011; Rubin, Bukowski, & Parker, 2006).

A wealth of empirical evidence has documented that around the second year of life, young children begin to experience guilt and engage in reparative behavior following transgressions (Aksan & Kochanska, 2005; Barrett et al., 1993; Barrett, 2005; Garner, 2003; Kochanska, Barry, Jimenez, Hollatz, & Woodard, 2009; Kochanska, Forman, Aksan, & Dunbar, 2005; Kochanska, Gross, Lin, & Nichols, 2002). To assess guilt in toddlers, most studies have

used *in vivo* transgression or “mishap” paradigms in which children are led to believe that they have broken an experimenter’s valued possession, and children’s guilt and reparative behaviors are then coded from videotape. Naturalistic studies have also documented the early emergence of guilt and reparative behavior; in a study by Zahn-Waxler et. al (1979), children as young as 15 months engaged in reparative behaviors such as verbalizations, making amends (i.e., offering food or a bandaid), and giving physical affection. Toddler children’s guilty displays are thought to be a precursor to later guilt (Kochanska et al., 2002).

Throughout early childhood, children’s capacity for guilt and reparative behavior continues to develop alongside cognitive skills such as self-awareness, theory of mind, and understanding of rules that define socially appropriate behavior (M. Lewis, 2008). For example, stable self-representations are thought to emerge between 18 and 24 months (M. Lewis, Sullivan, Stanger, & Weiss, 1989), and studies have found associations between self-awareness and both guilt (Kochanska et al., 2002) and reparative behavior (Zahn-Waxler, Radke-Yarrow, Wagner, & Chapman, 1992). Moreover, throughout early childhood, children develop abilities to infer a victim’s emotions and recognize that socialization agents (e.g., parents) have expectations about their behavior following transgressions; these developing abilities likely facilitate guilt and the use of reparative behaviors (Hay & Cook, 2007; Muris & Meesters, 2014). Shared socialization influences also underlie the development of both guilt and reparative behavior; indeed, authoritative parenting and discipline that is low in power assertion and includes inductive reasoning³ have been associated with greater guilt and reparative responding (Eisenberg, 2000;

³ Inductive reasoning is a discipline practice in which parents provide children with explanations about others’ needs and the effects of harmful acts. Although empirical evidence indicates that inductive reasoning is related to adaptive outcomes such as greater reparative responding in children, it should be noted that parental use of guilt induction has been associated with internalizing problems in children when the inductions are extremely frequent and/or intense, such as when parents emphasize disappointment and encourage an inappropriate sense of responsibility in the child (e.g., Donatelli, Bybee, & Buka, 2007; Rakow et al., 2009).

Kochanska et al., 2002; Kochanska, 1991; Scarnier, Schmader, & Lickel, 2009; Zahn-Waxler et al., 1979). For instance, one study found that mothers' use of inductive reasoning when children were 20-24 months old predicted their reparative behaviors 5 months later (Zahn-Waxler et al., 1979).

Thus, by middle childhood, children possess the cognitive skills necessary to experience guilt and engage in reparative behavior. Yet, guilt and reparative behaviors continue to mature from middle childhood to adolescence; indeed, existing studies suggest that guilt and reparative behaviors become more frequent and sophisticated with older age across this developmental period (Bybee, Merisca, & Velasco, 1998; M. Chapman, Zahn-Waxler, Cooperman, & Iannotti, 1987; Colasante et al., 2014). For example, older age was associated with greater parent-reported guilt and reparative behavior across 4-to 8-year olds (Colasante et al., 2014), and greater self-reported guilt across 10-to 17-year-olds (Bybee et al., 1998). Dimensions of guilt other than frequency and intensity also show age-related changes. Studies have found that whereas children in middle childhood typically experience guilt over externalizing incidents (e.g., aggression) and as a result of others' external evaluations, adolescents more commonly experience guilt over internal thoughts and inconsiderateness and as a result of their own internal standards of behavior (Ferguson, 1991; Gavazzi et al., 2011; Williams & Bybee, 1994). Thus, there is continuity in middle childhood in the types of incidents that elicit children's guilt, and these incidents appear to shift as children enter adolescence. Although very few studies have examined developmental changes in reparative behaviors, children generally engage in increasingly sophisticated prosocial behaviors with older age as they become progressively more sensitive to situations in which prosocial behaviors are needed and desired (Hay & Cook, 2007). One study of

specifically reparative prosocial behaviors found that children's apologies became increasingly elaborate with older age (Ely & Gleason, 2006).

Studies have demonstrated that children's understanding of guilt as an emotion term (i.e., understanding what the word "guilty" means) also increases throughout childhood. As children self-reported how "guilty" they felt in the current study, it was important to identify the age at which children can be expected to understand what the word "guilt" means. Although some researchers have concluded that it is only from age 7 that children understand guilt as an emotion term (Ferguson, 1991; Harris et al., 1987; Harter & Whitesell, 1989), one study found that 5- and 6-year-old children understood what "guilty" meant (Berti et al., 2000). In this study, 5- to 10-year-old children were asked to describe definitions, antecedents, thoughts, action tendencies, and strategies to regulate three different emotions: guilt, shame, and sadness (Berti et al., 2000). Children's responses were coded, and responses about guilt were compared to those of sadness and shame. Five- and six-year-old children accurately described guilt and distinguished guilt from both sadness and shame on all aspects of the emotional process and as well as the older children. The studies that concluded that 5- and 6-year-old children do not possess adequate understanding of guilt as an emotion term (Ferguson, 1991; Harris et al., 1987; Harter & Whitesell, 1989) either did not examine 5- to 6-year olds or exclusively asked children to generate possible antecedents of guilt. In contrast, the study by Berti et. al (2000) asked children about multiple aspects of the emotional process, which may have enabled them to more completely express their knowledge; for example, it is possible that some 5-year-olds have difficulty generating hypothetical guilt antecedents, but are able to describe definitions, thoughts, and actions associated with guilt. In sum, studies suggest that children age 5 and older display

understanding of guilt as an emotion term, which likely aids their ability to directly self-report their guilt.

1.3 The Measurement of Guilt and Reparative Behaviors in Childhood

Although assessment of moral emotions and behaviors is complex, validated measures have been developed to assess guilt and reparative behaviors in children. Most studies have relied on parent-report of guilt and reparative behavior, and existing child self-report measures of guilt exclusively examine children's trait rather than state guilt. The current study was the first to examine a) children's self-report of their state guilt and b) observed reparative behavior following an *in vivo* transgression in children over age three. Examinations of state guilt and resulting reparative behaviors are necessary in order to answer questions about the functional relations between guilt and reparative behavior. Thus, the current study adopted its measurement of guilt and reparative behavior from *in vivo* experimental studies of functional relations between state guilt and reparative behavior in adults.

Guilt and reparative behaviors arise from a specific wrongdoing and therefore must be measured in the context of a well-defined transgression (Cryder, Springer, & Morewedge, 2012). To date, existing studies in adults have utilized three primary methods to create a transgression scenario in a research setting: *autobiographical transgressions* in which participants recall past personal wrongdoings (e.g., hurtful words, betrayals), *hypothetical transgressions* in which participants read transgression vignettes, imagining themselves as the transgressor, and *in vivo transgressions* in which participants are led to believe that they have transgressed against a pretend "victim" in the laboratory. After a transgression scenario is recalled, imagined, or created, participants' state guilt and reparative behavior is self-reported or observed. Existing studies in adults assess state guilt using two primary types of self-report measures: single-

adjective measures in which participants rate the degree to which they feel “guilty” on a likert scale, and measures in which participants rate the degree to which a series of phenomenological descriptions of guilt fit their experience (e.g., *I feel bad about what I did*). Although phenomenological scales are considered slightly preferable, as even adults sometimes use the terms guilt and shame interchangeably (Lindsay-Hartz, 1984), and thus may conflate guilt and shame using single-adjective measures, single-adjective measures are typically preferred in studies using *in vivo* experimental transgressions (Amodio, Devine, & Harmon-Jones, 2007; Cryder et al., 2012), as they are efficient and also more effectively hide the study purpose. Finally, existing studies in adults typically measure reparative behavior using self-report questionnaires; however, studies in adults that use *in vivo* transgressions typically use dictator tasks in which the participant divides resources with the pretend “victim” as a measure of reparative behavior (e.g., Cryder et al., 2012).

Studies of guilt and reparative behavior in children have used differing transgression methodology and measures of guilt and reparative behavior depending on the age of the children studied. Studies examining guilt and reparative behaviors in early childhood have almost exclusively used the previously described *in vivo* “mishap” paradigms (Kochanska et al., 2002), and researchers code children’s state guilt and reparative responses from videotape. Guilty displays in toddlers have been operationalized as signals of discomfort such as avoidance, tension, arousal, distress, lessened positive affect and increased negative affect, and their reparative behaviors have been operationalized as attempts to fix the broken items as well as verbalizations such as confessions and apologies (Kochanska et al., 2002). In one study, children’s coded guilt at 22, 33, and 45 months was significantly and positively associated with

mothers' reports of guilt at 56 months, supporting the validity of the observational guilt measure (Kochanska et al., 2002).

In contrast, studies of guilt in middle childhood and adolescence have almost exclusively used hypothetical transgressions. Several validated hypothetical measures have been developed; these measures assess children's trait-like guilt-proneness rather than their state guilt (Ferguson & Stegge, 1998). In these measures, children are presented with a range of different transgression scenarios (e.g., "*you see your sibling in distress and don't stop to help her*"), and are asked to imagine that they are the transgressor. Following each scenario, children are presented with phenomenological descriptions of guilt (e.g., "*I would feel that I did something wrong*") and other self-conscious emotions (e.g., shame) and are asked to indicate on a likert scale how much each statement would describe themselves in the given scenario. The most frequently used measure of guilt-proneness in children is the Test of Self-Conscious Affect for Children (TOSCA-C; Tangney, Wagner, Burggraf, Gramzow, & Fletcher, 1990), which can be administered to children as young as 8 and exhibits good psychometric properties. Moreover, measures have been developed to assess trait-like guilt-proneness in children as young as 5 that display good internal consistency reliability and validity (e.g., the Self-Conscious Emotions—Maladaptive and Adaptive Scales (SCEMAS; (Stegge & Ferguson, 1994). Notably, the described hypothetical measures assess only guilt and not reparative behaviors in children, and no validated measure of state guilt in children has been developed. The few existing studies of reparative behaviors in school-aged children have almost exclusively relied on parent-report using the reparative behavior scale of the *My-Child-2* questionnaire (Kochanska, DeVet, Goldman, Murray, & Putnam, 1994), a measure of children's emotional and behavioral

responses to their own misbehaviors that displays good reliability and validity in that it was predictive of children's reparative behaviors in the laboratory.

Thus, existing self-report guilt measures in children are designed to tap children's trait-like guilt-proneness rather than state guilt, which precludes their use in an *in vivo* experiment. Further, although the existing hypothetical measures have been validated and display good psychometric properties, they are not without problems (Ferguson & Stegge, 1998). First, although hypothetical measures tell us how children expect that they would react in situations, we do not know how they would react to a real or simulated transgression. Second, guilt scores on these measures may be confounded by empathic perspective taking; for example, children may report high degrees of guilt not because they would truly feel that guilty after a transgression, but rather because they are able to imagine that a child who committed the hypothetical transgression would feel that guilty (Ferguson & Stegge, 1998). Finally, hypothetical scenarios do not enable the examination of causal mechanisms that are possible when using experimental designs. The current study's use of a novel *in vivo* transgression enabled examination of a) children's actual, rather than hypothetical affective and behavioral responses to transgressing, and b) causal relations between guilt and reparative behavior.

As there are no existing experimental studies of guilt and reparative behaviors in children, the current study adopted elements from existing observational studies of toddlers and experimental studies of adults. The transgression methodology employed is similar to the "mishap" paradigms used in toddlers in that children were led to believe that they had transgressed. However, a novel, developmentally appropriate paradigm was created for the current study, as researchers have reported that children over 45 months old were not convinced that they had transgressed during the "mishap" paradigms (e.g., older children realized that the

items were already broken and that they did not break them; Kochanska et al., 2002). In contrast to “mishap” paradigms used in toddlers that employ observational coding of guilt and reparative behavior, the older children examined in the current study self-reported their guilt feelings using a scale that combined aspects of phenomenological and single-adjective items. Finally, similar to experimental studies of adults (Cryder et al., 2012), the current study examined children’s actual reparative behavior following a transgression using a dictator task (see Measures, 2.4).

1.4 The Development of Guilt and Reparative Behaviors in Childhood

Although theorists have long hypothesized that guilt motivates reparative action and that reparative behaviors alleviate guilt (Tangney et al., 2007), nearly all of the empirical tests of these theoretical claims have occurred over the past ten years (e.g., Amodio et al., 2007; Cryder et al., 2012; de Hooge, 2012). This recent empirical evidence for theoretical claims that guilt motivates reparative behaviors and that reparative behaviors alleviate guilty feelings will now be reviewed.

1.4.1 Guilt motivates reparative behaviors.

All existing studies have found that greater guilt was related to greater reparative behavior, providing strong support for theories positing that guilt motivates repair. However, only two studies have demonstrated this association in children, and neither of the child studies tested causal mechanisms using an experimental design.

Across four cross-sectional studies using American, Dutch, Argentinian, Belgian, Finnish, and Polish samples, adult participants who reported experiencing greater guilt following autobiographical transgressions engaged in greater apology, confession, and making amends (Bassett et al., 2006; Leunissen, De Cremer, Folmer, & van Dijke, 2013; Riek, 2010; Silfver-Kuhalampi, Figueiredo, Sorthaix, & Fontaine, 2015). Studies have also found that adults’

greater report that they would experience guilt if they were to commit a hypothetical transgression (e.g., sleeping through a group presentation) was associated with their greater report that they would engage in reparative behaviors (Cryder et al., 2012; Ghorbani, Liao, Çayköylü, & Chand, 2013). To study relations between guilt and reparative behavior over time, one longitudinal study included only adult participants for whom an autobiographical transgression was not yet resolved (Riek et al., 2014). The American undergraduate participants self-reported guilt and reparative behavior at an initial lab visit and again, one month later. At baseline, guilt, even when controlling for shame and reparative behaviors, predicted making amends, apology and confession one month later, providing support that guilt predicts reparative behaviors.

Two studies have demonstrated that guilt predicts reparative behavior in adults using *in vivo* transgressions (Amodio et al., 2007; Cryder et al., 2012). In one study, American undergraduates students had an electroencephalogram (EEG) recorded during a guilt induction and subsequent repair opportunity (Amodio et al., 2007). After viewing racially diverse male faces, participants were given fake feedback that their EEG indicated negative brain responses to Black faces. Participants self-reported their guilt and other emotions (e.g., shame) at baseline and again following the fake feedback. Participants were then told that the experiment was over early, but were asked if they would help rate their interest in articles for a future study; reparative behavior was operationalized as participants' reported interest in reading articles related to prejudice reduction. Greater guilt following the fake feedback was significantly associated with greater reported interest in reading the prejudice reduction articles. Results of the EEG indicated that increases in guilt following the feedback (compared to baseline) predicted increased left-sided frontal cortical asymmetry, an indicator of approach motivation, and that this asymmetry

predicted greater desire to read the prejudice-reducing articles. Thus, this study demonstrated that guilt is associated with increased approach motivation and reparative behavior.

In an experimental study, American adults first viewed a computer screen that presented extensive “background information” describing study procedures (Cryder et al., 2012). It was assumed that participants would not read this information; this assumption was verified, as participants spent less than a minute reading the screen. Subsequently, participants chose whether to eat apple or vomit flavored jellybeans. Participants were randomly assigned to one of two conditions; they were told that, as described in the “background information,” another participant would have to eat either the flavor they chose (control condition) or did not choose (guilt condition)⁴. Participants rated their guilt and then completed a dictator task in which they divided five dollars with the other participant as the measure of reparative behavior. Participants in the guilt condition gave significantly more money to the other participant than control participants, demonstrating a causal relation between greater guilt and greater reparative behavior. Importantly, both studies that used *in vivo* transgressions demonstrated that the paradigms successfully induced guilt. For example, in the study by Cryder et. al (2012), participants in the guilt condition reported significantly more guilt following the transgression than any other negative emotion, including shame, and significantly more guilt than control participants.

Two studies have demonstrated that greater guilt was related to greater reparative behaviors in children ages 4-to 11 (Colasante et al., 2014; Ferguson, 1991). In the study by Ferguson (1991), Dutch fifth grade children ($M_{\text{age}} = 11.2$) imagined themselves as the transgressor in four hypothetical scenarios. Following each scenario, children indicated whether they would feel guilty and whether they would feel ashamed, and also created a resolution to

⁴ Participants who chose to eat the vomit flavored jellybeans were excluded from analyses.

each transgression scenario. Children's resolution narratives were coded for whether children indicated that they would repair the transgression or whether they would let another person repair the transgression. More children indicated a desire to repair the situation themselves in response to stories in which the majority of children reported guilt compared to stories in which the majority of children reported shame. Although this study supports a link between guilt and reparative behavior, it is unclear in the authors' reporting of the methodology whether children were prompted to discuss reparative behavior when creating a resolution to each scenario, which may have encouraged socially desirable responses and inflated the extent to which guilt was associated with reparative behavior.

In the study by Colasante (2014), researchers presented 4- and 8-year-old Canadian children with hypothetical transgression vignettes (e.g., grabbing a toy from a classmate) and asked children to label the emotion they would feel as the transgressor and indicate its intensity by pointing to a visual 3-point likert scale depicting squares that increased in size. Children's responses were coded as guilty if they indicated that they felt guilty, bad, or sad, and children received 1-3 points depending on which square they pointed to. Children's report that they felt any other emotion received a score of 0. Parents also reported on their children's typical guilt and reparative behavior following transgressions. Greater parent but not child-reported guilt was associated with children's greater reparative behavior. The authors discussed the possibility that child-reported guilt was not significantly associated with reparative behavior because the measure of reparation did not correspond to the hypothetical transgressions, as is recommended by guilt researchers (Tangney, 1996); for example, rather than asking children to report what they might do in response to the hypothetical transgressions, parents reported on children's general tendencies to engage in reparative behavior. In the current study, children were given an

opportunity to engage in reparative behavior that directly corresponded to the *in vivo* transgression (i.e., they were able to give stickers to a pretend peer that they were led to believe they had wronged).

About half of the reviewed studies also examined whether guilt is *uniquely* associated with reparative behavior relative to shame and other negative emotions. For example, using partial correlations, Silfver et. al (2015) found that guilt, when controlling for shame, predicted reparative behaviors, whereas shame, when controlling for guilt, was negatively and nonsignificantly associated with repair. Moreover, in the study by Amodio et. al (2007), the authors showed that no other negative emotion (shame, sadness or anxiety) was associated with frontal asymmetry or reparative behavior following the fake feedback. These findings support theories that shame does not motivate reparative behavior (Lewis, 1971) and suggest that it is guilt specifically and not negative affect generally that motivates reparative behavior.

To summarize, consistent with theory, greater guilt was related to greater use of reparative behaviors in multiple studies of adults from several cultures. Some studies also used longitudinal and experimental designs, supporting the direction of the effect. However, there have only been two studies demonstrating that guilt is associated with greater reparative behavior in children, and these studies were limited by less than ideal measurement of these constructs. This measurement weakness obscures whether or not guilt actually motivates reparative behavior in childhood and points to a need for additional examinations of the relation between guilt and reparative behavior during this developmental period. Finally, the current study was the first to use an experimental design to examine whether greater guilt causes greater reparative behavior in children.

1.4.2 Reparative behaviors alleviate guilt.

Guilt alleviation is clearly a crucial theoretical function of reparative behaviors. In contrast to studies examining guilt as a motivator of reparative behaviors, many fewer studies have examined whether reparative behaviors successfully alleviate guilt, and no studies have examined this association in children. However, all six studies that have investigated this relation in adults have found support for the claim that reparative behaviors alleviate guilt.

Two studies utilizing autobiographical transgressions have demonstrated that reparative behaviors alleviate guilt. In one study, American undergraduates experienced two imagery conditions, one in which they imagined a transgression and feelings associated with hurting the victim (transgression condition) and one in which they imagined taking steps to repair the transgression (forgiveness-seeking condition; Witvliet et al., 2002). Participants then rated their guilt. Participants reported less guilt following forgiveness-seeking than transgression imagery. However, because the study did not report baseline levels of guilt it could not demonstrate that guilt was *reduced*. In another study, participants wrote about ways they might repair an autobiographical transgression (Witvliet, Hinman, Exline, & Brandt, 2011) and completed baseline and post-writing paradigm measures of their current guilt over the transgression. Compared to baseline, participants' guilt decreased significantly following the writing paradigm; thus, in contrast to the previous study, this study suggests that contemplating reparative action reduces guilt.

The remaining four studies utilized hypothetical transgression methodology. Three studies demonstrated reductions in guilt by comparing participants' levels of guilt following a hypothetical transgression vignette to their guilt after they imagined engaging in reparative behavior; participants' guilt decreased significantly following imagined attempts to make

amends (de Hooge, 2012) and confess (Meek, Albright, & McMinn, 1995). For example, in one of the two studies reported in de Hooge et al. (2012), Dutch participants imagined that they lost a friend's bicycle, and then rated their guilt following the vignette and after imagining repairing the transgression (i.e., by offering a replacement bicycle); participants' guilt significantly decreased following repair. In the three studies using hypothetical transgression vignettes, measuring guilt immediately following the vignette served as a "baseline" and also demonstrated that the vignettes successfully induced guilt. The study by de Hooge et. al (2012) also demonstrated that the vignette *uniquely* induced guilt, as participants reported significantly more guilt than any other emotion, including related emotions such as shame. In another study, participants imagined that they let a friend get fired over something they did (Carpenter et al., 2014). Next, participants were randomly assigned to imagine that they made amends and apologized (repair condition) or did not do so (no-repair condition); repair participants reported significantly less guilt than no-repair participants (Carpenter et al., 2014).

Thus, all six studies supported the idea that reparative behaviors reduce guilt using both within and between subjects designs and in the context of samples from American and Dutch cultures. However, a primary limitation of this body of work is that no study has demonstrated that reparative behaviors successfully alleviate guilt prior to early adulthood. Additionally, the studies all examined participants' imagined, rather than actual reparative behaviors, and no study examined causal mechanisms in the context of an *in vivo* experimental design. The current study represents an important extension of this body of work as it was the first to examine a) relations between reparative behaviors and alleviated guilt in children, b) participants' actual repair following an *in vivo* transgression, and c) whether reparative behaviors *cause* reductions in guilt.

Thus, although existing studies of the functions of guilt and reparative behaviors provide promising support for theoretical perspectives, there have been surprisingly few investigations of these functions in children. However, as techniques to induce guilt through transgression scenarios have been developed and children have the capacity to self-report their guilt feelings, such examinations are now possible. The examination of functional relations between guilt and reparative behaviors is an important first step in evaluating the theoretical claim that guilt motivates reparative behaviors, which in turn effectively alleviate guilt and protect against psychopathology. Evidence indicating that reparative behaviors alleviate guilt in children may facilitate future investigations of whether these behaviors protect against psychopathology through this guilt alleviation. Findings may also begin to inform the development of reparative interventions. Finally, evidence that guilt motivates reparative behaviors in children might suggest a benefit in facilitating parents' and teachers' use of socialization practices known to foster children's development of healthy guilt.

1.5 Overview of the Current Study and Hypotheses

In summary, although guilt theoretically functions to motivate reparative behaviors and reparative behaviors are in turn theorized to alleviate guilty feelings, no studies have tested causal relations between guilt and reparative behaviors in children. The purpose of the current study is to examine whether guilt and reparative behaviors function as hypothesized in children. Examining whether reparative behaviors alleviate guilt in children is particularly important given research demonstrating that children can experience unalleviated, maladaptive guilt, and that maladaptive guilt is associated with psychopathology. The current study utilized a field experiment and transgression paradigm in which children were led to believe that they had transgressed to cause a peer's distress. Children were interviewed about their guilt feelings

following the pretend transgression, and then had the opportunity to repair the transgression by dividing stickers between themselves and the victimized peer. Finally, children were interviewed about their guilt a second time following the repair opportunity.

The *first hypothesis* is that children who are led to believe that they have transgressed to cause a peer's distress (i.e., children for whom guilt has been induced) will engage in greater reparative behaviors relative to children who are led to believe that they have caused a peer's positive emotions. The *second hypothesis* is that among children who are led to believe that they have transgressed, children who are given an opportunity to repair their transgression will be more likely to report reduced guilt relative to children who are not given a repair opportunity.

2 METHOD

2.1 Participants

Participants were a community sample ($N = 98$) of 6-to 10-year old children ($M_{age}=7.94$ years; $SD=1.37$ years). One child was excluded from analyses—the child laughed following the feedback that he had committed a transgression, suggesting that he may not have been deceived—resulting in a final sample of 97 children. Children participated at the Fernbank Museum of Natural History in Atlanta, GA. The sample size was selected based on a power analysis. The sample included 58 girls (59.2%). Parents reported the ethnicity of children as follows: 59.8% White, 11.3% Bi/multi-racial; 10.3% Black, 7.2% Other, 4.1% Hispanic/Latino, 3.1% Asian, and 4.1% not reported.

2.1.1 Recruitment

Child participants were recruited during their visit to the museum. Researchers invited children to participate in tasks related to child development, including the current study. Researchers described to the parents the purpose of the research and each of the tasks, and

described the required total time commitment (10 minutes) and child compensation (small toys collectively worth approximately \$2) for participating in all of the tasks. Then, researchers asked parents whether they would allow us to collect their child's data for research purposes. For parents who agreed, researchers completed informed consent and assent before beginning data collection (see Procedures, 2.3). The specific tasks that children participated in depended largely on their eligibility (i.e., age) and whether the children chose to complete more than one study. Thus, the sample of the current study is similar, but not identical to, the samples from the other studies.

2.2 Experimental Design

Figure 2 displays a flow chart of the overall study design. Children experienced either positive events or pretend transgressions against a peer, depending on experimental condition. The novel paradigm created for the current study was adapted from previous research in adults; it mirrored that of the previously described Cryder et. al (2012) study in that children were led to believe that a choice that they had made either caused another child to experience either slightly positive or negative emotions.

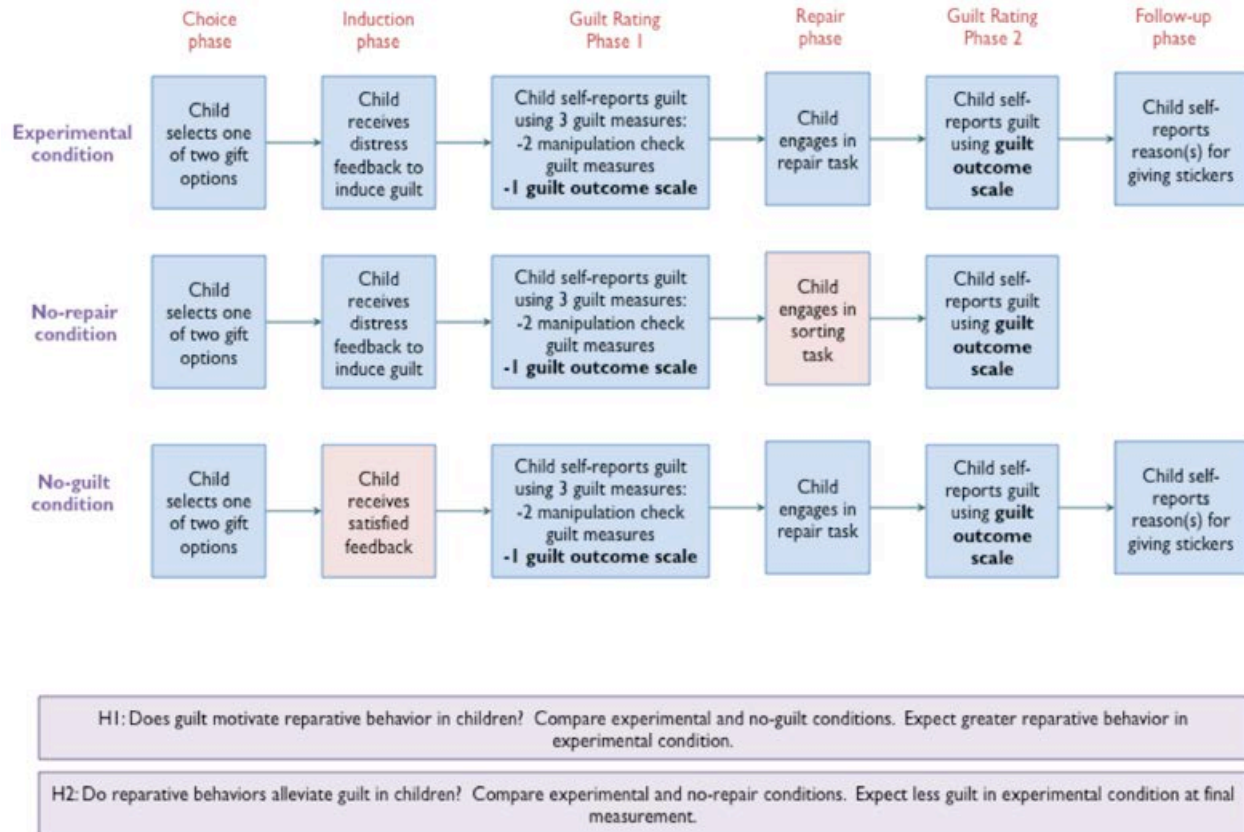


Figure 2. Flow chart of study design. Pink boxes indicate control elements of the paradigm.

2.3 Procedures

Informed consent and assent were first obtained when families approached the testing station near the Nature Quest exhibit. Parents and children consented and assented to participate in any/all of the tasks, including the current study. Parents were told that participation was voluntary, and that they were able to withdraw their children from the study at any point for any reason, including child distress, without loss of benefits. Parents were also told that if they did not want researchers to use their child's data for research purposes, children would still have the opportunity to engage in the tasks. Consent was also obtained from parents to gather demographic information. Assent was obtained from children ages 6 years and older;

researchers explained the tasks and told children that they could stop the tasks at any time. Children were then allowed to verbally assent to the study.

The researcher then led the child to the testing station. The current study's paradigm consisted of several phases (see Paradigm phases, 2.3.2). All children were given a choice between two gift options. Then, children were led to believe that their gift choice determined which gift a pretend peer received, and the researcher gave distress or satisfied feedback about how the pretend peer was feeling after receiving the gift using the distress or satisfied feedback image. Children then reported their guilt feelings using three guilt measures: two guilt measures served as manipulation checks to ensure that the distress feedback successfully induced guilt, and one measure, the guilt scale created for this paradigm, was used as the guilt outcome measure. Next, children either divided stickers between themselves and the pretend peer and wrote a note to the pretend peer (i.e., repaired the transgression), or divided stickers between two envelopes and wrote a note to the next child to visit the testing station. Children reported their guilt a second time using the guilt scale (outcome measure). Finally, children reported their reasons for giving stickers—a final manipulation check to ensure that giving stickers reflected reparative behavior. Children then either moved on to another task, or left the testing station. Thus, participants in the current study may have also participated in other tasks at the testing station such as tasks involving children's social decision-making. Tasks at the testing station did not occur in any standard order across participants. To reduce the potential for observer bias, a modified task without deception was used for children not eligible for the study (e.g., siblings of participants, children outside of the age range). In this task, children simply divided stickers between themselves and a pretend peer. No children indicated that they heard about the task parameters before participating.

2.3.1 *Experimental Conditions*

Prior to beginning the current study, children were assigned to one of three experimental conditions: experimental, no-repair, or no-guilt. Restricted random assignment ensured an approximately equal number of girls and boys in each condition, as some studies have indicated that as early as 32 months of age, girls display more guilt than boys during observational *in vivo* transgression paradigms (Kochanska et al., 2002). Table 1 displays a breakdown of the components of each experimental condition.

Table 1. *Components of experimental conditions.*

| | | Repair | No Repair |
|--------------|--------------------|--------------|-----------|
| Guilt | Distress feedback | Experimental | No-Repair |
| | Satisfied feedback | No-Guilt | N/A |

2.3.1.1 *Distress Feedback/Repair (Experimental) Condition.*

In this condition, the researcher induced children's guilt by giving children the distress feedback. Children were given the opportunity to divide stickers between themselves and the pretend peer and write a note to the pretend peer (i.e., repair the transgression).

2.3.1.2 *Distress Feedback/No Repair (No-Repair) Condition.*

In this condition, the researcher induced children's guilt by giving children the distress feedback. Children were not given the opportunity to repair the transgression; instead, they simply sorted stickers into two envelopes and wrote a note to the next child to visit the testing station. The experimental condition was compared to the no-repair condition to examine whether children reported reduced guilt when they had been given an opportunity to repair a transgression compared to when had not had such an opportunity. Children in this condition

were given an opportunity to repair the transgression after they completed the paradigm in order to minimize any distress.

2.3.1.3 Satisfied feedback/Repair (No-Guilt) Condition.

In this condition, guilt was not induced; instead, the researcher induced children's slightly positive feelings by giving children the satisfied feedback. Children were given the opportunity to divide stickers between themselves and the pretend peer and write a note to the pretend peer. The experimental condition was compared to the no-guilt condition to examine whether children who experienced guilt engaged in more reparative behavior than children who did not experience guilt.

2.3.2 Paradigm phases.

Children in all conditions progressed through the following phases:

2.3.2.1 Choice phase.

All children made a choice between the two gift options to take home. At this time, children also chose between two sets of five scratch and sniff stickers each (the two sets were different scents). Although these stickers were used later during the repair phase, they were given to the child early in the procedure to establish children's ownership over the stickers (Smith, Blake, & Harris, 2013). The researcher then put the toy and stickers aside and told the child he/she would get them back later, in order to ensure that children were not distracted during the remainder of the paradigm.

2.3.2.2 Induction phase.

The researcher pretended to receive a text message from another researcher. Using a script, the researcher explained to the participating child that the toy that he/she chose determined which toy another child who was participating in the study in a different room,

Emma, received. Using the pretend text message and either the distress or satisfied feedback image, the researcher delivered either distress or satisfied feedback to the participating child about Emma's emotions after receiving the toy in order to induce guilt or a slightly positive emotion in the participating child.

2.3.2.2.1 Distress feedback.

The researcher said, "Oh no! I just got a text message from my friend who is playing this game with another kid named Emma in the other room. She says that you took the last light up ball⁵, and so Emma has to take the small ball home. Emma really wanted a light up ball." The researcher showed the child the distress feedback image, and, in order to direct the participant's attention to Emma's emotions and ensure emotion understanding, asked the participating child, "how does Emma look like she feels about that?" After the child's response, the researcher said, "she looks like she's feeling sad and disappointed" to either acknowledge that the child is correct if he/she named a negative emotion, or correct the child if he/she did not indicate that Emma was distressed.

2.3.2.2.2 Satisfied feedback.

The researcher said, "Oh! I just got a text message from my friend who is playing this game with another kid named Emma in the other room. She says that because you picked the light up ball, Emma was given a light up ball, too. Emma really wanted a light up ball." The researcher showed the child the satisfied feedback image. Again, in order to ensure the child's attention and emotion understanding, the researcher asked the participating child, "how does Emma look like she feels about that?" After the child's response, the researcher said, "she looks like she's feeling a little happy" to either acknowledge that the child was correct if he/she named

⁵ Example script. The script differed if the participating child had chosen the small ball.

a positive emotion, or correct the child if he/she did not indicate that Emma was feeling at least slightly positively.

2.3.2.3 Guilt rating phase 1.

Children self-reported their guilt following the satisfied or distress feedback using three guilt measures. Two measures served as manipulation checks to ensure that the distress feedback induced guilt. The third measure was a guilt scale that served as the guilt outcome measure.

2.3.2.3.1 Children's fault question (manipulation check).

To assess children's awareness of fault/responsibility, a key component of guilt (Tilghman-Osborne et al., 2010), the experimenter then asked children in all experimental conditions, "Is Emma feeling [sad/happy] because of you, or not because of you?" The experimenter used the specific negative or positive emotion word that the child reported to describe Emma's feelings in the induction phase (2.3.2.2). The experimenter waited for the child's response and prompted twice by repeating the question if the child did not respond. Children's responses to this question were given a binary score corresponding to because of you (1) or not because of you (0).

2.3.2.3.2 Children's free-response report of their emotions (manipulation check).

To assess the other component of guilt, the experience of negative emotion (Tilghman-Osborne et al., 2010), the experimenter then asked children in all experimental conditions, "How do *you* feel about what happened with Emma?" Children's narrative responses to this question were coded (see 2.6, Coding and Rating of Data).

2.3.2.3.3 Guilt scale time 1 (outcome measure).

The researcher then said to the child, “I’m going to ask you some more questions about how you feel *right now* about what happened with Emma.” The researcher showed the child the distress or satisfied image to ensure that the child remembered Emma. As described in Measures (2.4), the researcher then asked the child to rate the extent to which he/she felt guilty using a 4-item, 5-point likert scale.

2.3.2.4 Repair phase.

The child then engaged in either the repair or sorting task described in Measures (2.4), depending on experimental condition.

2.3.2.5 Guilt rating phase 2.

2.3.2.5.1 Guilt scale time 2 (outcome measure).

The researcher interviewed children in all experimental conditions about their guilt a second time using the guilt scale (2.4).

2.3.2.6 Follow-up phase.

2.3.2.6.1 Children’s report of reasons for giving stickers.

Children in conditions that contained the repair task (experimental, no-guilt) were asked by the experimenter, “Why did you decide to share stickers with Emma?” Responses to this question were transcribed and coded (see 2.6, Coding and Rating of Data).

2.4 Measures

2.4.1 Demographic information.

Parents reported basic demographic information. Information gathered included children's age, sex, and race/ethnicity. Per an agreement with the museum, only children's ages in years were collected.

2.4.2 Guilt scale.

Because a validated measure of children's state guilt does not exist, a state guilt scale was created for the current study. The scale was composed of one single-adjective item and three phenomenological items, as both types of items afford strengths and weaknesses. Single-adjective items are those which explicitly ask participants how guilty they feel, using the word guilt (e.g., "do you feel guilty?"). The single-adjective item in the current study asks children, "Do you feel guilty about what happened with Emma?" A single-adjective item was included in order to very directly ask children about guilt feelings. As previously discussed, research indicates that children as young as five understand what the word "guilty" means (Berti et al., 2000). Moreover, most experimental studies in adults have used single-adjective scales to measure state guilt following an *in vivo* transgression (e.g., Cryder, Springer, & Morewedge, 2012; de Hooge, 2012).

The three phenomenological items were adapted from items from validated measures of trait guilt in adults and children. The first item, "Do you feel bad about what happened with Emma?" was adapted from an item from the adult State Shame and Guilt Scale, "I feel bad about something I have done" (SSGS; Marschall, Sanftner, & Tangney, 1994). In the child study by Colasante et. al (2014), researchers coded children's responses as "guilty" if the children referenced feeling "guilty" or "bad." The second item, "Do you feel sorry about what

happened with Emma?” was adapted from an item from the child TOSCA-C, “I would feel sorry, very sorry” (TOSCA-C; Tangney, Wagner, Burggraf, Gramzow, & Fletcher, 1990). The third item, “Do you feel happy about what happened with Emma?” was adapted from a reverse-coded item from the adult SSGS, “I feel pleased about something I have done.” The word “happy” was used as children demonstrate understanding of this positive emotion term earlier in development than other positive emotion terms such as proud or excited (Harris et al., 1987). This item was reverse-scored to reflect the degree to which children felt unhappy about what happened with Emma.

Items were presented in question format rather than the declarative format (e.g., “I feel sorry about what happened to Emma”) to reduce confusion children may face when they are required to verify declarative statements (J. W. Chapman & Tunmer, 1995). Finally, although items in some validated scales include vague references to a transgression (e.g., “I feel bad about *something I have done*”), it was determined explicitly referencing the interaction with Emma (e.g., by stating “about what happened with Emma”) would be important such that children’s reports would be specific to the transgression and not other potential sources of state guilt. It was determined that using the language “what happened *with* Emma” would be preferable to “what happened *to* Emma” so that questions would be equally applicable in conditions in which children do and do not transgress.

Thus, the researcher asked the child to report the extent to which he/she felt four different emotions: guilty, bad, sorry, and happy. Similar to the method used in the experimental study of guilt and reparative behavior by Cryder et. al. (2012), a composite score for guilt was calculated by averaging responses to “guilty,” “bad,” “sorry,” and the reverse-scored “happy.” The internal consistency reliability in the current sample was adequate at both time 1 ($\alpha=.77$) and time 2

($\alpha=.79$). Item-total statistics indicated that Chronbach's alpha would be lower if any item in the scale were deleted. Moreover, the internal consistency reliability was above .70 in each one-year age group of children at both time 1 and time 2. An unrotated exploratory principal axis factor analysis provided evidence of only one factor (see the scree plots in Figure 3) at both time 1 and time 2. The sole factor with an eigenvalue greater than 1 accounted for 59.19% of the variance at time 1 and 61.37% at time 2. No items had factor loadings less than .50 at either time 1 or time 2.

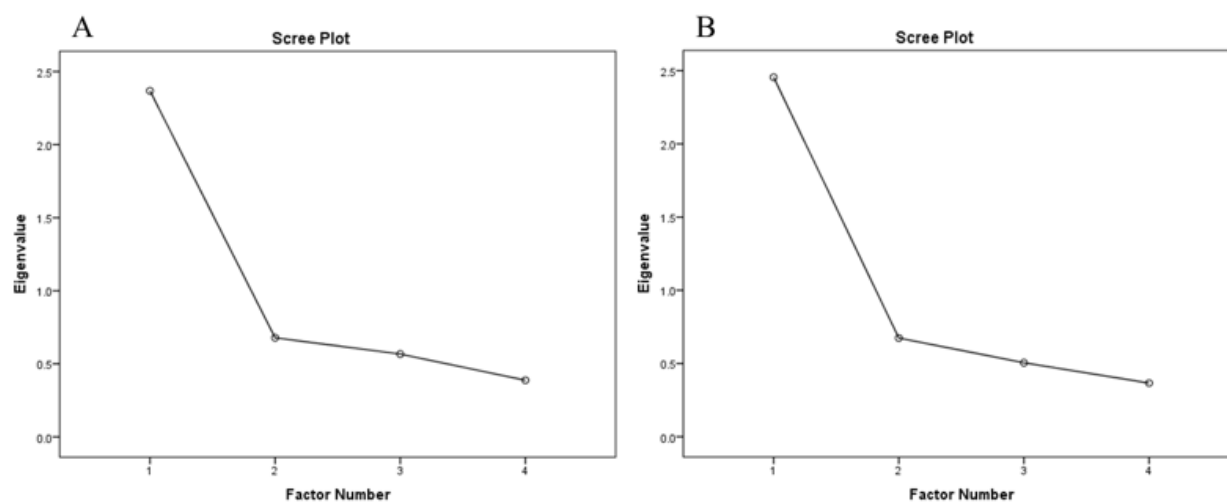


Figure 3. Scree plot for the factor analysis of the guilt scale at time 1 (A) and time 2 (B).

The order in which participants were asked about the four emotions was random across participants. Children rated each emotion on a 5-point likert scale ranging from 1-5. Because 5- to 10-year-old children have been found to have difficulty using likert scales, with tendencies to respond using extreme endpoints (Chambers & Johnston, 2002), a double binary response strategy was implemented that was adapted from that used in the Self Description Questionnaire for Preschoolers (SDQP; Marsh, Ellis, & Craven, 2002), a measure of self-concept in children that has good psychometric properties. The purpose of a double binary system is to obtain a

likert scale outcome by asking children a series of two yes or no questions. Children were asked, “do you feel guilty?” and responded “yes” or “no” to this initial question. A second binary response followed, in which children were asked to answer no “a lot” (1) or no “a little” (2) if they responded “no” initially, or yes “a little” (4) or yes “a lot” (5) if they initially responded “yes.” If a child indicated an understanding of the question but could not decide whether to answer “yes” or “no” to the first binary response, a 3 was recorded because it is midway between the responses of no “a little” (2) and yes “a little” (4).

Pictures were used to aid children’s understanding of the questions. This method had been successfully used by our lab group in a sample of 4-to 6-year-old children. The picture used to aid the first binary response depicted a green check mark to represent “yes” and a red ‘x’ to represent “no” (Figure 4). While the researcher asked the question, “do you feel guilty? Yes or no” she pointed to the green check mark while saying “yes” and the red “x” while saying “no.” Children were able to respond either verbally or by pointing to a picture. The picture used to aid the second binary response depicted either a large and a small red ‘x’ (if the child responded “no” to the first question) or a large and a small green check mark (if the child responded “yes” to the first question; Figure 5). While the researcher asked the question, “no a lot, or no a little?” she pointed to the large “x,” while saying “a lot,” and the small “x” while saying “a little.” The procedure was identical if the child indicated yes to the first question. While the researcher asked the question, “yes a little, or yes a lot?” she pointed to the small check while saying “a little” and the large check while saying “a lot.” The procedure was repeated for bad, sorry, and happy.

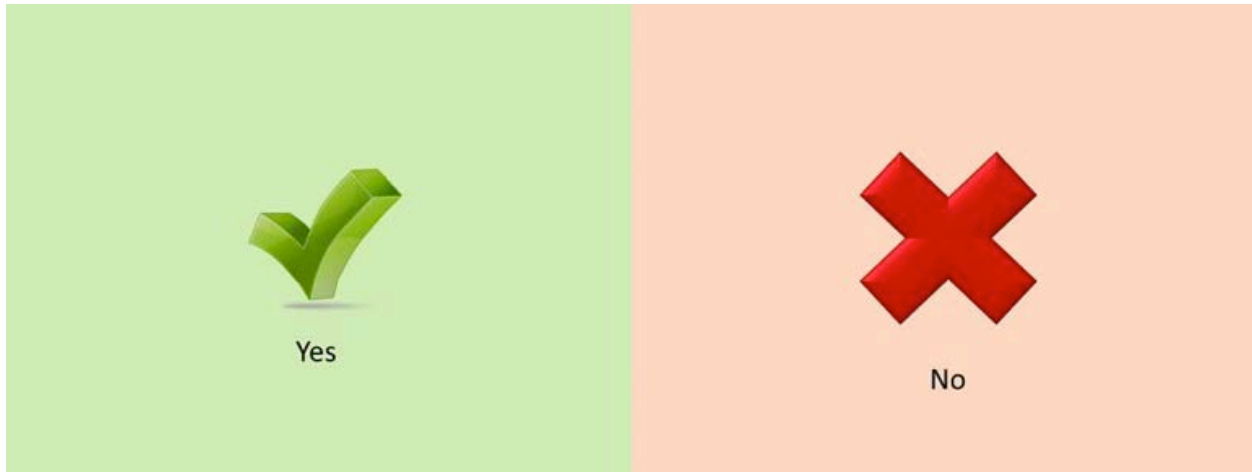


Figure 4. Visual aids for the first binary response.

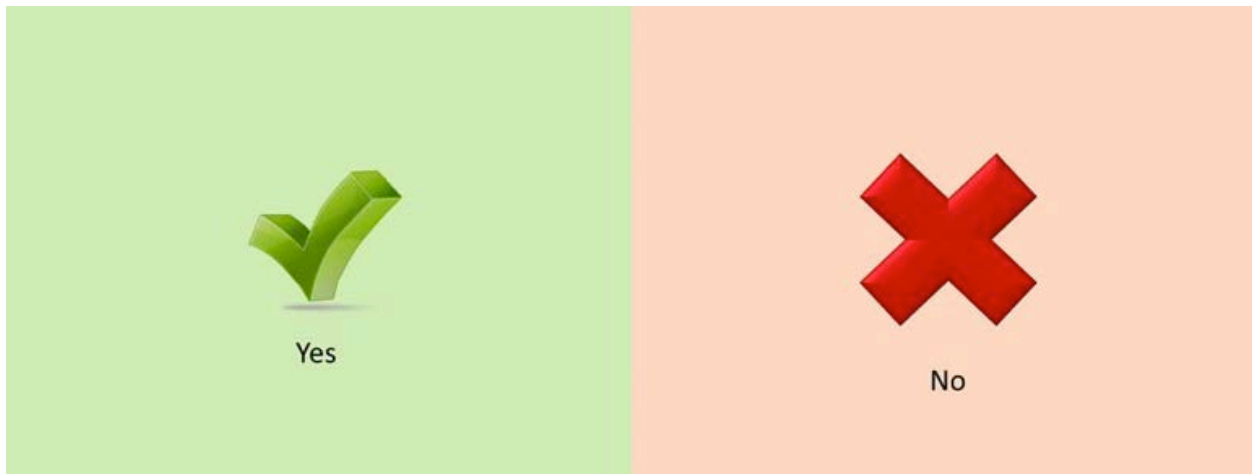


Figure 5. Visual aids for the second binary response (if the child chose “yes” for the first response).

2.4.3 *Repair task.*

As mentioned, existing correlational studies of reparative behaviors in children have either used observational coding of children’s behaviors or parent-report rating scales. The current study modeled one of its measures of reparative behavior from *in vivo* experimental studies of reparative behaviors in adults (Cryder et al., 2012) in that it used a dictator task. Although not used to study reparative behavior, dictator tasks have been widely used with 3-to 9-year-old children in studies of sharing and distributive justice using stickers as currency (Blake

& Rand, 2010; Fehr, Bernhard, & Rockenbach, 2008; Gummerum, Hanoch, Keller, Parsons, & Hummel, 2010; Smith, Blake, & Harris, 2013). The current study is the first to use a dictator task in a transgressor context as a measure of reparative behavior.

The current study's dictator task closely mirrored those used in previous studies in children. At the beginning of the study, children were first given five scratch and sniff stickers and were told that the stickers belonged to them. Establishing ownership is important in dictator tasks with young children, as believing the stickers belong to the researcher would alter the task (Smith et al., 2013). Moreover, previous studies have demonstrated that children's self-report of how much they value scratch and sniff stickers is high and does not change with age (Smith et al., 2013). Finally, although various numbers of currency and/or stickers have been used in previous research with dictator tasks (e.g., four, five and ten), five stickers were chosen to mirror the study of reparative behavior in adults by Cryder et al. (2012).

Later, during the repair phase, children were told that they could keep all of the stickers, or choose to share any number (1-5) with Emma, the pretend child. One, colored envelope with the participating child's name was placed by the child's side, and a second, white envelope marked "Emma" was placed by the researcher's side. To minimize socially desirable responses, the researcher explained that she would cover her eyes with her hands while the child placed the stickers in the envelopes, and that she would not know what the child chose to do with her stickers (Benenson, Pascoe, & Radmore, 2007). The researcher asked the child if she understood the instructions and repeated any as necessary (Benenson et al., 2007). After approximately a minute, she asked the child if she was finished. Once the child indicated that she was finished dividing the stickers, the researcher said, "Let's write a note to Emma to put in the envelope. What do you want to say to Emma?" The researcher transcribed the child's note on an index

card and placed it in the envelope. Finally, another researcher took the envelope marked “Emma” and pretended to deliver it to her. Thus, this measure of reparative behavior is operationalized as the number of stickers given to Emma on a 6-point likert scale from 0 to 5 stickers given. The note to Emma was coded and used as a second reparative behavior outcome measure (see 2.6).

2.4.4 Sorting task.

Instead of the repair task, children in the no-repair condition completed the sorting task in which children simply sorted stickers between two envelopes. The task mirrored the repair task closely. Early in the study, children were given five scratch and sniff stickers. During the repair phase, one colored and one white envelope was placed in front of the child. The child was told that she was to divide the stickers into the two envelopes, and that she could divide them any way she wanted. The researcher explained that she would cover her eyes with her hands while the child placed the stickers into the envelopes, and that she would not know what the child chose to do with the stickers. The researcher asked the child if she understood the instructions and repeated any as necessary. After approximately a minute, she asked the child if she was finished. Once the child indicated that she was finished dividing the stickers, she asked the child which envelope she wanted to take home. The researcher said, “Will you write a note to the next kid who comes to this table to play games? What do you want the note to say?” The researcher transcribed the child’s note on an index card and set it aside. The researcher then said that she wished there was a way for the child to share the remaining envelope with Emma, but a researcher was not available to take the envelope to her. This created a situation in which the child’s reparative behavior was blocked.

No data was collected during this task. A primary purpose of the task was to give children in the no-repair condition a task equivalent to the repair task. It was also used in order to ensure that there was an equivalent amount of time between the two guilt ratings for children in all experimental conditions; this was particularly important to examine whether reparative behaviors lead to alleviated guilt above and beyond guilt that is reduced due to simple passing of time.

2.5 Materials

2.5.1 Gift options.

Children chose between the following two gifts (Figure 6). The gifts were identical types of toys (bouncy balls), but one was brightly colored and attractive to children, and the other was ostensibly less attractive.

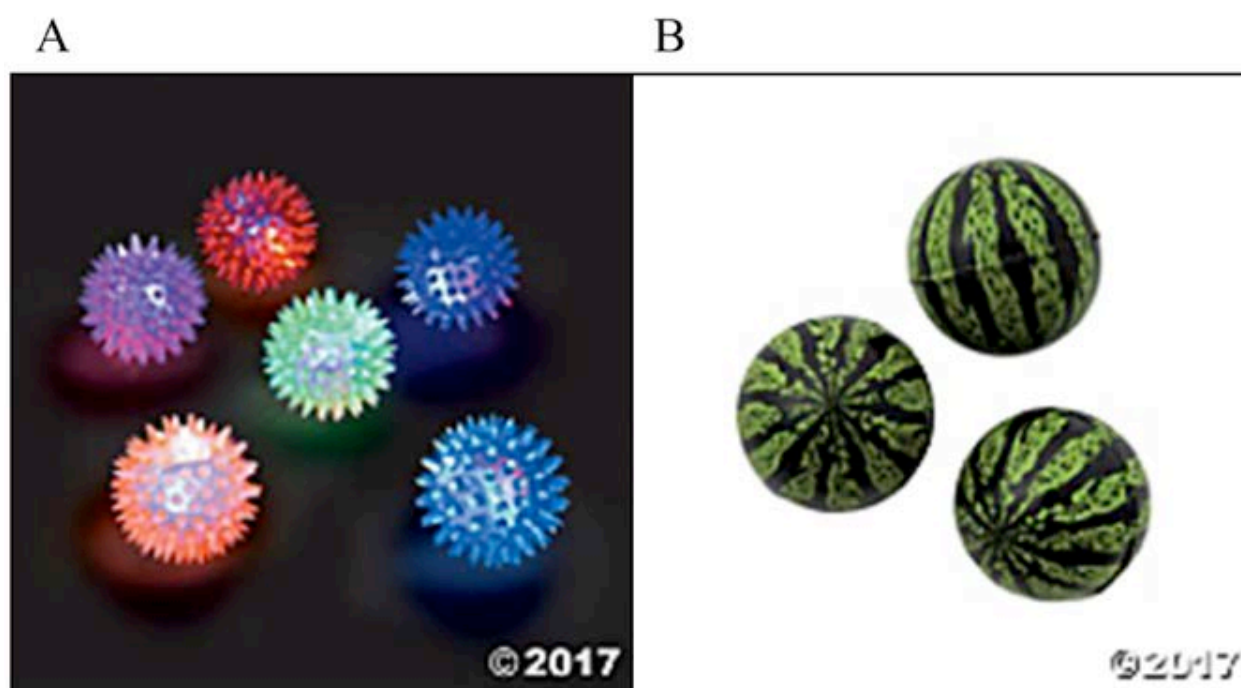


Figure 6. Still photos of the gift options: A) attractive toy and B) unattractive toy.

2.5.1.1 Attractive toy (A).

A colorful bouncy ball (2"). This ball lights up when bounced.

2.5.1.2 Unattractive toy (B).

A black and green bouncy ball (1"). This ball does not light up or have any special features.

2.5.2 Feedback images.

Two images (presented on an iPhone screen, 5.5 inches from corner to corner), each depicting a pretend text message exchange with another researcher (Figure 7). Both feedback images contained a text message from a pretend researcher and a photo of a female child either expressing distressed or slightly positive emotions.⁶

⁶ It was expected that the majority of children would choose the attractive toy. Indeed, in piloting the paradigm, 30 of the 33 pilot participants chose the attractive toy. However, the paradigm was still able to be completed if the participating child chose the unattractive toy; he/she was simply told that the pretend child really wanted the unattractive toy, too.

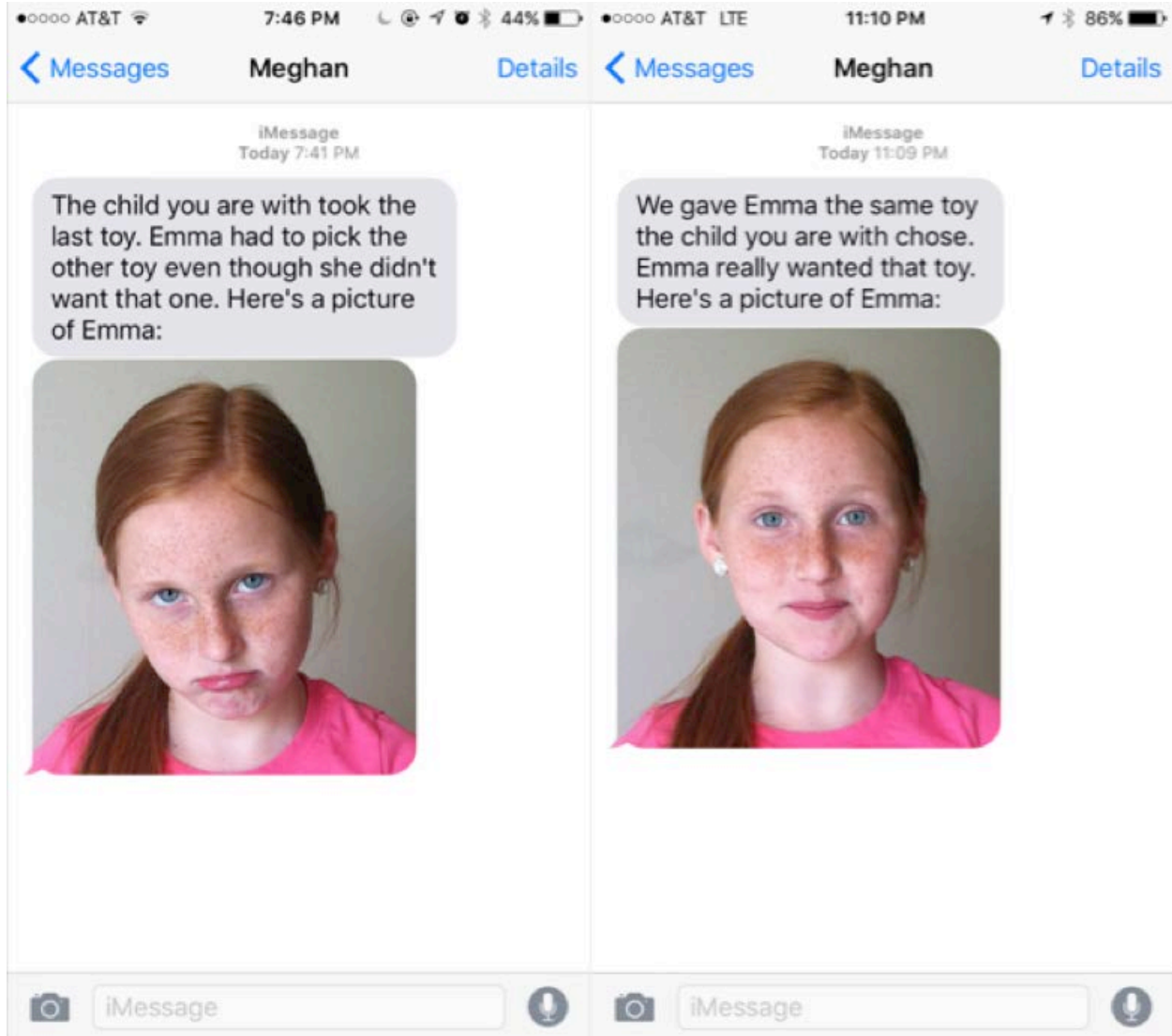


Figure 7. Still photos of the distress and slightly satisfied feedback images.

2.5.2.1 Distress feedback image.

The distress feedback image contained a picture of Emma's face, looking sad and disappointed.

2.5.2.2 Satisfied feedback image.

The satisfied feedback image contained a picture of Emma's face, looking slightly happy. It was determined that the emotion conveyed in this photo would be a slightly happy face rather

than a neutral face, as some research suggests that children with higher levels of social anxiety and depressive symptoms tend to interpret neutral faces as negative (e.g., Cooney, Atlas, Joormann, Eugene, & Gotlib, 2006), and it is possible that the community sample will include children with such symptoms.

2.5.3 Stickers.

Two kinds of colorful, scratch and sniff stickers were used in the repair and sorting tasks.

2.6 Coding and Rating of Data.

Research assistants (RAs) who were blind to the study hypothesis coded or rated four variables. Inter-rater reliability statistics are reported in Table 3.

2.6.1 Children's free-response report of their emotion.

In all conditions, the emotion word that each child free-responded following the induction phase (distress/satisfied feedback) was coded. The word was categorized into one of four groups: negative emotion words (i.e., sad, bad, not good/not happy, sorry), positive emotion words (i.e., good, happy/glad), neutral emotion words (e.g., okay), and "I don't know" responses. Each response could be coded as only one category. No child reported more than one emotion word. Responses that were off-topic or did not contain an emotion word were coded as a non-response (missing data). Inter-rater reliability, assessed with a Cohen's Kappa coefficient, was excellent (see Table 2).

Table 2. *Reliabilities for Scored or Rated Variables*

| Scored/Rated Variable | Training | Final |
|------------------------------------|----------|-------|
| Free-response report of emotions | 1.00 | 1.00 |
| Note ratings | .97 | 1.00 |
| Apology | 1.00 | 1.00 |
| <u>Reasons for giving stickers</u> | | |
| Mood Theme | .88 | .97 |
| Event Theme | .75 | .86 |
| Fault Theme | .93 | .93 |
| Desire Theme | .90 | .93 |
| Moral Obligation Theme | .93 | .93 |
| Other Theme | 1.00 | .90 |
| Don't Know | 1.00 | 1.00 |

Note. Final reliability includes all study items excluding training items.

2.6.2 Note ratings.

In the experimental and no-guilt conditions, the prosocial content of participants' notes to Emma was rated as a second measure of reparative behavior. The presence and sophistication of children's prosocial statements (e.g., comforting statements, statements of concern, apologies) were rated on a 4-point scale (0=*absent*, 3=*sophisticated*; see Appendix A). Inter-rater reliability, assessed with a two-way, mixed, absolute agreement intraclass correlation coefficient (ICC), was excellent. This statistic is identical to using a weighted kappa with quadratic weights for ordinal scales, and the two can be used interchangeably (Norman & Streiner, 2008).

2.6.3 *Children's apology.*

In the experimental and no-guilt conditions, participants' notes were coded for the presence (1) or absence (0) of an apology (i.e., a statement of sorrow for wrongdoing; in the current sample, all apologies included the word "sorry"). The presence of apology suggests children's perceptions of fault or responsibility for the transgression. Inter-rater reliability, assessed with a Cohen's Kappa coefficient, was excellent.

2.6.4 *Children's report of reasons for giving stickers.*

In the experimental and no-guilt conditions, children's report of reasons for giving stickers were transcribed. The transcript was coded for the presence of seven possible data-driven themes (see Table 3). More than one theme category could be present in a transcript (e.g., all categories that applied were coded). Mood, fault, and event themes were together considered themes that reflected prosocial intent or motivation. Thus, the Prosocial Themes variable is a binary variable reflecting the presence (1) or absence (0) of any mood, fault, or event theme in children's notes. Children's reasons were rated as a manipulation check in order to examine whether giving of stickers by children in the experimental condition was prosocially motivated (i.e., to confirm that this variable reflected reparative prosocial behavior). Inter-rater reliabilities, assessed with Prevalence-Adjusted Bias-Adjusted Kappa coefficients (PABAK; Byrt, Bishop, & Carlin, 1993), were excellent. This statistic is used to adjust for prevalence when one category is observed at a higher rate than other categories.

Table 3. *Categories of themes included in children's report of reasons for giving stickers*

| Reason Theme | Description | Example(s) |
|-------------------------|-------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------|
| <i>Prosocial Themes</i> | | |
| Mood Theme | References wanting to improve the other child's mood | "Because she's sad and I want to make her feel happy" "Because it will make her feel happier" |
| Fault Theme | Child acknowledges fault/responsibility | "Because I took the last ball" "Because I made her feel bad and I don't like hurting people's feelings" |
| Event Theme | Child references the event that happened to the other child without directly acknowledging responsibility for the event | "Because of what happened to Emma" "Cause when some person's sad I want to make it up to them" |
| <i>Other Themes</i> | | |
| Desire Theme | Child reports that they wanted to give stickers | "Because I wanted to" "I felt like it" |
| Moral Obligation Theme | Child references a moral norm | "Because it's nice to share" "Because it's fair" |
| Other Theme | Less common responses that could not be categorized into another theme | "Because most white kids like gummy bears" |
| Don't Know | Child states they do not know their reason for giving | "I don't know" |

2.6.5 Training and reliability of coders/raters

One trained researcher rated/coded all four variables. The researcher served as the master rater and was blind to all study hypotheses, child's demographics, and child's experimental condition. The researcher trained by reading the manual and meeting with the primary investigator (PI) to discuss ratings and examples and ask questions. The researcher first rated training items with the PI, who reviewed the rating. The researcher then rated a training item (for each of the four variables) on her own and met with the PI to review all ratings and reconcile any discrepancies. Next, the researcher was given one new training item at a time to rate independently. She met with the PI to discuss the ratings. After five items per variable were rated independently, training inter-rater reliability was assessed. Rating of items continued until

reliability of at least .75, considered in the excellent range (Cicchetti, 1994; Fleiss, 1981), was attained. The researcher and PI then independently rated all items for each variable (i.e., the design was fully crossed). Discrepancies between raters were resolved by retaining the rating of the master rater.

2.7 Manipulation checks and predictions.

Four manipulation checks assessed whether the distress feedback successfully induced guilt. Two manipulation checks assessed whether the distress feedback elicited the emotional component of guilt. First, condition differences in children's guilt scale time 1 scores were compared. It was expected that children in a condition that involved the distress feedback (experimental and no-repair conditions) would report greater guilt scale time 1 scores than children in the no-guilt condition. Second, condition differences in children's free-response report of their emotions was compared. It was expected that a greater proportion of children in a condition that involved the distress feedback would report a negative emotion than children in the no-guilt condition. Two manipulation checks assessed whether the distress feedback elicited the cognitive component of guilt. First, it was expected that the majority of children in a condition that involved the distress feedback would report fault on the children's fault question. Second, it was expected that the majority of notes of children in the experimental condition would contain an apology to Emma.

A final manipulation check assessed whether the number of stickers given was a valid measure of reparative prosocial behavior. Condition differences in children's report of reasons for giving stickers were examined. It was expected that a greater proportion of children in the experimental condition would report a prosocial theme than children in the no-guilt condition.

2.8 Piloting participants.

I piloted the study on a group of 5-to 10-year-old children ($N = 33$) to test aspects of the design. First, I examined the believability and effectiveness of the distress feedback in inducing guilt in children. The findings supported the prediction that children would believe that they had transgressed against a real child. No children laughed or stated that they thought the feedback was pretend; additionally, several children spontaneously discussed their emotions about what happened to the pretend peer, suggesting that the paradigm was believable to children. Children also reported greater guilt in response to guilt rating 1 in the conditions that involved a transgression (experimental and no-repair conditions; $M = 4.2$, $SD = .47$) than the no-guilt condition ($M = 2.3$, $SD = .84$); $t(15) = -5.85$, $p = .00$, suggesting that the paradigm successfully induced guilt.

Second, I examined the apparent comprehensibility of the guilt scale questions, for example whether children seemed confused by these questions (e.g., indicating that they didn't know what the word guilty means or many responses of, "I don't know"). Five-year-old children appeared visibly confused by the questions, and some children either gave identical responses to all of the three emotions (e.g., stating that they felt both very guilty and very happy), or asked me what "guilty" meant. Thus, I determined that five-year-olds were too young to understand the paradigm. Moreover, the visual aids appeared effective in maintaining children's attention during the ratings. Finally, I determined that four questions per guilt rating phase was the maximum that could be asked during the paradigm and maintain children's attention.

Regarding hypothesis 1, pilot children in the experimental condition engaged in more reparative behavior ($M = 2.4$, $SD = 1.14$) than children in the no-guilt condition ($M = 2.0$, $SD = 1.15$). Regarding hypothesis 2, although children had similar levels of guilt at the first rating in

experimental ($M = 4.00$, $SD = .45$) and no-repair conditions ($M = 4.40$, $SD = .42$), children in the experimental condition had lower levels of guilt at the second guilt rating ($M = 2.4$, $SD = 1.30$) than children in the no-repair condition ($M = 4.13$, $SD = 1.58$).

3 RESULTS

3.1 Descriptive Statistics

Means, standard deviations, and ranges for study variables are presented in Table 4.

Correlations among continuous variables are displayed in Table 5 and are discussed below.

Descriptive information about categorical variables is also presented in the text below.

Table 4. *Descriptive Statistics of Primary Study Variables: Means, Standard Deviations, and Ranges or Percentages*

| Variable | Observed Range | Mean (SD) or Frequency | | | |
|--------------------------------------|----------------|------------------------|--------------------------|--------------------------|--------------------------|
| | | Overall Sample N=97 | Experimental n=32 | No-repair n=32 | No-guilt n=33 |
| 1. Age (years) | 6-10 | 7.94 (1.37) | 7.88 (1.24) ^a | 8.13 (1.24) ^a | 7.82 (1.61) ^a |
| 2. Sex (% female) | | 59.8% | 62.5% ^a | 62.5% ^a | 54.5% ^a |
| 3. Race (% White) | | 59.8% | 56.3% ^a | 64.5% ^a | 60.6% ^a |
| 4. Children's fault question (% yes) | | 48.5% | 56.3% ^a | 51.9% ^a | 46.9% ^a |
| 5. Apology (% yes) | | 30.2% | 58.1% ^a | N/A | 3.1% ^b |
| 6. Guilt scale time 1 | 1-5 | 3.43 (1.18) | 4.00 (.82) ^a | 4.10 (.62) ^a | 2.22 (.93) ^b |
| 7. Guilt scale time 2 | 1-5 | 3.09 (1.27) | 3.26 (1.14) ^a | 4.08 (.67) ^b | 1.98 (.91) ^c |
| 8. Number of Stickers Given | 0-5 | 2.32 (1.13) | 2.66 (.90) ^a | N/A | 2.00 (1.25) ^b |
| 9. Note Rating | 0-3 | 1.54 (.96) | 2.06 (.93) ^a | N/A | 1.58 (.90) ^b |

Note. The observed ranges are identical to the possible ranges for each variable. Means/percentages in each row with differing superscripts are significantly different from each other ($p < .05$).

Table 5. *Correlations among Continuous Variables*

| Variable | 1 | 2 | 3 | 4 | 5 |
|-----------------------------|-------|------|------|-----|---|
| 1. Age | | | | | |
| 2. Guilt scale time 1 | -.05 | | | | |
| 3. Guilt scale time 2 | -.10 | .21 | | | |
| 4. Number of Stickers Given | .47** | -.24 | -.21 | | |
| 5. Note Rating | .43* | .28 | .30 | .30 | |

3.1.1 Correlations among variables.

Children's older age was significantly associated with a greater number of stickers given to the pretend peer. Children's older age was also significantly associated with higher note ratings (i.e., the notes contained more sophisticated prosocial content). The number of stickers given and the note ratings were moderately and positively but nonsignificantly correlated, suggesting that they are related but distinct measures of reparative prosocial behavior.

3.1.2 Group differences among variables.

Tests of differences in outcome variables by demographic variables were conducted in order to determine whether covariates should be included in tests of the study hypothesis. Independent samples t-tests revealed no significant differences in outcome variables (i.e., number of stickers given, or note ratings, guilt scale time 2 scores) by gender or race. A series of t-tests were conducted within each of the three conditions and also revealed that the outcome variables did not differ significantly by gender or race within any conditions, and all effect sizes were small ($d < .20$). Random assignment was expected to produce an equal distribution of demographic variables across experimental conditions. Statistical tests were conducted to confirm this assumption. Chi-square tests revealed that the proportion of boys and girls and the proportion of children from each racial group were not significantly different across experimental conditions. A one-way ANOVA revealed no significant difference in age by condition.

As a result of the above analyses, age was included as a covariate in the analyses examining the effect of condition on the number of stickers given and note ratings (Hypothesis 1).

3.2 Manipulation Checks

Four manipulation checks were conducted to examine whether a) the distress feedback successfully induced guilt (four manipulation checks) and b) the number of stickers given did indeed reflect reparative prosocial behavior (one manipulation check).

3.2.1 *Condition differences in children's guilt scale time 1 scores.*

First, condition differences in children's guilt scale time 1 scores were examined to ensure that children in a condition that involved the distress feedback (experimental and no-repair conditions) felt guilty following the transgression. Guilt scale time 1 scores were significantly different among conditions, $F(2,94) = 56.54, p < .001$, partial $\eta^2 = .55$. Post-hoc analysis using the Bonferroni correction revealed that there was no significant difference in children's rated guilt between the two conditions that utilized the distress feedback ($M_{diff} = -.10, SE = .20, p = 1.00$). However, children in both the experimental ($M_{diff} = -1.78, SE = .20, p < .001$) and no-repair ($M_{diff} = -1.88, SE = .20, p < .001$) conditions reported greater guilt in response to guilt rating 1 than children in the no-guilt condition. This result suggests that the distress feedback successfully induced guilt and thus supports the validity of the paradigm.

3.2.2 *Children's report of their perceived fault.*

3.2.2.1 *Children's fault question.*

Second, children's responses to the children's fault question asking "is Emma feeling [happy/sad] because of you?" were examined to ensure that the majority of children in a condition involving the distress feedback (experimental, no-repair conditions) would indicate that they were at fault for causing Emma's distress. Eight of the 97 child participants did not respond to the question after repeated prompting, and were thus entered as missing data. Of the children who responded to the question in a condition involving the distress feedback, 56% ($n =$

32), responded “yes,” indicating that they were responsible for causing Emma’s emotions. Of children who responded to the question in the no-guilt condition, 47% ($n = 15$) indicated that they responsible for causing Emma’s emotions. As expected, given that in all conditions children were told that their selection of a toy resulted in Emma receiving a particular toy and thus her emotions, a Chi-square test revealed no significant effect of condition on children’s report of their perceived fault, $\chi^2(1) = 71, p = .40$. Thus, children in a distress feedback condition did not report significantly more fault than children in the no-guilt condition. Consistent with expectations, the majority of children in a condition in which they were led to believe they transgressed reported that they were at fault for causing Emma’s distress.

3.2.2.2 Children’s apology.

Third, the presence of an apology in children’s notes was examined to ensure that the majority of children in the experimental condition apologized to Emma. As expected, a slight majority of children’s notes in the experimental condition contained an apology ($n = 18$; 58%); in contrast, only one child in the no-guilt condition apologized (3.1%). A Chi-square analysis revealed that significantly more children in the experimental than the no-guilt condition apologized in their notes, $\chi^2(1) = 22.56, p < .001$. The children’s fault question was not related to children’s apology; the proportion of children who apologized was not significantly different depending on whether children had previously reported fault on the fault question, $\chi^2(1) = .08, p = .79$.

Of the 32 children in a distress feedback condition who have both measures of fault (children in the experimental condition), 84.4% of children ($n = 27$) reported fault on at least one measure. Children who admitted fault were more likely to report fault on one measure (66.67%) than both measures (33.33%). Taken together, the data indicates that children in the

experimental condition were highly likely to report fault, but tended to do so on one, but not both fault measures.

3.2.3 Children's free-response report of their emotions.

Fourth, children's free-responses to the question that asked "How do you feel about what happened with Emma?" were examined to ensure that children in a condition involving the distress feedback experienced negative affect following the transgression. In the no-guilt condition, 1 child had missing data (the child did not respond to the question). Of children in this condition with data, 81.5% named a positive emotion word, 11.1% named a negative emotion word, and 7.4% named a neutral emotion word. Table 6 provides a breakdown of the specific words used by children in each emotional category. In a condition that contained the distress feedback (experimental and no-repair conditions), 4 children had missing data (2 did not respond, and 2 did not respond with an emotion word). Of children in these conditions with data, 87.0% named a negative emotion word, 11.1% said they did not know, and 1.9% named a neutral emotion word. Notably, no child named a positive emotion word.

Table 6. *Emotion Words Contained in Children's Free Responses by Experimental Condition*

| Emotion Word Variables | Percent Reported | |
|-------------------------------|------------------------------------|-----------------------------------------------------------------------|
| | No-guilt condition <i>n</i> =32 | Distress feedback condition (Experimental, No-repair) <i>n</i> =60 |
| <i>Negative Emotion Words</i> | 11.1% | 87.0% |
| 1. Sad | 7.4% | 62.2% |
| 2. Bad | 0% | 15.1% |
| 3. Not good/Not happy | 3.7% | 5.7% |
| 4. Sorry | 0% | 3.8% |
| <i>Positive Emotion Words</i> | 81.5% | 0% |
| 5. Good | 44.4% | 0% |
| 6. Happy/Glad | 37.0% | 0% |
| <i>Neutral Emotion Words</i> | 7.4% | 1.9% |
| 7. Okay | 7.4% | 0% |
| 8. Other neutral word | 0% | 1.9% |
| <i>Other responses</i> | 0% | 11.1% |
| 8. Don't know | 0% | 11.1% |

As expected, a chi square analysis revealed that a significantly greater proportion of children in a condition involving the distress feedback reported experiencing a negative emotion than children in the no-guilt condition, $\chi^2(1) = 45.92, p < .001$, supporting the validity of the paradigm for eliciting the negative emotion component of guilt. Of children in a condition involving the distress feedback, 24.6% spontaneously generated an emotion word that was specifically subsequently probed through items in the guilt rating scale created for this study (i.e., the children said they felt bad, sorry, or unhappy⁷). The majority of these children (62.2%) reported feeling sad following the transgression, an emotion that was not specifically probed in the guilt rating scale.

⁷ Recall that the item probing for children's happy emotions is reverse-scored to reflect the degree that children felt unhappy about what happened with Emma.

3.2.4 *Children's report of reasons for giving stickers.*

Finally, children's free-responses to the question that asked "Why did you decide to give stickers to Emma?" were examined to ensure that children's giving of stickers was prosocially motivated and is therefore a valid measure of reparative behavior⁸. Table 7 provides a breakdown by experimental condition of the percent of each theme reported in the transcripts of children's reasons for giving stickers. The most common theme for children in the no-guilt condition was that they had a moral obligation to share with Emma (e.g., "Because it's good to share," "Because it's fair"), which was endorsed by 31% of children. In contrast, the most common theme for children in the experimental condition was that they wanted to improve Emma's mood (e.g., "Because she's sad and I want to help her feel happy"), which was endorsed by 34.5% of children. A series of Chi-square tests were conducted to examine condition differences in the themes endorsed. Significantly more children in the no-guilt than experimental condition reported that they gave stickers because of a moral obligation, $\chi^2(1) = 5.50, p = .02$. More children in the experimental than no-guilt condition reported that they gave stickers because of something that happened to Emma (event theme), $\chi^2(1) = 3.63, p = .06$, an effect that was marginally significant.

The majority of children in the experimental condition endorsed a prosocial theme for giving (75.8%). In contrast, 37.9% of children in the no-guilt condition endorsed a prosocial theme. A significantly greater proportion of children in the experimental than no-guilt condition endorsed a prosocial theme, $\chi^2(1) = 6.11, p = .01$. Taken together, these results suggest that giving of stickers by children in the experimental condition reflected reparative prosocial behavior.

⁸ Recall that only children in the no-guilt and experimental conditions were asked this question, as children in the no-repair condition were not given an opportunity to give stickers to Emma until after the paradigm was completed.

Table 7. *Themes Endorsed in Children’s Report of Reasons for Giving Stickers by Experimental Condition*

| Theme Variable | Percent Endorsed | |
|---------------------------|------------------------------------|----------------------------------------|
| | No-guilt condition <i>n</i> =29 | Experimental condition <i>n</i> =29 |
| <i>Prosocial Themes</i> | 37.9% ^a | 75.8% ^b |
| 1. Mood Theme | 27.6% ^a | 34.5% ^a |
| 2. Fault Theme | 10.3% ^a | 17.2% ^a |
| 3. Event Theme | 3.4% ^a | 24.1% ^b |
| <i>Other Themes</i> | | |
| 4. Desire Theme | 13.8% ^a | 13.8% ^a |
| 5. Moral Obligation Theme | 31.0% ^a | 6.9% ^b |
| 6. Other Theme | 10.3% ^a | 10.3% ^a |
| 5. Don’t Know | 6.9% ^a | 0% ^a |

Note. Percentages in each row with differing superscripts are significantly or marginally significantly different from each other ($p < .07$).

3.3 Tests of Study Hypotheses.

3.3.1 *Guilt motivates children’s greater reparative behaviors.*

Tests of hypothesis 1 involved examining the effect of condition (experimental and no-guilt) on two distinct measures of reparative prosocial behavior—number of stickers given and note ratings (of prosocial content). The purpose was to investigate whether children in the experimental condition in which guilt was induced engaged in more reparative behavior than children in the no-guilt condition in which guilt was not induced.

3.3.1.1 *Number of stickers given.*

An Analysis of Covariance (ANCOVA) that included age as a covariate was conducted to examine the effect of condition on the number of stickers given. The number of stickers given was a 6-point continuous variable and was approximately normally distributed, as assessed by visual inspection of a histogram. Assumptions of ANCOVA were tested (Huitema, 2011). Standardized residuals of the number of stickers given were normally distributed for the

experimental condition, $W(32) = .97, p = .42$, and the no-guilt condition, $W(33) = .95, p = .14$, as assessed by the Shapiro-Wilk's test of normality. Standardized residuals for the overall model were also normally distributed, as assessed by Shapiro-Wilk's test of normality, $W(65) = .97, p = .10$. There was homogeneity of regression slopes as the interaction of condition and age was not statistically significant, $F(1, 62) = 1.16, p = .29$. There was homoscedasticity, as assessed by visual inspection of the standardized residuals plotted against the predicted values. There was homogeneity of variances, as assessed by Levene's test of homogeneity of variance, $F(1,64) = 1.42, p = .24$. There were no outliers in any cells of the design, as assessed by no cases with standardized residuals greater than ± 3 standard deviations. There was a linear relation between age and number of stickers given for both the experimental and no-guilt conditions, as assessed by visual inspection of a scatterplot.

When controlling for age, there was a significant effect of condition on the number of stickers given, $F(1,62) = 6.41, p = .01, \eta_p^2 = .09$ (see Figure 8A). Children in the experimental condition for whom guilt was induced gave significantly more stickers than children in the no-guilt condition for whom guilt was not induced ($M_{diff} = .64, SE = .25$). The size of the effect was medium in magnitude, $d = .59$.

3.3.1.2 Note ratings.

A second ANCOVA that included age as a covariate was conducted to examine the effect of condition on the note ratings. The number of stickers given was a 4-point continuous variable and was approximately normally distributed, as assessed by visual inspection of a histogram. Assumptions of ANCOVA were tested. Standardized residuals of the note ratings were normally distributed for the experimental condition, $W(31) = .16, p = .08$, and the no-guilt condition, $W(33) = .12, p = .10$, as assessed by the Shapiro-Wilk's test of normality.

Standardized residuals for the overall model were also normally distributed, as assessed by the Shapiro-Wilk's test of normality, $W(64) = .11, p = .19$. There was homogeneity of regression slopes as the interaction of condition and age was not statistically significant, $F(1, 60) = .41, p = .53$. There was homoscedasticity, as assessed by visual inspection of the standardized residuals plotted against the predicted values. There was homogeneity of variances, as assessed by Levene's test of homogeneity of variance, $F(1,62) = .02, p = .89$. There were no outliers in any cells of the design, as assessed by no cases with standardized residuals greater than ± 3 standard deviations. Regarding the assumption of linearity, there was a linear relation between age and note ratings for both the experimental and no-guilt conditions, as assessed by visual inspection of a scatterplot.

When controlling for age, there was a significant effect of condition on the note ratings, $F(1,61) = 5.51, p = .02, \eta_p^2 = .08$ (see Figure 8B). The notes of children in the experimental condition were rated to be significantly more prosocial than the notes of children in the no-guilt condition ($M_{diff} = .49, SE = .21$). The size of the effect was medium in magnitude, $d = .54$.

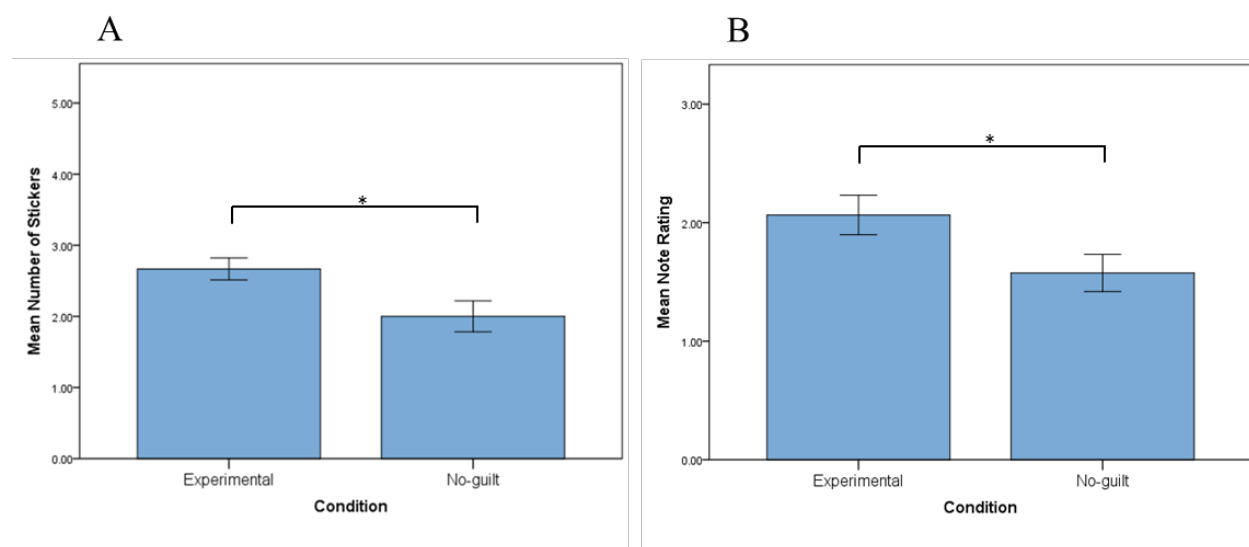


Figure 8. Mean ($\pm SE$) number of stickers (A) and mean ($\pm SE$) note ratings of prosocial themes (B) as a function of experimental condition. * = $p < .05$.

3.3.2 *Children's reparative behaviors alleviate guilt.*

To examine hypothesis 2, a mixed 2x2 factorial ANOVA with time as a 2-level within-subjects factor (two timepoints), condition as a 2-level between-subjects factor (experimental and no-repair) and guilt scores as the dependent variable was used to test the interaction of time and condition on guilt scores. The purpose of this analysis was to investigate whether, among children for whom guilt was induced, children who were given a repair opportunity reported greater reductions in guilt over time compared to children who did not have a repair opportunity.

Assumptions of mixed ANOVA were tested. Standardized residuals were normally distributed for each combination of the levels of the between- and within-subjects factors, as assessed by Shapiro-Wilk's test of normality. Specifically, guilt rating 1 scores were normally distributed for the experimental condition, $W(29) = .91, p = .10$, and the no-repair condition, $W(31) = .96, p = .21$; guilt rating 2 scores were normally distributed for the experimental condition, $W(29) = .96, p = .21$, and the no-repair condition, $W(31) = .94, p = .07$, as assessed by the Shapiro-Wilk's test of normality. There were four outliers in cells of the design exerting undue influence on the analysis (3 in the experimental condition, 1 in the no-repair condition), as assessed by examination of studentized residuals for values greater than ± 3 . The outliers were removed from the analysis⁹. There was homogeneity of variances for guilt rating 1 scores, $F(1,59) = .004, p = .95$, but not guilt rating 2 scores, $F(1,59) = 12.88, p = .001$, as assessed by Levene's test of homogeneity of variance. There was not homogeneity of covariances, as assessed by Box's test of equality of covariance matrices ($p < .001$). Square root and logarithmic transformations did not result in data that met the assumptions of homogeneity of variance and covariance. Mauchly's test of sphericity indicated that the assumption of sphericity was violated for the two-way interaction, $\chi^2(0) = 1.00, p < .001$. As assumptions of mixed ANOVA were

⁹ The result of the analysis of hypothesis 2 was the same when outliers were included.

violated, linear mixed modeling, a statistical alternative to mixed ANOVA, was used (West, Galecki, & Welch, 2014). Linear mixed models require normally distributed residuals for each combination of the levels of the between- and within-subjects factors and are sensitive to outliers, but are useful in data with heterogeneity of variance and/or missing datapoints (West et al., 2014).

As is recommended for linear mixed models, the step-up strategy to model building was used in which the simplest model is fitted first and more complex models (i.e., models with a greater number of parameters) are retained only if they significantly improve model fit (Raudenbush & Bryk, 2002; Twisk, 2006). Model fit was measured using the -2 log-likelihood criteria; significantly smaller values, relative to the degrees of freedom, indicate better fit of that model. This statistic follows a Chi-square distribution. The number of degrees of freedom is obtained by subtracting the number of parameters in the less complex model from the more complex model that is being compared (West et al., 2014). A diagonal covariance structure was specified; this variance structure is used for repeated measures data in which variances are assumed to be heterogeneous (Field, 2009). Model 1 compared the effect of condition as a fixed effect on guilt scores. The model had 3 parameters and a -2 log-likelihood of 307.63. Model 2 added the effect of time as a fixed effect. The model had 5 parameters and significantly improved model fit, $\chi^2_{\text{diff}}(2) = 17.35, p < .05$. Model 3 added an interaction between condition x time. The model had 6 parameters and significantly improved model fit, $\chi^2_{\text{diff}}(1) = 7.94, p < .05$. Model 4 added a random intercept. The model had 7 parameters and significantly improved model fit, $\chi^2_{\text{diff}}(1) = 10.31, p < .05$. Model 5 added a random slope. The model had 9 parameters and did not significantly improve model fit $\chi^2_{\text{diff}}(2) = .032, p > .05$. Thus, the best fitting model

included condition and time as fixed effects, a condition x time interaction term, and a random intercept.

There was a statistically significant interaction between condition and time on guilt scores, $F(1, 61) = 14.464, p < .001$ (see Figure 9). The nature of the interaction was probed using additional linear mixed models. Two separate linear mixed models were conducted to test for differences in guilt between timepoints for each condition. The models specified were the same as the main model but excluded the main effect term for condition as well as the interaction term. Consistent with hypothesis 2, there was a statistically significant effect of time on guilt for the experimental condition, $F(1, 29) = 15.88, p < .001$, but not the no-repair condition, $F(1, 32) = .09, p = .76$. This indicated that in the experimental condition, as time increased, guilt scores decreased, $b = -.88, t(29) = -3.99, p < .001$. The size of the effect was large in magnitude, $d = 1.05$.

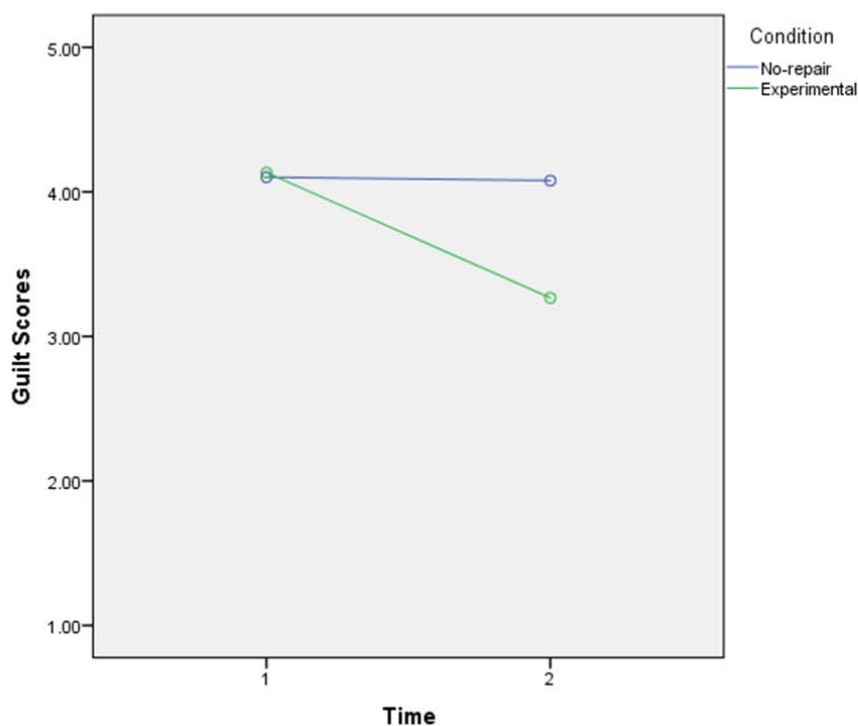


Figure 9. Plot of the interaction of condition and time on children's rated guilt.

4 DISCUSSION

The purpose of the current study was to investigate functional relations between guilt and reparative behaviors in children using a novel methodology in which children were led to believe that they had transgressed against a peer. Although previous research in adults has demonstrated that guilt motivates reparative behaviors (Cryder et al., 2012) and that, in turn, reparative behaviors alleviate guilt (de Hooge, 2012), these theoretical assumptions remain largely untested in children. Children in this study who were led to believe that they had transgressed to cause a peer's distress reported significantly greater guilt than children who were led to believe that they had caused a peer's slightly positive emotions. Children who experienced guilt engaged in significantly greater reparative behaviors across two separate measures than children who did not experience guilt. Further, among children who experienced guilt, children who were given the

opportunity to repair their transgression reported reduced guilt over time, whereas guilt did not decrease in children who were prevented from repairing their transgression. Before interpreting the primary analyses, findings related to the study design are discussed.

4.1 Validity of the transgression paradigm.

Four manipulation checks were used to establish the validity of the transgression paradigm. Three manipulation checks were used to confirm that the distress feedback successfully induced both the affective and cognitive components of guilt. First, children who were led to believe that they had transgressed to cause a peer's distress self-reported significantly greater guilt immediately following the transgression on a rating scale than children who were led to believe that they caused a peer's slightly positive emotions. Second, when asked to freely describe their emotions, a significantly greater proportion of children in a transgression condition used a negative emotion word than children who were not led to believe that they had transgressed. However, no transgressor children freely responded that they felt "guilty" and the majority reported feeling "sad" following the transgression. Although by middle childhood, children can provide narratives of their autobiographical guilt experiences (Gavazzi et al., 2011; Williams & Bybee, 1994) and can provide accurate answers when asked what the word "guilt" means (Berti et al., 2000; Harris et al., 1987), children may struggle to use the emotion term "guilt" to describe guilt experiences. In one study, 6-to 10-year-old children heard stories about a child protagonist who committed a transgression and felt "guilty"; 41% of children used basic emotion labels, particularly "sad" when recalling the emotion of the protagonist, rather than the term "guilt" (Davidson, 2006). Children may similarly struggle to use the emotion label "guilt" to describe their own emotions following transgressions.

Third, the vast majority of children in a condition including the distress feedback displayed understanding of responsibility for causing the pretend peer's distress—the cognitive component of guilt—by either self-reporting that they were at fault in response to a yes/no question or by apologizing in the note they wrote to the pretend peer. Children tended to indicate fault on one, but not both, of these measures. It is possible that some children self-reported that they were not at fault due to shame that was aroused by direct questioning rather than failure to understand their responsibility for the transgression; indeed, shame is associated with children's avoidance and denial of fault (Barrett et al., 1993). Studies that utilize paradigms in which children's guilt is experimentally manipulated should not rely solely on direct questioning to assess children's understanding of fault due to the potential interference of shame.

One, final manipulation check was used to confirm that the number of stickers given to the pretend peer was a valid measure of reparative behavior. Although a similar measure of currency, number of dollars given, was used as a reparative behavior measure in an adult experimental study (Cryder et al., 2012), manipulation checks were not conducted to assess the validity of currency as a measure of reparative behavior in that study and currency has never before been used as a reparative behavior measure in a study of children. A significantly greater proportion of children in a transgression condition than the condition in which children were not led to believe they had transgressed reported a prosocial reason for giving stickers to the distressed child, supporting the variable as a measure of reparative behavior.

In sum, the current study provides indications that children experienced both the affective and cognitive components of guilt following the transgression and that sharing stickers with the peer victim was a valid measure of reparative behavior. Given that the two previous studies of children's guilt and reparative behaviors assessed children's emotions and behaviors in response

to *hypothetical* transgressions, it was particularly important to establish the validity of the transgression paradigm and reparation measure in the current study, the first to experimentally study children's *actual* guilt and reparative behaviors.

4.2 Guilt motivates children's reparative behaviors.

Consistent with expectations, children who were led to believe that they had harmed a peer (i.e., children who experienced guilt) engaged in more reparative behavior across two separate measures than children who did not experience guilt. Specifically, children who experienced guilt gave more stickers to the pretend peer victim and also wrote notes to the peer that contained significantly more sophisticated prosocial content than children who did not experience guilt. Although several previous studies in adults have demonstrated that greater guilt is associated with greater reparative behaviors (Bassett et al., 2006; Ghorbani et al., 2013; Leunissen et al., 2013; Riek et al., 2014; Riek, 2010; Silfver-Kuhlampi et al., 2015), only two demonstrated a causal link using an experimental design (Amodio et al., 2007; Cryder et al., 2012). Moreover, only two studies have demonstrated the association in children (Colasante et al., 2014; Ferguson, 1991). The current finding provides evidence that guilt functions as theorized, by causing reparative behavior aimed to counter the wrongful act, in middle childhood (Barrett, 1995; Tangney et al., 2007; Tracy & Robins, 2004).

The finding that guilt causes reparative behavior is important given the adaptive benefit of reparative behaviors for psychological functioning. Reparative behaviors have been positively associated with a range of healthy outcomes, such as social acceptance, positive affect, self-compassion and self-esteem, and negatively associated with poorer outcomes, such as externalizing and internalizing problems, across children and adults (Bafunno & Camodeca, 2013; J. A. Bybee, Zigler, & Berliner, 1996; Caprara, Barbaranelli, Pastorelli, Cermak, & Rosza,

2001; Cohen, Wolf, Panter, & Insko, 2011; Howell, Dopko, Turowski, & Buro, 2011; Kochanska, Koenig, Barry, Kim, & Yoon, 2010; Luyten et al., 2002; Wolf, Cohen, Panter, & Insko, 2010). Thus, the finding that guilt causes these adaptive behaviors suggests a potential benefit of parent socialization of children's healthy guilt following transgressions. Authoritative parenting and discipline that is low in power assertion—punishment or threat of punishment—and includes inductive reasoning—wherein parents teach children about others' needs and the effects of transgressions on victims—have been associated with children's greater guilt (Eisenberg, 2000; Kochanska et al., 2002; Kochanska, 1991; Scarnier et al., 2009; Zahn-Waxler et al., 1979). The current finding may inform the development of parenting interventions such as those that provide psychoeducation on the potential benefits of children's experience of guilt following transgressions. Future investigations might also empirically establish components of effective inductions to inform potential interventions aimed at the socialization of healthy guilt.

As previously described, whether guilt functions adaptively or maladaptively remains highly debated. The current finding that guilt causes positive social behaviors lends support to perspectives that emphasize the adaptive potential of acute experiences of guilt following a transgression. Previous studies have examined linear relations between guilt and reparative behavior, finding that greater guilt is typically associated with greater reparative behavior; however, theories also describe that guilt can be maladaptive when it is extreme and/or excessive (Tilghman-Osborne et al., 2012). Thus, future studies should test the possibility of a negative quadratic relationship in which moderately high levels of guilt motivate reparative behavior, whereas extremely high and low levels of guilt are associated with low levels of reparation. Indeed, in one study of adults, participants who reported extremely high guilt following a hypothetical transgression considered the transgression irreparable (Nencini & Meneghini,

2013). The novel paradigm created for the current study was mild and unlikely to elicit the kind of extreme, excessive guilt described in theories of maladaptive guilt, likely explaining the positive association between guilt and reparative behavior. Indeed, greater perceived transgression severity has been associated with greater guilt (Riek et al., 2014), and highly severe transgressions may elicit maladaptively intense guilt that may impede reparation. Future studies should test possible negative quadratic associations between guilt and reparative behavior through examining individual differences in guilt and reparative responses to an identical transgression.

Future studies should also examine potential moderators of the association between guilt and reparative behavior as well as additional motivators of reparative behavior. The current study examined children's guilt and reparation as a result of a transgression against an unknown peer. As greater interpersonal closeness between the transgressor and victim has been related to greater reparative behavior in adults (Ghorbani et al., 2013; Riek et al., 2014; Riek, 2010), future research should investigate whether the relation between guilt and reparative behavior is stronger when a child has transgressed against a close relational partner. Similarly, although this study focuses on guilt as a critical predictor of reparative behaviors, other predictors warrant investigation, particularly individual difference factors that may influence whether a child is willing/able to use reparative behaviors following a transgression. One individual difference factor that has been investigated in relation to young children's prosocial behavior more generally is behavioral inhibition. Perhaps surprisingly, studies have found that children's levels of behavioral inhibition do not directly predict their willingness to engage in prosocial behavior (Hastings, Rubin, & DeRose, 2005). Other temperamental factors may be more relevant; for example, temperamentally shame-prone children likely have greater reparative difficulty, as

shame may overshadow concurrent adaptive guilt and is associated with withdrawal and avoidance rather than approach and prosocial behavior (Barrett et al., 1993). Moreover, children's socialization histories likely impact their reparative abilities. Children's willingness to behave prosocially depends on their self-efficacy, or belief in their ability to help (Bandura, 1977), and socialization practices likely equip children with specific reparative strategies, increasing their self-efficacy and reparative responding. Finally, studies have found that children raised in families characterized by lower socioeconomic status (SES) and/or parental education are less prosocial than children from higher SES backgrounds (e.g., Lichter, Shanahan, & Gardner, 2002), which may reflect stress and deprivation that increases children's self-concern or reduced availability of caregivers to teach and model behaviors. Future research should examine whether SES similarly impacts children's reparative prosocial responding.

4.3 Reparative behaviors alleviate children's guilt.

Among children who were led to believe that they had harmed a peer, only children who were given the opportunity to repair their transgression (i.e., by giving stickers and writing a note to the victimized peer) reported reduced guilt over time; in contrast, levels of guilt remained high in children who were prevented from engaging in reparative behavior. Although previous studies in adults have demonstrated that reparative behaviors alleviate guilt (Carpenter et al., 2014; de Hooge, 2012; Meek et al., 1995; Witvliet et al., 2011, 2002), the current study is the first to demonstrate this relation in children and to examine a causal mechanism in the context of an *in vivo* transgression. This finding supports the theory that reparative behaviors alleviate negative, potentially harmful guilt emotions in middle childhood (Quiles & Bybee, 1997; Tangney et al., 2007). The current study also highlights critical understudied intrapersonal benefits of reparative behaviors for transgressors, as previous studies have overwhelmingly

focused on relationship-mending functions (Baumeister, Stillwell, & Heatherton, 1994) or positive consequences for victims (e.g., improved mood, reduced anger, forgiveness; Kirchhoff et al., 2012; Smith & Harris, 2012).

The current finding that reparative behaviors successfully reduce children's guilt coupled with previous findings that unalleviated, maladaptive guilt is associated with psychopathology, particularly depression, in children (Luby, Belden, Sullivan, et al., 2009; Tilghman-Osborne et al., 2012) point to potential benefits of developing reparative skills and willingness to use such skills in childhood. Reparative behaviors may protect against the development of unalleviated, maladaptive guilt and depression, or may help mitigate maladaptive guilt symptoms in currently depressed children. Interventions may facilitate development and use of these important skills. Only one intervention has examined reparative behaviors, finding that adults who participated in a 2.5-hour intervention teaching communication skills to use when apologizing reported higher relationship quality than control participants who watched 2.5 hours of nature videos (Miller, 2009); however, benefits for transgressors were not measured. Although there are several evidence-based prosocial interventions for children (Caprara, Luengo Kanacri, Zuffiano, Gerbino, & Pastorelli, 2015; Kramer, Caldarella, Young, Fischer, & Warren, 2014; McCarty, Teie, McCutchen, & Geller, 2016; Menting, Orobio de Castro, & Matthys, 2013), these programs often target prosocial behaviors with the goal of reducing externalizing problems (e.g., aggression, bullying) and no known interventions or components of interventions specifically target reparative behaviors. Studies that examine the effects of reparative interventions in alleviating guilt and protecting against psychopathology in children may provide important information about how to reduce maladaptive intrapersonal and interpersonal consequences of transgressions. Such interventions could contain components to teach reparative behavioral

skills, encourage children to approach and use reparative skills in transgressor contexts, and challenge cognitions that transgressions are irreparable.

This study found that reparative behaviors successfully reduced guilt despite the fact that the second and final guilt measurement was taken mere minutes following children's repair opportunity, highlighting the effectiveness of reparative behaviors in alleviating guilt. Future studies should measure guilt after a delay, as it is possible that the effect of reparative behaviors in reducing guilt is even greater at a timepoint more distal to the transgression than examined in this study. Reparative behaviors were also effective in reducing guilt in this study despite the fact that children did not receive feedback that their repair attempts were successful at improving the pretend peer's mood, nor did they receive feedback that the pretend peer had forgiven them for the transgression. Future studies should also investigate the possibility that children may gain additional guilt-alleviation benefits when they receive feedback that their attempts at alleviating another's mood and/or gaining forgiveness have been successful.

4.4 Toward an empirical model of guilt, reparative behaviors, and psychopathology.

As detailed previously and illustrated in Figure 1, theorists posit that guilt motivates reparative behaviors, which alleviate guilt and are thus associated with better psychological outcomes. On the other hand, lack of reparative behavior leads to maladaptive guilt and subsequent psychopathology. The current study's findings that guilt causes reparative behaviors, which in turn alleviate guilt, provide an important foundation for future investigations of the role of reparative behaviors in protecting against maladaptive guilt and psychopathology.

This study demonstrated causal links between guilt and reparative behaviors in middle childhood. Future research should examine whether these findings extend to other developmental periods, perhaps particularly very young children, for whom maladaptive guilt is

a highly specific depressive symptom (Luby, Belden, Pautsch, et al., 2009), and adolescents, who experience higher levels of depression compared to other age groups of children (Thapar et al., 2012). Similarly, only one study has demonstrated that reparative behaviors explain relations between guilt and healthy outcomes in adults (Luyten et al., 2002); replication of this finding in children is needed. Further, despite evidence that reparative behaviors are associated with a range of healthy psychological outcomes in children (e.g., Bafunno & Camodeca, 2013; Luby, Belden, Sullivan, et al., 2009) the theoretical assumption that reparative behaviors are related to children's healthy psychological functioning in part *because* of their guilt-alleviating function¹⁰ has not been studied. Evidence that reparative behaviors lead to healthy outcomes through guilt alleviation will be important for establishing the potential utility of reparative interventions. Future studies should track autobiographical transgressions longitudinally or utilize laboratory transgressions at multiple time points to assess how relations between transgressors' typical guilt, maladaptive guilt, and reparative responses unfold across time; individuals who display greater reparative tendencies across time should be protected against maladaptive guilt through the alleviation of acute guilt that reparative behaviors accomplish, and should thus display better psychological outcomes. In contrast, individuals who experience repeated reparative difficulty and maladaptive guilt may display greater psychopathology.

4.5 Limitations.

Several limitations of the study should be noted. It is always possible that paradigms that include deception may appear artificial to some children; however, children's behavioral and affective responses to the paradigm suggest that it was realistic to the majority of children. For ethical reasons, the transgression designed for this paradigm was quite mild, which may have

¹⁰ There are likely other mechanisms through which reparative behaviors are related to healthy outcomes (e.g., better social functioning).

limited the intensity of guilt emotions induced. Moreover, children were led to believe that they had committed a transgression against a White, female peer; although children's guilt and reparative behavior did not differ by gender or race of the participating child, utilizing images of children of varied genders and races would demonstrate the generalizability of the current study's findings related to functional relations between guilt and reparative behavior. Although the guilt scale created for the current study demonstrated good internal consistency, it was not a validated scale; future work on the validity of this scale should be done, perhaps by comparing children's scores on the scale to a) parents' reports of children's typical guilt responses to transgressions and/or b) children's coded facial and gestural indications of guilt during the paradigm. Further, although practical limitations related to conducting research in a fast-paced, museum setting prohibited asking children about emotions other than guilt, future studies might also ask children about shame to empirically demonstrate the theoretical assumption that only guilt, and not shame, motivates reparative behaviors. Finally, the sample was predominantly White and although data on socioeconomic status were not collected, it is possible that many children came from families of relatively high SES given the cost of museum admission. Studies with samples of differing demographic characteristics will be important for understanding the generalizability of functional relations between guilt and reparative behaviors in children, particularly given findings indicating that SES may impact children's prosocial responding.

4.6 Summary.

The current study presents a novel and valid paradigm through which functional relations between guilt and reparative behaviors in children were examined. This study found that children who experienced guilt engaged in more reparative behavior and that children's reparative behaviors were effective in alleviating guilt. This study's findings may facilitate the

future investigation of the role of reparative behaviors in determining whether guilt is alleviated and associated with better psychological outcomes, or remains unalleviated and associated with psychopathology.

REFERENCES

- Aksan, N., & Kochanska, G. (2005). Conscience in childhood: Old questions, new answers. *Developmental Psychology, 41*(3), 506–516. <http://doi.org/10.1037/0012-1649.41.3.506>
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*. *American Journal of Psychiatry*.
<http://doi.org/10.1176/appi.books.9780890425596.744053>
- Amodio, D. M., Devine, P. G., & Harmon-Jones, E. (2007). A dynamic model of guilt: Implications for motivation and self-regulation in the context of prejudice. *Psychological Science, 18*(6), 524–530. <http://doi.org/10.1111/j.1467-9280.2007.01933.x>
- Bafunno, D., & Camodeca, M. (2013). Shame and guilt development in preschoolers: The role of context, audience and individual characteristics. *European Journal of Developmental Psychology, 10*(2), 128–143. <http://doi.org/10.1080/17405629.2013.765796>
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215.
- Barrett, K. C. (1995). A functionalist approach to shame and guilt. In J. P. Tangney & K. W. Fischer (Eds.), *Self-conscious emotions: The psychology of shame, guilt, embarrassment, and pride* (pp. 25–63). New York, NY: Guilford Press.
- Barrett, K. C. (2005). The origins of social emotions and self-regulation in toddlerhood: New evidence. *Cognition & Emotion, 19*(7), 953–979.
<http://doi.org/10.1080/02699930500172515>
- Barrett, K. C., Zahn-Waxler, C., & Cole, P. M. (1993). Avoiders vs. amenders: Implications for the investigation of guilt and shame during toddlerhood? *Cognition & Emotion, 7*(6), 481–505. <http://doi.org/10.1080/02699939308409201>

- Bassett, R. L., Bassett, Kelly, M., Lloyd, M. W., & Johnson, J. L. (2006). Seeking forgiveness: Considering the role of moral emotions. *Journal of Psychology and Theology, 34*(2), 111–124.
- Baumeister, R. F., Stillwell, A. M., & Heatherton, T. F. (1994). Guilt: an interpersonal approach. *Psychological Bulletin, 115*(2), 243–267. <http://doi.org/10.1037/0033-2909.115.2.243>
- Benenson, J. F., Pascoe, J., & Radmore, N. (2007). Children's altruistic behavior in the dictator game. *Evolution and Human Behavior, 28*(3), 168–175.
<http://doi.org/10.1016/j.evolhumbehav.2006.10.003>
- Berti, A. E., Garattoni, C., & Venturini, B. (2000). The understanding of sadness, guilt and shame in 5-, 7-, and 9-year-old children. *Genetic, Social, and General Psychology Monographs, 126*(3), 293–318.
- Blake, P. R., & Rand, D. G. (2010). Currency value moderates equity preference among young children. *Evolution and Human Behavior, 31*(3), 210–218.
<http://doi.org/10.1016/j.evolhumbehav.2009.06.012>
- Bybee, J. A., Merisca, R., & Velasco, R. (1998). The development of reactions to guilt-producing events. In J. Bybee (Ed.), *Guilt and Children* (p. 185). San Diego, CA: Academic Press.
- Bybee, J. A., Zigler, E., & Berliner, D. (1996). Guilt, guilt-evoking events, depression and eating disorders. *Current Psychology: Developmental, Learning, Personality, Social, 15*(2), 113–127. <http://doi.org/10.1007/BF02686944>
- Byrt, T., Bishop, J., & Carlin, J. B. (1993). Bias, prevalence and kappa. *Journal of Clinical Epidemiology, 46*(5), 423–429.

- Caprara, G. V., Barbaranelli, C., Pastorelli, C., Cermak, I. V. O., & Rosza, S. (2001). Facing guilt: Role of negative affectivity, need for reparation, and fear of punishment in leading to prosocial behaviour and aggression. *European Journal of Personality, 15*(3), 219–237. <http://doi.org/10.1002/per.402>
- Caprara, G. V., Luengo Kanacri, B. P., Zuffiano, A., Gerbino, M., & Pastorelli, C. (2015). Why and how to promote adolescents' prosocial behaviors: Direct, mediated and moderated effects of the CEPIDEA school-based program. *Journal of Youth and Adolescence, 44*(12), 2211–2229. <http://doi.org/10.1007/s10964-015-0293-1>
- Carpenter, T. P., Carlisle, R. D., & Tsang, J.-A. (2014). Tipping the scales: Conciliatory behavior and the morality of self-forgiveness. *The Journal of Positive Psychology, 9*(5), 389–401. <http://doi.org/10.1080/17439760.2014.910823>
- Chambers, C. T., & Johnston, C. (2002). Developmental differences in children's use of rating scales. *Journal of Pediatric Psychology, 27*(1), 27–36. <http://doi.org/10.1093/jpepsy/27.1.27>
- Chapman, J. W., & Tunmer, W. E. (1995). Development of young children's reading self-concepts: An examination of emerging subcomponents and their relationship with reading achievement. *Journal of Educational Psychology, 87*(1), 154–167. <http://doi.org/10.1037/0022-0663.87.1.154>
- Chapman, M., Zahn-Waxler, C., Cooperman, G., & Iannotti, R. (1987). Empathy and responsibility in the motivation of children's helping. *Developmental Psychology, 23*(1), 140–145. <http://doi.org/10.1037/0012-1649.23.1.140>
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*(4), 284–290. <http://doi.org/10.1037//1040-3590.6.4.284>

- Cohen, T. R., Wolf, S. T., Panter, a T., & Insko, C. a. (2011). Introducing the GASP scale: A new measure of guilt and shame proneness. *Journal of Personality and Social Psychology*, *100*(5), 947–966. <http://doi.org/10.1037/a0022641>
- Colasante, T., Zuffiano, A., Bae, N. Y., & Malti, T. (2014). Inhibitory control and moral emotions: Relations to reparation in early and middle childhood. *The Journal of Genetic Psychology*, *175*(6), 511–527. <http://doi.org/10.1017/CBO9781107415324.004>
- Cooney, R. E., Atlas, L. Y., Joormann, J., Eugene, F., & Gotlib, I. H. (2006). Amygdala activation in the processing of neutral faces in social anxiety disorder: Is neutral really neutral? *Psychiatry Research - Neuroimaging*, *148*(1), 55–59. <http://doi.org/10.1016/j.pscychresns.2006.05.003>
- Cryder, C. E., Springer, S., & Morewedge, C. K. (2012). Guilty feelings, targeted actions. *Personality & Social Psychology Bulletin*, *38*(5), 607–18. <http://doi.org/10.1177/0146167211435796>
- Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York: G. P. Putnam's Sons.
- Davidson, D. (2006). The role of basic, self-conscious and self-conscious evaluative emotions in children's memory and understanding of emotion. *Motivation and Emotion*, *30*(3), 232–242. <http://doi.org/10.1007/s11031-006-9037-6>
- De Hooge, I. E. (2012). The exemplary social emotion guilt: Not so relationship-oriented when another person repairs for you. *Cognition & Emotion*, *26*(7), 1189–1207. <http://doi.org/10.1080/02699931.2011.640663>

- Dearing, R. L., Stuewig, J., & Tangney, J. P. (2005). On the importance of distinguishing shame from guilt: Relations to problematic alcohol and drug use. *Addictive Behaviors, 30*(7), 1392–1404. <http://doi.org/10.1016/j.addbeh.2005.02.002>
- Donatelli, J.-A. L., Bybee, J. A., & Buka, S. L. (2007). What do mothers make adolescents feel guilty about? Incidents, reactions, and relation to depression. *Journal of Child and Family Studies, 16*, 859–875. <http://doi.org/10.1007/s10826-006-9130-1>
- Eisenberg, N. (2000). Emotion, regulation, and moral development. *Annual Review of Psychology, 51*, 665–697. <http://doi.org/10.1146/annurev.psych.51.1.665>
- Eisenberg, N., & Fabes, R. A. (1998). Prosocial development. In W. Damon & N. Eisenberg (Eds.), *Handbook of child psychology, 5th ed.: Vol 3. Social, emotional, and personality development*. (pp. 701–778). Hoboken, NJ: John Wiley & Sons, Inc.
- Ely, R., & Gleason, J. B. (2006). I'm sorry I said that: Apologies in young children's discourse. *Journal of Child Language, 33*(3), 599–620.
<http://doi.org/http://dx.doi.org/10.1017/S0305000906007446>
- Exline, J. J., Deshea, L., & Holeman, V. T. (2007). Is apology worth the risk? Predictors, outcomes, and ways to avoid regret. *Journal of Social and Clinical Psychology, 26*(4), 479–504. <http://doi.org/10.1521/jscp.2007.26.4.479>
- Exline, J. J., Root, B. L., Yadavalli, S., Martin, A. M., & Fisher, M. L. (2011). Reparative behaviors and self-forgiveness: Effects of a laboratory-based exercise. *Self and Identity, 10*(1), 101–126. <http://doi.org/10.1080/15298861003669565>
- Fehr, E., Bernhard, H., & Rockenbach, B. (2008). Egalitarianism in young children. *Nature, 454*(7208), 1079–1083. <http://doi.org/10.1038/nature07155>

- Ferguson, T. J. (1991). Children's understanding of guilt and shame. *Child Development, 62*(4), 827–839.
- Ferguson, T. J., & Stegge, H. (1998). Measuring guilt in children: A rose by any other name still has thorns. In J. A. Bybee (Ed.), *Guilt and Children* (pp. 19–74). San Diego, CA: Academic Press. <http://doi.org/10.1016/B978-012148610-5/50003-5>
- Field, A. (2009). *Discovering Statistics Using SPSS*. Sage Publishing.
<http://doi.org/10.1234/12345678>
- Garner, P. W. (2003). Child and family correlates of toddlers' emotional and behavioral responses to a mishap. *Infant Mental Health Journal, 24*(6), 580–596.
<http://doi.org/10.1002/imhj.10076>
- Gavazzi, I. G., Ornaghi, V., & Antoniotti, C. (2011). Children's and adolescents' narratives of guilt: Antecedents and mentalization. *European Journal of Developmental Psychology, 8*(3), 311–330. <http://doi.org/10.1080/17405629.2010.491303>
- Ghorbani, M., Liao, Y., Çayköylü, S., & Chand, M. (2013). Guilt, shame, and reparative behavior: The effect of psychological proximity. *Journal of Business Ethics, 114*(2), 311–323. <http://doi.org/10.1007/s10551-012-1350-2>
- Gummerum, M., Hanoch, Y., Keller, M., Parsons, K., & Hummel, A. (2010). Preschoolers' allocations in the dictator game: The role of moral emotions. *Journal of Economic Psychology, 31*(1), 25–34. <http://doi.org/10.1016/j.joep.2009.09.002>
- Haapasalo, J., Tremblay, R. E., Boulerice, B., & Vitaro, F. (2000). Relative advantages of person- and variable-based approaches for predicting problem behaviors from Kindergarten assessments. *Journal of Quantitative Criminology, 16*(2), 145–168.
<http://doi.org/http://dx.doi.org/10.1023/A:1007512521780>

- Harder, D. W., Cutler, L., & Rockart, L. (1992). Assessment of shame and guilt and their relationships to psychopathology. *Journal of Personality Assessment*, *59*(3), 584–604.
http://doi.org/10.1207/s15327752jpa5903_12
- Harris, P. L., Olthof, T., Terwogt, M. M., & Hardman, C. E. (1987). Children's knowledge of the situations that provoke emotion. *International Journal of Behavioral Development*, *10*(3), 319–343.
- Harter, S., & Whitesell, N. R. (1989). Developmental changes in children's understanding of single, multiple, and blended emotion concepts. In C. Saarni & P. L. Harris (Eds.), *Children's understanding of emotions* (pp. 81–116). Cambridge, England: Cambridge University Press.
- Hastings, P. D., Rubin, K. H., & DeRose, L. (2005). Links among gender, inhibition, and parental socialization in the development of prosocial behavior. *Merrill-Palmer Quarterly*, *51*(4), 467–493. <http://doi.org/10.1353/mpq.2005.0023>
- Hay, D. F., & Cook, K. V. (2007). The transformation of prosocial behavior from infancy to childhood. In C. A. Brownell & C. B. Kopp (Eds.), *Socioemotional development in the toddler years: Transitions and transformations* (pp. 100–131). New York, NY: Guilford Press.
- Hay, D. F., Payne, A., & Chadwick, A. (2004). Peer relations in childhood. *Journal of Child Psychology and Psychiatry*, *45*(1), 84–108. <http://doi.org/10.1046/j.0021-9630.2003.00308.x>
- Howell, A. J., Dopko, R. L., Turowski, J. B., & Buro, K. (2011). The disposition to apologize. *Personality and Individual Differences*, *51*(4), 509–514.
<http://doi.org/10.1016/j.paid.2011.05.009>

- Hymel, S., Rubin, K. H., Rowden, L., & LeMare, L. (1990). Children's peer relationships: Longitudinal prediction of internalizing and externalizing problems from middle to late childhood. *Child Development, 61*(6), 2004. <http://doi.org/10.2307/1130854>
- Jones, W. H., & Kugler, K. (1993). Interpersonal correlates of the guilt inventory. *Journal of Personality Assessment, 61*(2), 246–258. http://doi.org/10.1207/s15327752jpa6102_6
- Kim, S., Thibodeau, R., & Jorgensen, R. S. (2011). Shame, guilt, and depressive symptoms: A meta-analytic review. *Psychological Bulletin, 137*(1), 68–96. <http://doi.org/10.1037/a0021466>
- Kirchhoff, J., Wagner, U., & Strack, M. (2012). Apologies: Words of magic? The role of verbal components, anger reduction, and offence severity. *Peace and Conflict: Journal of Peace Psychology, 18*(2), 109–130. <http://doi.org/10.1037/a0028092>
- Kochanska, G. (1991). Socialization and temperament in the development of guilt and conscience. *Child Development, 62*(6), 1379–1392. <http://doi.org/10.2307/1130813>
- Kochanska, G., Barry, R. A., Jimenez, N. B., Hollatz, A. L., & Woodard, J. (2009). Guilt and effortful control: Two mechanisms that prevent disruptive developmental trajectories. *Journal of Personality and Social Psychology, 97*(2), 322–333. <http://doi.org/10.1037/a0015471>
- Kochanska, G., DeVet, K., Goldman, M., Murray, K., & Putnam, S. P. (1994). Maternal reports of conscience development and temperament in young children. *Child Development, 65*(3), 852–868. <http://doi.org/10.2307/1131423>
- Kochanska, G., Forman, D. R., Aksan, N., & Dunbar, S. B. (2005). Pathways to conscience: Early mother-child mutually responsive orientation and children's moral emotion, conduct,

- and cognition. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 46(1), 19–34. <http://doi.org/10.1111/j.1469-7610.2004.00348.x>
- Kochanska, G., Gross, J. N., Lin, M.-H., & Nichols, K. E. (2002). Guilt in young children: Development, determinants, and relations with a broader system of standards. *Child Development*, 73(2), 461–82. <http://doi.org/10.1111/1467-8624.00418>
- Kochanska, G., Koenig, J. L., Barry, R. A., Kim, S., & Yoon, J. E. (2010). Children's conscience during toddler and preschool years, moral self, and a competent, adaptive developmental trajectory. *Developmental Psychology*, 46(5), 1320–1332. <http://doi.org/10.1037/a0020381.Children>
- Kramer, T. J., Caldarella, P., Young, K. R., Fischer, L., & Warren, S. (2014). Implementing strong kids school-wide to reduce internalizing behaviors and increase prosocial behaviors. *Education and Treatment of Children*, 37(4), 659–680. <http://doi.org/10.1353/etc.2014.0031>
- Lancy, D. F., & Grove, M. A. (2011). Getting noticed: Middle childhood in cross-cultural perspective. *Human Nature*, 22(3), 281–302. <http://doi.org/10.1007/s12110-011-9117-5>
- Leitenberg, H., Yost, L. W., & Carroll-Wilson, M. (1986). Negative cognitive errors in children: Questionnaire development, normative data, and comparisons between children with and without self-reported symptoms of depression, low self-esteem, and evaluation anxiety. *Journal of Consulting and Clinical Psychology*, 54(4), 528–536. <http://doi.org/10.1037/0022-006X.54.4.528>
- Leunissen, J. M., De Cremer, D., Folmer, C. P. R., & van Dijke, M. (2013). The apology mismatch: Asymmetries between victim's need for apologies and perpetrator's willingness to apologize. *Journal of Experimental Social Psychology*, 49(3), 315–324. <http://doi.org/10.1016/j.jesp.2012.12.005>

- Lewis, H. B. (1971). Shame and guilt in neurosis. *Psychoanalytic Review*.
[http://doi.org/10.1016/S0002-7138\(09\)61341-8](http://doi.org/10.1016/S0002-7138(09)61341-8)
- Lewis, M. (2008). Self-conscious emotions: Embarrassment, pride, shame, and guilt. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (pp. 742–756). New York: Guilford Press.
- Lewis, M., Sullivan, M. W., Stanger, C., & Weiss, M. (1989). Self development and self-conscious emotions. *Child Development*, *60*(1), 146–56. <http://doi.org/10.2307/1131080>
- Lichter, D. T., Shanahan, M. J., & Gardner, E. L. (2002). Helping others? The effects of childhood poverty and family instability on prosocial behavior. *Youth & Society*, *34*(1), 89–119. <http://doi.org/10.1177/0044118X02034001004>
- Lindsay-Hartz, J. (1984). Contrasting experiences of shame and guilt. *American Behavioral Scientist*, *27*(6), 689–704. <http://doi.org/10.1177/000276484027006003>
- Luby, J. L., Belden, A. C., Pautsch, J., Si, X., & Spitznagel, E. (2009). The clinical significance of preschool depression: Impairment in functioning and clinical markers of the disorder. *Journal of Affective Disorders*, *112*(1-3), 111–9. <http://doi.org/10.1016/j.jad.2008.03.026>
- Luby, J. L., Belden, A., Sullivan, J., Hayen, R., McCadney, A., & Spitznagel, E. (2009). Shame and guilt in preschool depression: Evidence for elevations in self-conscious emotions in depression as early as age 3. *Journal of Child Psychology and Psychiatry*, *50*(9), 1156–1166. <http://doi.org/10.1111/j.1469-7610.2009.02077.x>
- Luyten, P., Fontaine, J. R. J., & Corveleyn, J. (2002). Does the Test of Self-Conscious Affect (TOSCA) measure maladaptive aspects of guilt and adaptive aspects of shame? An empirical investigation. *Personality and Individual Differences*, *33*(8), 1373–1387.
[http://doi.org/10.1016/S0191-8869\(02\)00197-6](http://doi.org/10.1016/S0191-8869(02)00197-6)

- Marschall, D. E., Sanftner, J. L., & Tangney, J. P. (1994). *The State Shame and Guilt Scale (SSGS)*. Fairfax, VA: George Mason University.
- Marsh, H. W., Ellis, L. a., & Craven, R. G. (2002). How do preschool children feel about themselves? Unraveling measurement and multidimensional self-concept structure. *Developmental Psychology*, *38*(3), 376–393. <http://doi.org/10.1037//0012-1649.38.3.376>
- McCarty, S., Teie, S., McCutchen, J., & Geller, E. S. (2016). Actively caring to prevent bullying in an elementary school: Prompting and rewarding prosocial behavior. *Journal of Prevention & Intervention in the Community*, *44*(3), 164–176. <http://doi.org/10.1080/10852352.2016.1166809>
- Meek, K. R., Albright, J. S., & McMinn, M. R. (1995). Religious orientation, guilt, confession, and forgiveness. *Journal of Psychology and Theology*, *23*(3), 190–197.
- Menting, A. T. a, Orobio de Castro, B., & Matthys, W. (2013). Effectiveness of the Incredible Years parent training to modify disruptive and prosocial child behavior: A meta-analytic review. *Clinical Psychology Review*, *33*(8), 901–913. <http://doi.org/10.1016/j.cpr.2013.07.006>
- Miller, A. J. (2009). *An intervention to improve communication of reproach and apology after transgressions*. Virginia Commonwealth University.
- Muris, P., & Meesters, C. (2014). Small or big in the eyes of the other: On the developmental psychopathology of self-conscious emotions as shame, guilt, and pride. *Clinical Child Family Psychology Review*, *17*(1), 19–40. <http://doi.org/10.1007/s10567-013-0137-z>
- Nencini, A., & Meneghini, A. M. (2013). How relational bonds influence strategies for coping with guilt. *Journal of Relationships Research*, *4*, 1–7. <http://doi.org/10.1017/jrr.2013.9>

- Quiles, Z. N., & Bybee, J. A. (1997). Chronic and predispositional guilt: Relations to mental health, prosocial behavior, and religiosity. *Journal of Personality Assessment*, 69(1), 104–126. http://doi.org/10.1207/s15327752jpa6901_6
- Rakow, A., Forehand, R., McKee, L., Coffelt, N., Champion, J., Fear, J., & Compas, B. (2009). The relation of parental guilt induction to child internalizing problems when a caregiver has a history of depression. *Journal of Child and Family Studies*, 18(4), 367–377. <http://doi.org/10.1007/s10826-008-9239-5>
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Sage Publishing.
- Riek, B. M. (2010). Transgressions, guilt, and forgiveness: A model of seeking forgiveness. *Journal of Psychology & Theology*, 38(4), 246–254.
- Riek, B. M., Luna, L. M. R., & Schnabelrauch, C. A. (2014). Transgressors' guilt and shame: A longitudinal examination of forgiveness seeking. *Journal of Social and Personal Relationships*, 31(6), 751–772. <http://doi.org/10.1177/0265407513503595>
- Rubin, K. H., Bukowski, W. M., & Parker, J. G. (2006). Peer interactions, relationships, and groups. *Handbook of Child Psychology, 3rd Edition*, 571–645. <http://doi.org/10.1002/jor.22544>
- Scarnier, M., Schmader, T., & Lickel, B. (2009). Parental shame and guilt: Distinguishing emotional responses to a child's wrongdoings. *Personal Relationships*, 16(2), 205–220. <http://doi.org/10.1111/j.1475-6811.2009.01219.x>
- Silfver, M. (2007). Coping with guilt and shame: A narrative approach. *Journal of Moral Education*, 36(2), 169–183. <http://doi.org/10.1080/03057240701325274>

- Silfver-Kuhlampi, M., Figueiredo, A., Sortheix, F., & Fontaine, J. (2015). Humiliated self, bad self or bad behavior? The relations between moral emotional appraisals and moral motivation. *Journal of Moral Education, 44*(2), 213–231.
<http://doi.org/10.1080/03057240.2015.1043874>
- Smith, C. E., Blake, P. R., & Harris, P. L. (2013). I should but I won't: Why young children endorse norms of fair sharing but do not follow them. *PLoS ONE, 8*(3).
<http://doi.org/10.1371/journal.pone.0059510>
- Smith, C. E., & Harris, P. L. (2012). He didn't want me to feel sad : Children's reactions to disappointment and apology. *Social Development, 21*(2), 215–228.
<http://doi.org/10.1111/j.1467-9507.2011.00606.x>
- Stegge, H., & Ferguson, T. J. (1994). *Self-Conscious Emotion: Maladaptive and Adaptive Scales (SCEMAS)*. Unpublished instrument, Vrije Universiteit Amsterdam and Utah State University.
- Tangney, J. P. (1991). Moral affect: The good, the bad, and the ugly. *Journal of Personality and Social Psychology, 61*(4), 598–607. <http://doi.org/10.1037/0022-3514.61.4.598>
- Tangney, J. P. (1996). Conceptual and methodological issues in the assessment of shame and guilt. *Behavior Research and Therapy, 34*(9), 741–754. [http://doi.org/10.1016/0005-7967\(96\)00034-4](http://doi.org/10.1016/0005-7967(96)00034-4)
- Tangney, J. P. (1998). How does guilt differ from shame? In J. Bybee (Ed.), *Guilt and children* (pp. 1–17). San Diego, CA: Academic Press.
- Tangney, J. P., Stuewig, J., & Mashek, D. J. (2007). Moral emotions and moral behavior. *Annual Review of Psychology, 58*, 345–372.
<http://doi.org/10.1146/annurev.psych.56.091103.070145>

- Tangney, J. P., Wagner, P. E., Burggraf, S. A., Gramzow, R., & Fletcher, C. (1990). *Test of Self-Conscious Affect for Children (TOSCA-C)*. Fairfax, VA: George Mason University.
- Tangney, J. P., Wagner, P. E., Hill-Barlow, D., Marschall, D. E., & Gramzow, R. (1996). Relation of shame and guilt to constructive versus destructive responses to anger across the lifespan. *Journal of Personality and Social Psychology*, *70*(4), 797–809.
<http://doi.org/10.1037/0022-3514.70.4.797>
- Tangney, J. P., Wagner, P., & Gramzow, R. (1992). Proneness to shame, proneness to guilt, and psychopathology. *Journal of Abnormal Psychology*, *101*(3), 469–478.
<http://doi.org/10.1037/0021-843X.101.3.469>
- Thapar, A., Collishaw, S., Pine, D. S., & Thapar, A. K. (2012). Depression in adolescence. *The Lancet*, *379*(9820), 1056–1067. [http://doi.org/10.1016/S0140-6736\(11\)60871-4](http://doi.org/10.1016/S0140-6736(11)60871-4)
- Tilghman-Osborne, C., Cole, D. a., & Felton, J. W. (2012). Inappropriate and excessive guilt: Instrument validation and developmental differences in relation to depression. *Journal of Abnormal Child Psychology*, *40*(4), 607–20. <http://doi.org/10.1007/s10802-011-9591-6>
- Tilghman-Osborne, C., Cole, D. a., & Felton, J. W. (2010). Definition and measurement of guilt: Implications for clinical research and practice. *Clinical Psychology Review*, *30*(5), 536–546.
<http://doi.org/10.1016/j.cpr.2010.03.007>
- Tracy, J. L., & Robins, R. W. (2004). Putting the self into self-conscious emotions: A theoretical model. *Psychological Inquiry*, *15*(2), 103–125.
http://doi.org/10.1207/s15327965pli1502_01
- Twisk, J. (2006). *Applied multilevel analysis: a practical guide. Technology*. Cambridge University Press.

- Warneken, F., Lohse, K., Melis, A. P., & Tomasello, M. (2011). Young children share the spoils after collaboration. *Psychological Science, 22*(2), 267–73.
<http://doi.org/10.1177/0956797610395392>
- West, B. T., Galecki, A. T., & Welch, K. B. (2014). *Linear mixed models*. CRC Press.
- Williams, C., & Bybee, J. A. (1994). What do children feel guilty about? Developmental and gender differences. *Developmental Psychology, 30*(5), 617–623.
<http://doi.org/10.1037/0012-1649.30.5.617>
- Witvliet, C. vanOyen, Hinman, N. G., Exline, J. J., & Brandt, T. (2011). Responding to our own transgressions: An experimental writing study of repentance, offense rumination, self-justification, and distraction. *Journal of Psychology and Christianity, 30*(3), 223–238.
- Witvliet, C. vanOyen, Ludwig, T. E., & Bauer, D. J. (2002). Please forgive me: Transgressors' emotions and physiology during imagery of seeking forgiveness and victim responses. *Journal of Psychology and Christianity, 21*(3), 219–233.
- Wolf, S. T., Cohen, T. R., Panter, A. T., & Insko, C. A. (2010). Shame proneness and guilt proneness: Toward the further understanding of reactions to public and private transgressions. *Self and Identity, 9*(4), 337–362. <http://doi.org/10.1080/15298860903106843>
- Zahn-Waxler, C., & Kochanska, G. (1990). The origins of guilt. In R. A. Thompson (Ed.), *Nebraska symposium on motivation, 1988: Socioemotional development*. (pp. 183–258). Lincoln, NE: University of Nebraska Press.
- Zahn-Waxler, C., & Radke-Yarrow, M. (1990). The origins of empathic concern. *Motivation and Emotion, 14*(2), 107–130. <http://doi.org/10.1007/BF00991639>

Zahn-Waxler, C., Radke-Yarrow, M., & King, R. A. (1979). Child rearing and children's prosocial initiations toward victims of distress. *Child Development*, 50(2), 319–330.

<http://doi.org/10.2307/1129406>

Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., & Chapman, M. (1992). Development of concern for others. *Developmental Psychology*, 28(1), 126–136.

<http://doi.org/10.1037//0012-1649.28.1.126>

APPENDIX

PROSOCIAL NOTES CODING MANUAL

Prosocial behavior is behavior that is helpful, supportive, affiliative, and intends to benefit another person.

Two types:

Altruistic prosocial behavior – prosocial behavior initiated when someone witnesses someone in need of help

Reparative prosocial behavior – prosocial behavior initiated when someone has caused another person's distress

We are interested in both. Neither altruistic nor reparative behavior is better or stronger than the other.

We will read a child's note to another child and determine how prosocial the content of the note is. That is, we will rate how strong/sophisticated any prosocial statements in the note are.

- Examples of prosocial behaviors include (not exhaustive):
 - Statements of affection - e.g., "I love you." "Are you OK?"
 - Statements that reference sharing
 - Apology (e.g., "I'm sorry." "I apologize.")

(Continue to next page for guidelines on ratings)

0 = none

- Example: “I don’t know”

1 = minimal/unclear

POSSIBLE prosocial behavior (unclear whether the behavior is truly prosocial) or clear attempt to affiliate

- Examples: “have fun!” “have a nice day!” “hello!”

Informational with the intent of possibly making another child happy or excited

- Example: “You will like this game because it has prizes”

2 = moderate**EITHER**

Clear altruistic or reparative prosocial behavior

- Example reparative: “I’m sorry” (apologizes)
- Example altruistic: “I gave you two of my stickers” (references sharing)

OR

References wanting to improve the other child’s emotions or their own emotions about something that happened to the child

- Example: “I hope you feel better;” “I am happy for you that you got what you wanted” “I am really sad you didn’t get what you wanted”

3 = strong

Clear altruistic or reparative prosocial behavior (as in rating of 2)

AND ONE OR MORE OF THE FOLLOWING:

1. References wanting to improve the other child’s emotions OR references their own emotions about something that happened to the child
2. Clearly explains the connection between the child’s own behavior and the prosocial behavior (i.e., gives clear reasoning for being prosocial)
3. Otherwise clearly a standout prosocial statement in terms of sophistication
 - Example: “I took the last ball so I’m going to give you some of my stickers” (clear prosocial statement – giving stickers + clearly explains connection between child’s own behavior and the prosocial behavior – child states that they are giving stickers because they harmed the other child)
 - Example: “I am giving you some stickers and I hope that makes you happy” (clear prosocial statement – giving stickers + references wanting to improve the other child’s mood)