

Reparative Prosocial Behaviors Alleviate Children's Guilt

Meghan Rose Donohue and Erin C. Tully
Georgia State University

Guilt theoretically functions to motivate reparative behaviors, which, in turn, theoretically alleviate guilt and prevent depression. Although empirical research supports these theories in adults, studies have 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 10-year olds ($N = 97$) were randomly assigned to 1 of 3 conditions. Children in the experimental condition were led to believe they had transgressed to cause a peer's distress. Children self-reported their guilt following the 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 than children in a no-guilt condition who were led to believe they had caused a peer's slightly positive emotions. Further, children in the experimental condition reported reduced guilt across the first to second guilt measurement, whereas children in the no-repair condition (who transgressed but did not have a repair opportunity) did not report reduced guilt over time. Results demonstrate that guilt and reparative behaviors function as theorized in middle childhood and may begin to inform reparative interventions aimed at preventing maladaptive guilt and depression.

Keywords: guilt, reparative behavior, transgressions, prosocial behavior, moral development


Transgressions and the guilt feelings that often follow are a common human experience. Theorists have 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 internalizing psychopathology, particularly depression (Quiles & Bybee, 1997). Although studies of adults have demonstrated that guilt motivates reparative behaviors and that reparative behaviors alleviate guilt (de Hooge, 2012), few studies have investigated how early in life guilt and reparative behaviors function as theorized. Investigating reparative behaviors as a means of guilt alleviation in childhood is critical, as children who display greater unalleviated, maladaptive guilt exhibit greater depression (Tilghman-Osborne, Cole, & Felton, 2012). The purpose of this 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).

Functionalist Perspectives of Guilt

Guilt is an emotional and cognitive experience triggered by a transgression that involves two necessary components: experiencing empathy—for example, negative emotion in response to another's emotional distress—and an awareness of personal responsibility for the transgression (Tilghman-Osborne, Cole, & Felton, 2010). Reparative behaviors are prosocial actions transgressors direct toward victims of their own wrongdoings that include making amends (i.e., undoing the transgression's consequences), confessing, and apologizing (Tangney, Stuewig, & Mashek, 2007). Empathy underlies both reparative behaviors as well as the general prosocial behaviors that children use when they have not caused but instead are bystanders to another's distress; thus, it is this sense of personal responsibility for causing a victim's distress that theoretically distinguishes guilt from empathy (Tilghman-Osborne et al., 2010).

Guilt is related to yet distinct from shame; although the two emotions are typically moderately correlated (Tangney et al., 2007), they theoretically lead to differing motivations and behaviors. 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 defective self cannot be mended (Tracy & Robins, 2004). Indeed, studies across children and adults have demonstrated that shame is associated with less prosocial behavior (Barrett, Zahn-Waxler, & Cole, 1993; Tangney et al., 2007). Because of the unique theoretical relationship between guilt and reparative behaviors, this study focuses on guilt rather than shame.

This article was published Online First July 25, 2019.

 Meghan Rose Donohue and Erin C. Tully, Department of Psychology, Georgia State University.

Meghan Rose Donohue's work was supported by a Georgia State University Research Services and Administration Dissertation Grant, HRSA fellowship D40HP19643 (PI: Cohen), and NIH Grant T32 MH100019 (PIs: Barch and Luby). We thank the families who participated in the study and the research assistants who assisted with coding.

Correspondence concerning this article should be addressed to Meghan Rose Donohue, who is now at the Department of Psychiatry, Washington University School of Medicine, 4444 Forest Park Avenue, Suite 2500, St. Louis, MO 63110. E-mail: rdonohue@wustl.edu

Perhaps surprisingly, developmental researchers typically present guilt as an adaptive emotion, particularly during childhood (Tilghman-Osborne et al., 2010). Indeed, guilt's primary function is to channel a transgressor's negative feelings into reparative behaviors that mend relationships and prevent internalizing psychopathology for the transgressor (Tangney et al., 2007). Studies have demonstrated the adaptive potential of guilt; some measures of the trait-like tendency to experience guilt following transgressions (i.e., guilt-proneness) have been associated with lower levels of depression and anxiety, and higher levels of self-esteem across children, adolescents, and adults (Luyten, Fontaine, & Corveleyn, 2002; Tangney et al., 2007).

Yet, despite this adaptive view of guilt, functionalist theorists posit that emotions are not inherently adaptive or maladaptive; rather, their functionality depends on how effectively they are managed (Barrett, 1995). Experiences of state guilt following a transgression that motivate reparative behavior are theorized to prevent internalizing psychopathology. In contrast, when individuals do not engage in reparative behaviors to alleviate guilt, maladaptive guilt, a destructive guilt variant that is excessive (Tilghman-Osborne et al., 2012) is theorized to develop; in a vicious cycle, maladaptive guilt further blocks reparative action, leading to chronically unalleviated guilt and, over time, depression (Bybee, Zigler, Berliner, & Merisca, 1996; Quiles & Bybee, 1997). Indeed, adults who used fewer reparative behaviors following a past transgression reported greater current maladaptive guilt over the transgression (Riek, Luna, & Schnabelrauch, 2014; Silfver, 2007), and maladaptive guilt has been associated with higher levels of depression and anxiety and poorer social functioning (Jones & Kugler, 1993; Luby et al., 2009).

The Development of Guilt and Reparative Behaviors in Childhood

Although research is limited, studies suggest that guilt and reparative behaviors become increasingly complex throughout childhood (Muris & Meesters, 2014). Studies have documented that by the second year of life, children begin to experience guilt and engage in reparative behavior following transgressions (Barrett et al., 1993; Kochanska, Gross, Lin, & Nichols, 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 social standards (Lewis, 2008). Whereas youth in middle childhood typically experience guilt as a result of others' external evaluations, adolescents typically experience guilt as a result of their own internal standards (Ferguson, 1991; Gavazzi, Ornaghi, & Antoniotti, 2011). Very few studies have examined developmental changes in reparative behaviors. Although study found that 8-year-old children engaged in greater reparative behaviors than 4-year-olds (Colasante, Zuffianò, Bae, & Malti, 2014), the relation between reparative behaviors and age may be more complex, as children's general prosocial behaviors tend to become more selective across development (Hay & Cook, 2007).

Middle childhood is a critical period in which to investigate children's responses to transgressing against another child, as children gain social independence and peer relationships become increasingly related to children's psychological adjustment during this developmental period (Hay, Payne, & Chadwick, 2004; Lancy

& Grove, 2011). By middle childhood, children also demonstrate understanding of what the word *guilt* means, enabling them to self-report their own guilt (Ferguson, 1991; Berti, Garattoni, & Venturini, 2000). For example, in one study, 5- and 6-year-old children accurately described guilt and distinguished guilt from sadness and shame on all aspects of the emotional process (e.g., definition, antecedents, action tendencies, strategies to regulate) and as well as older children aged 7- to 10-years, indicating that children as young as 5 display understanding of guilt as an emotion term (Berti et al., 2000).

Children also have the capacity to experience maladaptive guilt, which may place them at risk for depression. In one study, depressed preschoolers were characterized by both low levels of reparative behaviors and high levels of maladaptive guilt (Luby et al., 2009). In another study, 7- to 16-year-old children's greater self-reported maladaptive guilt was associated with greater self-reported depression at all ages, and was more depressotypic with older age (Tilghman-Osborne et al., 2012).

The Measurement of Guilt and Reparative Behaviors in Childhood

The measures utilized to assess children's guilt and reparative behaviors typically differ depending on the age of the children studied. Guilt and reparative behaviors arise from a specific wrongdoing and therefore researchers recommend that they be measured in the context of a specific transgression (Cryder, Springer, & Morewedge, 2012). To measure reparative behaviors in childhood, studies have almost exclusively relied on parent-report using the My-Child-2 questionnaire (Kochanska, DeVet, Goldman, Murray, & Putnam, 1994). To measure guilt in toddlers, studies have almost exclusively used *in vivo* "mishap" paradigms in which children are led to believe that they have broken an experimenter's valued possession through items that are rigged to break. Indications of guilt (e.g., avoidance, tension, arousal) and reparative behaviors (e.g., attempts to fix the items) are coded from video (Kochanska et al., 2002). All existing *in vivo* studies of toddlers examine unintentional transgressions; theorists argue that guilt primarily arises from unintentional transgressions (Baumeister, Stillwell, & Heatherton, 1994), and intentional harm may reflect aggression, perhaps particularly in children. Kochanska et al. (2002) found that children's coded guilt was longitudinally stable from 22 to 45 months and was associated with greater mother-reported guilt, supporting the stability and validity of the measure. A strength of *in vivo* paradigms is that they allow for examination of children's actual emotions and behaviors following a transgression. Although these paradigms have been used to examine the development of early guilt, no study has utilized such a paradigm to examine functional relations between children's guilt and reparative behaviors.

In contrast, studies of guilt in middle childhood and adolescence use measures in which children are presented with hypothetical transgression scenarios and rate how guilty they would feel as the transgressor. Although there are several hypothetical measures available (e.g., Test of Self-Conscious Affect for Children; Tangney, Wagner, Burggraf, Gramzow, & Fletcher, 1990), they assess children's trait-like guilt proneness rather than state guilt following a specific transgression, which precludes their use in an *in vivo* paradigm. Further, although hypothetical measures tell us how children expect they would feel and behave after transgressing, we

do not know how they would react to a real or simulated transgression.

Guilt Motivates Reparative Behaviors

Although theorists have long hypothesized that guilt motivates reparative behaviors and that reparative action alleviates guilt, nearly all of the empirical tests of these claims have occurred over the past 10 years, potentially because researchers have often conflated guilt and reparative behavior, which prohibits examination of whether guilt affect motivates reparative action. Several longitudinal and cross-sectional studies have found that adults' reports of experiencing greater guilt following autobiographical transgressions (i.e., past personal wrongdoings) as well as their reports that they would experience greater guilt if they were to commit a hypothetical transgression (e.g., sleeping through a group presentation) were associated with their greater reparative behaviors (Riek et al., 2014; Silfver, 2007).

Two studies have demonstrated that guilt predicts reparative behavior in adults using *in vivo* transgressions (Amodio, Devine, & Harmon-Jones, 2007; Cryder et al., 2012). In one experimental study, adults viewed a computer screen that presented extensive "background information" describing study procedures (Cryder et al., 2012). It was verified that participants did not read this information. Participants then chose to eat either 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). Participants rated their guilt and then completed a dictator task in which they divided five dollars with the other participant. Participants in the guilt condition gave significantly more money than control participants, demonstrating a causal relation between greater guilt and greater reparative behavior.

Studies of children have typically examined associations between children's guilt and measures of general prosocial tendencies (Caprara, Barbaranelli, Pastorelli, Cermak, & Rosza, 2001; Drummond, Hammond, Satlof-Bedrick, Waugh, & Brownell, 2017) rather than specifically reparative behaviors. Investigating whether guilt following a specific transgression motivates reparative behaviors to mend that transgression is important in order to provide a direct test of theoretical claims. Three studies have demonstrated that greater guilt in children ages 4 to 17 was related to greater specifically reparative behaviors (Colasante et al., 2014; Ferguson, 1991; Tangney, Wagner, Hill-Barlow, Marschall, & Gramzow, 1996). In a study of 4- and 8-year-old children, greater parent-reported guilt was associated with greater parent-reported reparative behaviors (Colasante et al., 2014). However, the parent-report measures tapped children's general tendencies to experience guilt and use reparative behaviors rather than children's affect and behavior in response to a specific transgression.

One experimental study used a 2×2 design, varying whether or not a mishap caused an experimenter's distress and whether or not children caused the mishap; 3-year-old children engaged in the greatest reparative behaviors when they were at fault for a mishap that caused distress (Vaish, Carpenter, & Tomasello, 2016). Although this finding suggests a potential effect of guilt on reparative behavior, only reparative behavior and not guilt was measured in the study. In sum, consistent with theory, greater guilt was related

to greater use of reparative behaviors in multiple studies of adults and three studies of children. However, the studies of children have exclusively used either parent reports or child reports using hypothetical transgression scenarios.

Reparative Behaviors Alleviate Guilt

In contrast to studies examining guilt as a motivator of reparative behaviors, many fewer studies have tested whether reparative behaviors alleviate guilt. Two studies of autobiographical transgressions have demonstrated that reparative behaviors alleviate guilt in adults (Witvliet, Hinman, Exline, & Brandt, 2011; Witvliet, Ludwig, & Bauer, 2002). Four studies demonstrated guilt reductions by comparing participants' guilt following a hypothetical transgression vignette to their guilt after they imagined engaging in reparative behavior; participants' guilt decreased significantly following imagined attempts to repair the transgressions (de Hooge, 2012). For example, in one study, participants imagined that they let a friend get fired over something they did (Carpenter, Carlisle, & Tsang, 2014). Participants were 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.

Despite this evidence in adults, no study has examined how early in development reparative behaviors serve to effectively alleviate guilt. In one study, toddler children's changes in pupil dilation, a measure of internal arousal, decreased after repairing a transgression and increased if they were unable to repair (Hepach, Vaish, & Tomasello, 2017), suggesting potential internal benefits of repair for transgressors. Although guilt was not measured, guilt, like arousal, may similarly decrease following the use of reparative behaviors.

Overview of the Current Study and Hypotheses

In summary, although empirical studies of adults support theoretical claims that guilt functions to motivate reparative behaviors, which in turn alleviate guilt, very few studies have examined how early in life guilt and reparative behaviors function as hypothesized. The purpose of this study is to examine whether guilt and reparative behaviors function as theorized in middle childhood; this question is particularly important given research demonstrating that children's unalleviated, maladaptive guilt is associated with depression.

This study is the first to use an experimental design to examine causal relations between children's guilt and reparative behaviors. In this study, children were led to believe that they had transgressed, and their guilt and reparative behaviors were measured. This study is also the first to use an *in vivo* transgression in children older than three to examine children's actual, rather than hypothetical, guilt and reparative behaviors. Researchers have reported that the *in vivo* toddler "mishap" paradigms are not effective for children over 45 months old, as they are not convinced that they have transgressed (Kochanska et al., 2002). Thus, we created a novel, developmentally appropriate paradigm that mirrored that of the adult Cryder et al. (2012) study; children were led to believe that a choice they made caused another child to experience either negative or slightly positive emotions. Existing *in vivo* studies of toddlers examine parent-reported or observation-

ally coded guilt, which may not fully capture highly internal guilt feelings and cognitions. The school-age children in our study were able to self-report their own guilt using a scale adapted from Cryder et al. (2012). Finally, as in Cryder et al. (2012), we used a dictator task as a measure of reparative behavior in which children divided resources between themselves and the victim; thus, as is recommended, this task involved a targeted behavior that directly corresponded to the specific transgression.

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 behavior than children who are led to believe that they have caused a peer's slightly positive emotions. The second hypothesis is that among children who are led to believe they have transgressed, children who are given a repair opportunity will report greater reductions in guilt over time relative to children who are not given a repair opportunity.

Method

Participants

Participants were a sample ($N = 97$) of 6- to 10-year old children ($M_{age} = 7.94$ years; $SD = 1.37$ years). The sample size was selected based on power analyses conducted a priori using G*Power 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007). Effect sizes based on prior research in adults (Cryder et al., 2012; de Hooge, 2012) were used. Alpha was set at .05. First, power for a one-way ANOVA of condition (experimental vs. no-guilt) predicting reparative behavior was examined (Hypothesis 1). Previous experimental studies suggest medium to large effect sizes for differences in reparative behavior between guilt and no-guilt control conditions ($d = .42, 1.02, .61$ across three experiments in Cryder et al., 2012). Using an average effect size of $d = .68$, a sample size of 80 was needed for .85 power and 115 for .95 power. Second, power for a mixed ANOVA examining the interaction of time (Time 1, Time 2) and condition on guilt scores was examined (Hypothesis 2). Previous experimental studies reported large effect sizes for differences in guilt between repair and no-repair conditions (e.g., $d = 1.25$ in the experiment reported in de Hooge, 2012, which is most similar to the current study). Using an effect size of $d = 1.25$, a sample size of 28 participants was needed for .95 power. Thus, the analysis that required the largest sample is that of Hypothesis 1, requiring 80–115 participants. Power analyses indicated that across the planned statistical tests a sample size of 97 would yield power of at least .90 to detect the hypothesized effects.

Children participated at a children's museum in a large city in the Southeastern United States. The experimenter (E) approached parents to ask if they were interested in having their children participate in tasks related to child development. The sample included 58 girls (59.2%). Parents reported children's race/ethnicity as follows: 59.8% White, 11.3% Bi/multiracial; 10.3% Black, 7.2% Other, 4.1% Hispanic/Latinx, 3.1% Asian, and 4.1% unreported. Age was collected in years per a museum policy.

Materials

Children chose between two toys that were both bouncy balls, but one was colorful and lit up when bounced and the other was

less colorful, smaller, and did not light up or have any special features. Two images (presented on an iPhone screen, 5.5 in. from corner to corner) each depicted a pretend text message exchange between two experimenters and a photo of a pretend female peer, "Emma." Emma displayed a sad facial expression in the *distress feedback image*, and a slightly happy facial expression in the *satisfied feedback image* (see Figure 1).

Procedure

All procedures were performed in accordance with the ethical standards of Georgia State University and approved by its IRB (protocol name: Children's social learning and decision making). Parents and children provided verbal consent and assent, respectively. Children progressed through several study phases. Children chose between two toys and were led to believe that their choice determined which toy a pretend peer (Emma), who E explained was participating in the study in another room, received. Guilt was induced or not induced: E gave feedback that Emma felt either distressed or satisfied about the toy she received. Children reported their guilt using three measures: two guilt measures served as manipulation checks to ensure that the distress feedback induced guilt, and one measure served as the guilt outcome measure. Next, children engaged in either a reparative behavior task (giving stickers and a note to Emma) or a control task (sorting stickers). Children reported their guilt a second time using the guilt outcome measure. Finally, children reported their reasons for giving stickers as a manipulation check to ensure that giving stickers reflected reparative behavior.

Experimental conditions. Children were assigned to one of three experimental conditions using restricted random assignment to ensure an approximately equal number of girls and boys in each condition. For readability, procedures are described in full as children in each experimental condition experienced them.

In the Experimental Condition, E induced children's guilt, and children were then given the opportunity to give stickers and a note to Emma (i.e., repair their transgression).

Choice phase. Children chose between the two toys, and also chose between two sets of five scratch and sniff stickers; although the stickers were used later during the repair phase, they were given early to establish that the children owned the stickers (C. E. Smith, Blake, & Harris, 2013).

Induction phase. Children in the experimental condition received the distress feedback: E showed the child the distress feedback image and said, "Oh no! I got a text message from my friend who is playing this game with a kid named Emma in the other room. My friend says that you took the last light up ball, and so Emma has to take the small ball. Emma really wanted a light up ball."¹ If the participating child chose the small ball, the script varied slightly. E asked the child how Emma seemed to be feeling in order to direct attention to Emma's emotions and ensure emotion understanding. No child responded incorrectly to this question

¹ Interestingly, after the distress feedback but before the repair opportunity, a small number of children ($n = 5$) offered to give their ball to Emma. When this occurred, E said, "Let's talk about that later." No child brought up wanting to give Emma their ball after the repair phase. We interpreted this as evidence that the children were satisfied with the repair opportunity they were given.

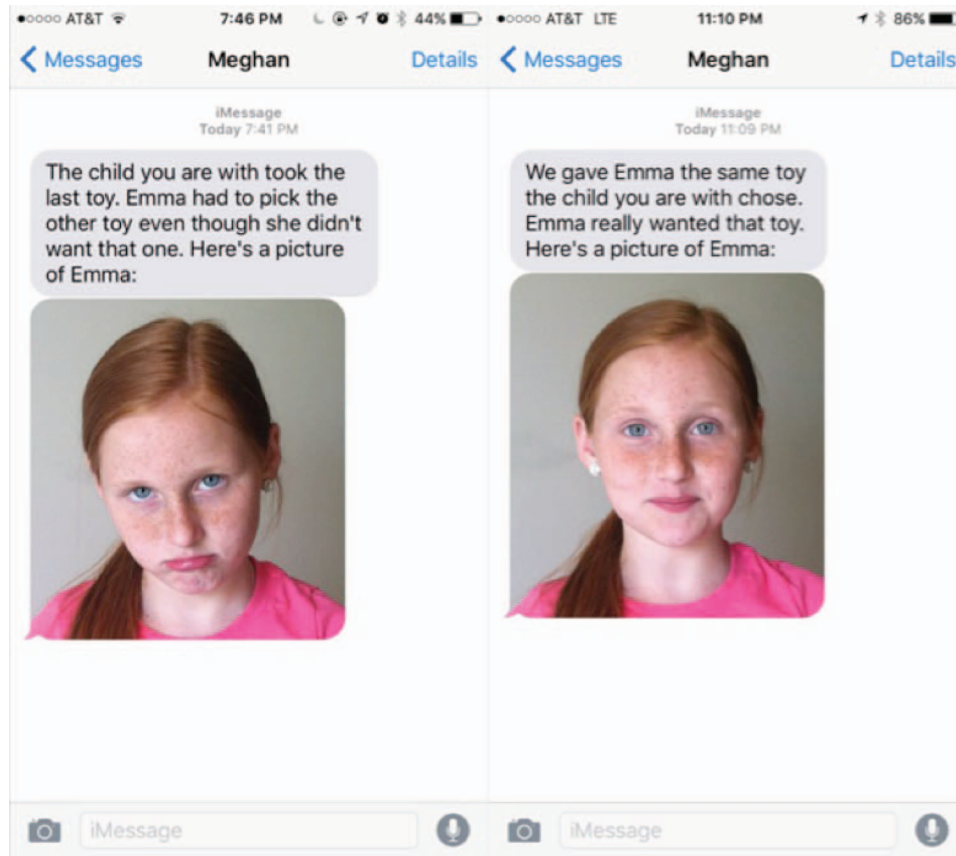


Figure 1. Photos of the distress and satisfied feedback images. The authors received signed consent from the parent and child for the child's likenesses to be published in this article. See the online article for the color version of this figure.

(e.g., all children who received the distress feedback reported that Emma was feeling a negative emotion).

Guilt rating phase 1. Children self-reported their guilt using three measures. Two measures served as manipulation checks. For the fault question, E asked, "Is Emma feeling [sad/happy] because of you, or not because of you?" E used the specific emotion word that the child used to describe Emma's feelings in the induction phase. Children's responses yielded a binary score corresponding to because of you (1) or not because of you (0). Eight children did not respond to this question after repeated prompting and were thus entered as missing data. For the free-response report of emotion, E asked, "How do *you* feel about what happened with Emma?" Children's responses were coded (see Coding and Rating of Data, below).

The Guilt Scale was the guilt outcome measure. E said, "I'm going to ask you more questions about how you feel right now about what happened with Emma." Because a validated measure of children's state guilt does not exist, one was created. As in the child study by Colasante et al. (2014) and adult study by Cryder et al. (2012), we asked children to rate the extent to which they felt "guilty", "bad", and "sorry" on a 5-point scale. Importantly, these three emotion terms are commonly used in validated measures of trait guilt in adults and children. For example, an item from the adult State Shame and Guilt Scale reads, "I feel bad about some-

thing I have done" (SSGS; Marschall, Sanftner, & Tangney, 1994), and an item from the child TOSCA-C reads, "I would feel sorry, very sorry" (Tangney et al., 1990). The fourth item in our scale asked children how "happy" they felt, which was adapted from an item from the adult SSGS, "I feel pleased about something I have done." The term *happy* was used as children demonstrate understanding of happy earlier in development than other positive emotion terms (e.g., proud, excited; Harris, Olthof, Terwogt, & Hardman, 1987). As in the SSGS, this item was reverse-scored.

Because 6- to 10-year-olds have difficulty using Likert scales, with tendencies to respond using extreme endpoints, a double binary response strategy was used (Marsh, Ellis, & Craven, 2002) in which children were asked a series of two yes or no questions. Children were asked, "Do you feel [guilty] about what happened with Emma?" and responded "yes" or "no." Children then answered no *a lot* (1) or no *a little* (2) if they initially responded "no"; or yes *a little* (4) or yes *a lot* (5) if they initially responded "yes." A 3 was given if a child understood the initial question but could not decide how to answer (Marsh et al., 2002). A 3 was given infrequently, on 1.2% of items. This procedure was repeated for "bad," "sorry," and "happy." A guilt composite score was calculated by averaging responses to the four items (Cryder et al., 2012). The order in which E asked about the four emotions was random across participants.

Repair phase. Children then engaged in the Repair task. As in experimental studies of reparative behavior in adults (Cryder et al., 2012), a dictator task was used. Our dictator task mirrored those used in previous studies of children's sharing (C. E. Smith et al., 2013) given that dictator tasks have not been used to measure children's reparative behavior. E told children that they could keep all of their stickers, or share any number with Emma. Children were told to put any stickers for themselves in their bag and any stickers for Emma in a presented envelope. To minimize socially desirable responses, E covered her eyes while the child divided the stickers. Once the child was finished, E asked what the child wanted to say in a note to Emma. E transcribed the child's note on an index card and placed it in the envelope. A research assistant (RA) took the envelope and pretended to deliver it to Emma. The number of stickers given on a 6-point scale (0–5) was the first reparative behavior measure. Prosocial content in the note was rated and used as a second reparative behavior measure (see Coding and Rating of Data).

Guilt rating phase 2. Children rated their guilt a second time using the Guilt Scale described above. Thus, the Guilt Scale was used at two time points, hereafter referred to as Time 1 and Time 2. The Guilt Scale appeared to have good psychometric properties: internal consistency reliability was adequate at Time 1 ($\alpha = .77$) and Time 2 ($\alpha = .79$). Item-total statistics indicated that Cronbach's alpha would be lower if any scale item were deleted. Moreover, the internal consistency reliability was above .70 in each 1-year age group of children at both Time 1 and 2. An unrotated exploratory principal axis factor analysis suggested one factor at both Time 1 and 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 2.

Follow-up phase. Children reported their motivation for giving stickers; E asked, "Why did you share stickers with Emma?" Responses were transcribed and coded (see Coding and Rating of Data).

In the No-Guilt Condition, procedures were identical except that after choosing a toy, children received the Satisfied feedback: E showed the child the satisfied feedback image and said, "Oh! I got a text message from my friend who is playing this game with a kid named Emma in the other room. My friend says because you picked the light up ball, Emma was given a light up ball, too. Emma really wanted a light up ball." Thus, in this condition, guilt was not induced, and then children could give stickers and a note to Emma. The experimental and no-guilt conditions were compared to examine whether children who felt guilty engaged in more reparative behavior than children who did not experience guilt (Hypothesis 1).

In the No-Repair Condition, procedures were identical to that of the experimental condition except that instead of the repair task they completed the Sorting task, a control task that mirrored the repair task. Two envelopes were placed in front of children. E told the child to divide the stickers between the envelopes in any way they wanted. E covered her eyes during the task. Once children finished, they were told to choose an envelope to keep. E then asked the child to write a note to the next child to participate at the testing station. E transcribed the note on an index card and set it aside. E then stated that she wished there was a way for the child to share the remaining envelope with Emma, but no one was

available to deliver it to her. Children in this condition did not complete the follow-up phase (i.e., were not asked why they gave stickers to Emma). Thus, in this condition, E induced guilt, but children were not given a repair opportunity.² The experimental and no-repair conditions were compared to examine whether children reported greater guilt reductions when they had been given a repair opportunity compared to when they had not (Hypothesis 2).

Coding and rating of data. RAs who were blind to study hypotheses coded or rated four variables. One RA served as the master rater and a second RA served as an additional rater. Using pilot data, RAs rated at least five items per variable independently during a training period until interrater reliability (IRR) of at least .75, considered in the excellent range (e.g., Cicchetti, 1994), was attained. The two RAs then independently rated all items for each variable (i.e., the design was fully crossed). Raters met weekly with the PI to review ratings to minimize observer drift (i.e., the implicit change in code definitions made by observers over time; Kazdin, 1977; Smith, 1986). Discrepancies between raters were resolved by retaining the rating of the master rater. IRR for observer drift was calculated.

Prosocial note ratings. The prosocial content of children's notes to Emma was rated as a reparative behavior outcome measure. The presence and sophistication of prosocial statements (e.g., statements of comfort or concern, apologies) were rated on a 4-point scale (0–3). A rating of 0 (*absent*) was given if the note did not contain any prosocial content. A rating of 1 (*minimal*) was given if the note contained a prosocial behavior that reflected an attempt to affiliate with Emma (e.g., "Have fun!") or provided information with the possible intent of improving Emma's mood (e.g., "You will like this game because it has prizes"). A rating of 2 (*moderate*) was given if the note contained a clear prosocial or reparative behavior (e.g., "I wanted to give you two of my stickers," referencing sharing; "I'm sorry") OR referenced Emma's or the child's own emotions (e.g., "I am happy for you that you got what you wanted"; "I am really sad you didn't get what you wanted"). A rating of 3 (*strong*) was given if the note contained a clear prosocial or reparative behavior AND one or more of the following: (a) referenced Emma's or the child's own emotions (e.g., "I am giving you some stickers and I hope that makes you happy"), (b) provided clear reasoning for being prosocial (e.g., "I took the last ball so I'm going to give you some of my stickers"), or (c) was otherwise a standout prosocial statement in terms of sophistication. IRR, assessed with a two-way, mixed, absolute agreement intraclass correlation coefficient (ICC), was .97 for training and 1.00 for drift reliability.

Free-response report of emotion. To assess the experience of empathy, the affective component of guilt (Tilghman-Osborne et al., 2010), RAs coded the free-response emotion that children reported in guilt rating Phase 1. Emotion words were categorized into one of four groups: negative (sad, bad, not good/not happy, sorry), positive (good, happy/glad), neutral (okay), and "I don't know" responses. Each response was coded as only one category. No child reported more than one emotion word, and five children's responses were off-topic and coded as missing. IRR, assessed with a Cohen's Kappa coefficient, was 1.00 for training and drift reliability.

² These children were given a repair opportunity after the paradigm was completed to minimize any distress.

Apology. To assess children's awareness of fault, the cognitive component of guilt (Tilghman-Osborne et al., 2010), children's notes were coded for the presence (1) or absence (0) of an apology (i.e., a statement of sorrow for wrongdoing; in our sample, all apologies included the word *sorry*). IRR, assessed with a Cohen's Kappa coefficient, was 1.00 for training and drift reliability.

Prosocial motivation for giving stickers. To assess whether giving stickers reflected reparative behavior, children's reasons for giving stickers were coded for the presence (1) or absence (0) of three possible data-driven themes reflecting prosocial motivation. All themes present in a transcript were coded. The Mood Theme reflected referencing wanting to improve Emma's mood (e.g., "Because she's sad and I want to make her happy"). The Fault theme reflected referencing fault/responsibility (e.g., "Because I took the last ball"). The Event theme reflected referencing that an event happened to Emma without acknowledging fault (e.g., "Because of what happened to Emma"). IRRs, assessed with Prevalence-Adjusted Bias-Adjusted Kappa coefficients (PABAK; Byrt, Bishop, & Carlin, 1993), ranged from .75 to 1.00 ($M = .91$) for training and .86 to 1.00 ($M = .93$) for drift reliability.

Results

Descriptive Statistics

Group differences among variables. Means, standard deviations, ranges for study variables, and results of tests of equal distributions of demographic variables across experimental conditions are presented in Table 1. Independent samples *t* tests revealed no significant differences in outcome variables by gender or race. Chi-square tests revealed that the proportion of boys and girls and children from each racial group were not significantly different across conditions. A one-way ANOVA revealed no significant difference in age by condition.

Correlations among variables. Children's older age was significantly associated with a greater number of stickers given to the pretend peer (see Table 2). Children's older age was also signifi-

cantly associated with higher prosocial note ratings (i.e., more prosocial notes). The number of stickers and note ratings were moderately and positively but nonsignificantly correlated, suggesting that they are related but distinct measures of reparative prosocial behavior. As a result of these analyses, age was included as a covariate in the analyses examining the effect of condition on the number of stickers given and prosocial note ratings (Hypothesis 1).

Number of stickers given. The majority (71.88%) of children in the experimental condition gave three or more stickers to Emma ($n = 23$). In contrast, a minority (33.33%) of children in the no-guilt condition gave three or more stickers ($n = 11$). Note that although children were not required to give stickers, no child in the experimental condition gave 0 stickers; thus, all children in the experimental condition repaired their transgression.

Manipulation Checks

Manipulation checks were conducted to examine whether the distress feedback induced affective guilt (two checks), cognitive guilt (two checks), and whether the number of stickers given reflected reparative behavior (one check).

Affective guilt (empathy). Guilt Scale Time 1 scores were significantly different among conditions, $F(2, 94) = 56.54, p < .001$, partial $\eta^2 = .55$. There was not a significant difference in guilt between the two conditions that utilized the distress feedback (experimental and no-repair conditions; $M_{diff} = -.10, SE = .20, p = 1.00$). 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 at Time 1 than children in the no-guilt condition.

Children's free-response report of emotion in response to the question that asked "How do you feel about what happened with Emma?" was examined. The number of children with missing data was not significantly different among conditions, $\chi^2(2) = .10, p = .95$. As expected, a chi-square analysis revealed that a significantly greater proportion of children in a condition receiving the distress feedback reported experiencing a negative emotion than children

Table 1

Descriptive Statistics of Primary Study Variables: Means, Standard Deviations, and Ranges or Percentages

Variable	Observed range	Overall sample ($N = 97$)	Mean (<i>SD</i>) or Frequency		
			Experimental condition ($n = 32$)	No-repair condition ($n = 32$)	No-guilt condition ($n = 33$)
1. Age (years)	6–10	7.94 (1.37)	7.88 (1.24)	8.13 (1.24)	7.82 (1.61)
2. Sex (% female)		59.8%	62.5%	62.5%	54.5%
3. Race (% White)		59.8%	56.3%	64.5%	60.6%
4. Children's fault question (% yes)		48.5%	56.3%	51.9%	46.9%
5. Free-response report of emotion (% negative)		61.7%	88.89% ^a	85.19% ^a	11.1% ^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. Number of stickers given	0–5	2.32 (1.13)	2.66 (.90) ^a	N/A	2.00 (1.25) ^b
8. Prosocial note ratings	0–3	1.54 (.96)	2.06 (.93) ^a	N/A	1.58 (.90) ^b
9. Guilt Scale Time 2	1–5	3.09 (1.27)	3.26 (1.14) ^a	4.08 (.67) ^b	1.98 (.91) ^c
10. Prosocial motivation for giving stickers (% present)		54.24%	70.0% ^a	N/A	37.9% ^b
11. Apology (% present)		30.2%	58.1% ^a	N/A	3.1% ^b

Note. Observed ranges are identical to the possible ranges for each variable. Within rows, significant ($p \leq .05$) condition differences in means/percentages are indicated with different superscript letters; for example, a condition with a superscript 'a' is significantly different from a condition with a superscript 'b', and a condition with a superscript 'c' is significantly different from conditions with superscripts 'a' and 'b'.

Table 2
Pearson's Zero-Order Correlations

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. Prosocial note ratings	.43*	.28	.30	.30	

* $p < .05$. ** $p < .01$.

in the no-guilt condition, $\chi^2(1) = 45.92, p < .001$. In sum, results of both a self-report scale and a free-response question support that the distress feedback successfully induced affective guilt.

Cognitive guilt (awareness of responsibility). Children's responses to the fault question asking "is Emma feeling [happy/sad] because of you?" were examined. The number of children with missing data was not significantly different among conditions, $\chi^2(2) = 3.73, p = .16$. Of children in a distress feedback condition, 56% ($n = 32$) responded "yes," indicating responsibility for causing Emma's emotions. Of children in the no-guilt condition, 47% ($n = 15$) indicated that they felt responsible. A chi-square analysis 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. This was expected given that in all conditions children were told that their own toy choice resulted in Emma receiving a particular toy and thus caused her happiness or sadness.

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, $\chi^2(1) = 22.56, p < .001$. In sum, of the 32 children who have both measures of cognitive guilt (i.e., children in the experimental condition), 84.4% ($n = 27$) reported fault on at least one measure. Children who reported fault were more likely to do so through either the fault question or apology measure (66.67%) than both measures (33.33%). Thus, as expected, children in the experimental condition were highly likely to report fault, and tended to do so on one, but not both fault measures.

Prosocial motivation for giving stickers. Children's free responses to the question that asked "Why did you decide to give stickers to Emma?" were examined. A significantly greater proportion of children in the experimental than no-guilt condition endorsed a prosocial theme for giving stickers, $\chi^2(1) = 6.11, p = .01$, indicating that giving stickers was prosocially motivated and is therefore a valid measure of reparative behavior.

Tests of Study Hypotheses³

Children's guilt causes reparative behaviors. Tests of Hypothesis 1 involved using two separate Analyses of Covariance (ANCOVAs) that each included age as a covariate to examine the effect of condition on each of the two measures of reparative behavior. Assumptions of ANCOVA were tested and met for both analyses.

Number of stickers given. There was a significant effect of condition on the number of stickers given, $F(1, 62) = 6.41, p = .01, \eta_p^2 = .09$. Children in the experimental condition (i.e., children who felt guilty) gave significantly more stickers than children in the no-guilt condition ($M_{diff} = .64, SE = .25$). The size of the effect was medium in magnitude, $d = .59$.

Prosocial note ratings. There was a significant effect of condition on the prosocial note ratings, $F(1, 61) = 5.51, p = .02, \eta_p^2 = .08$. 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, $d = .54$.

Children's reparative behaviors alleviate guilt. Tests of Hypothesis 2 involved examining the interaction of time and condition on guilt scores to compare how guilt changed across the two guilt measurements (i.e., before and after the repair phase) in children who transgressed and were given a repair opportunity versus children who transgressed but did not have an opportunity to repair. 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. Four outliers in cells of the design exerted 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 and were removed; results did not differ depending on whether or not outliers were removed. 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. Thus, linear mixed modeling, a statistical alternative to mixed ANOVA that is less sensitive to heterogeneity of variance, was used. A diagonal covariance structure was specified; this variance structure is used for repeated measures data in which variances are assumed to be heterogeneous (West, Welch, & Galecki, 2014).

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 improve model fit (Raudenbush & Bryk, 2002). Model fit was measured using the -2 log-likelihood criteria, which follows a chi-square distribution. Chi-square difference tests that compared the less complex model to the next more complex model were used to determine if the additional parameters significantly improved model fit (West et al., 2014). The best fitting model included condition and time (i.e., two time points) as fixed effects, a Condition \times Time interaction term, and a random intercept (see Table 3).

There was a statistically significant interaction between condition and time on guilt, $F(1, 61) = 14.464, p < .001$ (see Figure 2). 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 time points for the two conditions separately. The models specified were the same as the

³ All analyses were repeated excluding the small number of participants who did not report fault on either fault measure ($n = 5$). Results remained the same in that all tests of hypotheses remained significant and the effect sizes were slightly larger. Results presented in text include these participants.

Table 3
Comparison of Linear Mixed Models for Tests of Hypothesis 2

Model	-2lnL	df	$\Delta\chi^2$	Δdf	p
Model 1: Condition	307.63	3			
Model 2: Condition, Time	324.98	5	17.35	2	<.05
Model 3: Condition, Time, Condition \times Time	332.92	6	7.94	1	<.05
Model 4: Condition, Time, Condition \times Time, Random intercept	343.23	7	10.31	1	<.05
Model 5: Condition, Time, Condition \times Time, Random intercept, Random slope	343.26	9	.03	2	>.05

Note. -2lnL = -2 log-likelihood criteria. The -2lnL for each model was compared to the preceding (less complex) model.

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 only, guilt decreased from the first to second time point, $b = -.88, t(29) = -3.99, p < .001$. In other words, guilt only decreased from the first to second measurement for children who were able to repair their transgression; in contrast, guilt remained high and stable at the second measurement for children who did not engage in reparative behavior. The size of the effect was large in magnitude, $d = 1.05$.

Discussion

The purpose of this study was to investigate functional relations between guilt and reparative behaviors in children using a novel experimental methodology in which children were led to believe they had harmed a peer. Children who were led to believe that they had transgressed to cause a peer's distress (i.e., children who felt guilty) engaged in significantly greater reparative behaviors across two separate measures than children who were led to believe that they had caused a peer's slightly positive emotions. Further, among transgressor children, children who were given the opportunity to repair their transgression reported reduced guilt over time, whereas guilt remained high in children who did not have a repair opportunity.

Several manipulation checks established the validity of the study paradigm. Compared to children who were given feedback that they caused a peer's positive emotions, children who were given feedback that they had transgressed to cause a peer's distress reported significantly greater guilt following the feedback on a rating scale. Moreover, a significantly greater proportion of transgressor children used a negative emotion word when freely describing their emotion after the feedback. Further, the vast majority of transgressor children either self-reported fault for the transgression or apologized in a note to the peer. Thus, evidence across four measures indicated that the paradigm induced the affective and cognitive components of guilt—empathy and understanding of personal responsibility for the transgression; this was critical to demonstrate, as neither component is sufficient to give rise to guilt, and personal responsibility distinguishes guilt from empathy (Tilghman-Osborne et al., 2010). Finally, compared to children who did not transgress, a significantly greater proportion of trans-

gressor children reported a prosocial reason for giving stickers to the peer, supporting that giving stickers was a valid measure of reparative behavior. It was particularly important to establish the validity of the transgression paradigm and reparation measure as the current study was the first to study children's guilt *in vivo* using child report of guilt and a reparative behavior task rather than studying children's hypothetical guilt and repair or relying on parent report. The manipulation checks in this study support the feasibility of experimentally manipulating children's guilt *in vivo* during middle childhood and provides a valid method through which to do so.

Consistent with expectations, children who felt guilty engaged in more reparative behavior by giving more stickers to the peer and writing notes to the peer that were more prosocial than children who did not experience guilt. Older children engaged in greater reparative behavior; this finding adds to a small body of literature documenting increases in reparative behaviors with older age (Colasante et al., 2014) and provides the first evidence of increases in reparative behaviors across middle childhood, specifically. Importantly, the causal relationship between guilt and reparative behavior held even after accounting for age-related increases in reparative behaviors. The prosocial notes of many children were illuminating in that they demonstrated that children often made explicit connections between the transgression, their guilt, and engaging in reparative behavior. A 7-year-old boy wrote, "Dear Emma, I'm sorry I took the glitter ball. Here's something for the apology"; this child made a clear connection between his transgression, feelings sorry, and engaging in reparative behavior to apologize. A 9-year-old girl wrote, "Dear Emma, I didn't want you to feel bad so I gave the things I wanted to you," indicating that this child gave stickers to help Emma alleviate her distress. Our finding that guilt causes children's greater reparative behavior is notable given evidence that reparative behaviors are positively associated with a range of healthy outcomes, such as self-esteem, and negatively associated with poorer outcomes, such as internalizing problems, across children and adults (Bafunno & Camodeca,

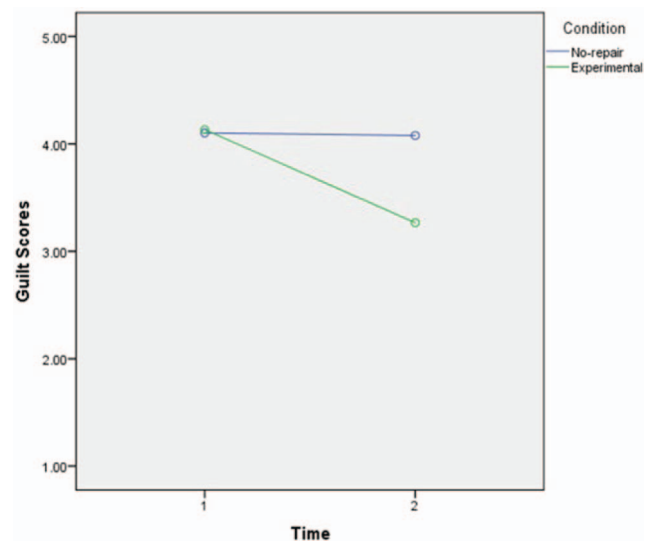


Figure 2. Plot of the interaction of condition and time on children's guilt. See the online article for the color version of this figure.

2013; Cohen, Wolf, Panter, & Insko, 2011). The finding that guilt causes these adaptive behaviors suggests a potential benefit of parents socializing children's healthy guilt following transgressions. Parenting practices such as inductive reasoning have been associated with children's greater guilt (Kochanska, 1991). Our finding may inform the development of parenting interventions that provide psychoeducation on the benefits of children experiencing guilt after transgressing and teach components of effective guilt inductions.

Our finding that guilt causes positive social behaviors lends support to perspectives that emphasize the adaptive potential of acute guilt following a transgression. Previous studies have examined linear relations between guilt and reparative behavior, finding that greater guilt is associated with greater reparative behavior; however, theories also describe that guilt can be maladaptive when it is 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). Our novel paradigm was mild and unlikely to elicit extreme guilt, likely explaining the positive association between guilt and reparative behavior. Future studies should test possible negative quadratic associations between guilt and reparative behavior through examining individual differences in responses to an identical transgression.

Among transgressor children, only children who were given a repair opportunity reported reduced guilt; in contrast, levels of guilt remained high in children who did not have a repair opportunity. Our finding that reparative behaviors alleviate children's guilt coupled with findings that unalleviated, maladaptive guilt is associated with depression in children (Luby et al., 2009) point to a potential benefit of reparative interventions in helping children develop or use these important skills while reparative abilities are still maturing (Muris & Meesters, 2014). These interventions may be particularly beneficial in middle childhood to help prevent increases in depression in adolescence. Although there are several evidence-based prosocial interventions for children (e.g., Caprara, Luengo Kanacri, Zuffianò, Gerbino, & Pastorelli, 2015), these programs typically aim to reduce externalizing problems. Only one known intervention has specifically targeted reparative behaviors. In this study, depressed preschoolers received Parent-Child Interaction Therapy that added an Emotion Development module (PCIT-ED); this module included teaching parents how to help children alter excessive guilt cognitions and engage in reparative behavior (Luby, Barch, Whalen, Tillman, & Freedland, 2018). Compared to children in a wait-list control group, reparative behaviors significantly increased for children in the PCIT-ED group. Future studies should examine the prevention or alleviation of maladaptive guilt as a mechanism through which reparative interventions have an effect on children's psychological outcomes.

Future Directions and Limitations

In this study, we designed a task in which school-age children are led to believe that a victim is distressed, self-report their emotions, and are given an opportunity to use prosocial behaviors

to help the victim. Future studies could modify our design to ask novel questions about children's actual emotions and behaviors in the context of a distressed peer. For example, although both guilt and empathy theoretically motivate prosocial behavior, findings are mixed as to whether children are more prosocial as transgressors or bystanders (e.g., Demetriou & Hay, 2004; Dunn & Brown, 1994). Although this study did not aim to compare children's prosocial behaviors in bystander versus transgressor contexts, this question might illuminate the relative motivational influence of guilt versus empathy on prosocial behavior.

Our 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 internalizing psychopathology. Research has established that reparative behaviors are associated with lower levels of internalizing problems in children (Bafunno & Camodeca, 2013). However, studies are needed to empirically demonstrate the theoretical assumption that reparative behaviors are related to healthy outcomes in part because of their guilt-alleviating function; such evidence would more firmly demonstrate the potential utility of reparative interventions. Studies should track autobiographical transgressions longitudinally or use laboratory transgressions at multiple time points to assess how relations between children's typical guilt, maladaptive guilt, and reparative responses unfold across time; children who display greater reparative tendencies 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, children who experience repeated reparative difficulty and maladaptive guilt may display greater internalizing problems.

Limitations of the study should be noted. For ethical reasons, the transgression in this study was mild, which may have limited the intensity of guilt induced. Although the created Guilt Scale was adapted closely from validated scales and demonstrated good internal consistency, future work on its validity should be done, perhaps by comparing children's scores on the scale to parent-reports of typical guilt. The design of this study necessitated giving children in the experimental condition a repair opportunity; future studies should replicate the current findings using paradigms in which children are not offered an obvious means to repair the transgression. Photos of a White, female peer were used; although participant gender and race were not significantly associated with guilt or reparative behavior, using images of children of varied genders and races would demonstrate the generalizability of functional relations between guilt and reparative behavior. Similarly, the sample was predominantly White; studies with more diverse samples will be important for understanding the generalizability of our results.

Summary

This study presents a novel and valid paradigm through which functional relations between guilt and reparative behaviors in children were examined. We found that guilt motivated children's reparative behaviors, which in turn effectively alleviated their guilt. Our findings may facilitate the investigation of the role of reparative behaviors in determining whether guilt is alleviated, or remains unalleviated and is associated with depression.

References

- 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*, 524–530. <http://dx.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*, 128–143. <http://dx.doi.org/10.1080/17405629.2013.765796>
- 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., Zahn-Waxler, C., & Cole, P. M. (1993). Avoiders vs. amenders: Implications for the investigation of guilt and shame during toddlerhood? *Cognition and Emotion, 7*, 481–505. <http://dx.doi.org/10.1080/02699939308409201>
- Baumeister, R. F., Stillwell, A. M., & Heatherton, T. F. (1994). Guilt: An interpersonal approach. *Psychological Bulletin, 115*, 243–267.
- 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*, 293–318.
- Bybee, J. A., Zigler, E., Berliner, D., & Merisca, R. (1996). Guilt, guilt-evoking events, depression and eating disorders. *Current Psychology, 15*, 113–127. <http://dx.doi.org/10.1007/BF02686944>
- Byrt, T., Bishop, J., & Carlin, J. B. (1993). Bias, prevalence and kappa. *Journal of Clinical Epidemiology, 46*, 423–429. [http://dx.doi.org/10.1016/0895-4356\(93\)90018-V](http://dx.doi.org/10.1016/0895-4356(93)90018-V)
- 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*, 219–237. <http://dx.doi.org/10.1002/per.402>
- Caprara, G. V., Luengo Kanacri, B. P., Zuffianò, 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*, 2211–2229. <http://dx.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*, 389–401. <http://dx.doi.org/10.1080/17439760.2014.910823>
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*, 284–290. <http://dx.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*, 947–966. <http://dx.doi.org/10.1037/a0022641>
- Colasante, T., Zuffianò, 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*, 511–527. <http://dx.doi.org/10.1080/00221325.2014.976535>
- Cryder, C. E., Springer, S., & Morewedge, C. K. (2012). Guilty feelings, targeted actions. *Personality and Social Psychology Bulletin, 38*, 607–618. <http://dx.doi.org/10.1177/0146167211435796>
- de Hooge, I. E. (2012). The exemplary social emotion guilt: Not so relationship-oriented when another person repairs for you. *Cognition and Emotion, 26*, 1189–1207. <http://dx.doi.org/10.1080/02699931.2011.640663>
- Demetriou, H., & Hay, D. F. (2004). Toddlers' reactions to the distress of familiar peers: The importance of context. *Infancy, 6*, 299–318. http://dx.doi.org/10.1207/s15327078in0602_9
- Drummond, J. D. K., Hammond, S. I., Satlof-Bedrick, E., Waugh, W. E., & Brownell, C. A. (2017). Helping the one you hurt: Toddlers' rudimentary guilt, shame, and prosocial behavior after harming another. *Child Development, 88*, 1382–1397. <http://dx.doi.org/10.1111/cdev.12653>
- Dunn, J., & Brown, J. (1994). Affect expression in the family, children's understanding of emotions, and their interactions with others. *Merrill-Palmer Quarterly, 40*, 120–137.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191. <http://dx.doi.org/10.3758/BF03193146>
- Ferguson, T. J. (1991). Children's understanding of guilt and shame. *Child Development, 62*, 827–839.
- Gavazzi, I. G., Ornaghi, V., & Antonietti, C. (2011). Children's and adolescents' narratives of guilt: Antecedents and mentalization. *European Journal of Developmental Psychology, 8*, 311–330. <http://dx.doi.org/10.1080/17405629.2010.491303>
- 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*, 319–343.
- 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*, 84–108. <http://dx.doi.org/10.1046/j.0021-9630.2003.00308.x>
- Hepach, R., Vaish, A., & Tomasello, M. (2017). Children's intrinsic motivation to provide help themselves after accidentally harming others. *Child Development, 88*, 1251–1264. <http://dx.doi.org/10.1111/cdev.12646>
- Jones, W. H., & Kugler, K. (1993). Interpersonal correlates of the Guilt Inventory. *Journal of Personality Assessment, 61*, 246–258. http://dx.doi.org/10.1207/s15327752jpa6102_6
- Kazdin, A. E. (1977). Artifact, bias, and complexity of assessment: The ABCs of reliability. *Journal of Applied Behavior Analysis, 10*, 141–150. <http://dx.doi.org/10.1901/jaba.1977.10-141>
- Kochanska, G. (1991). Socialization and temperament in the development of guilt and conscience. *Child Development, 62*, 1379–1392. <http://dx.doi.org/10.2307/1130813>
- 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*, 852–868. <http://dx.doi.org/10.2307/1131423>
- 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*, 461–482. <http://dx.doi.org/10.1111/1467-8624.00418>
- Lancy, D. F., & Grove, M. A. (2011). Getting noticed: Middle childhood in cross-cultural perspective. *Human Nature, 22*, 281–302. <http://dx.doi.org/10.1007/s12110-011-9117-5>
- 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, NY: Guilford Press.
- Luby, J. L., Barch, D. M., Whalen, D., Tillman, R., & Freedland, K. E. (2018). A randomized controlled trial of parent-child psychotherapy targeting emotion development for early childhood depression. *American Journal of Psychiatry, 175*, 1102–1110. <http://dx.doi.org/10.1176/appi.ajp.2018.1803.0321>
- Luby, J., 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*, 1156–1166. <http://dx.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*, 1373–1387. [http://dx.doi.org/10.1016/S0191-8869\(02\)00197-6](http://dx.doi.org/10.1016/S0191-8869(02)00197-6)
- Marschall, D., Sanftner, J., & Tangney, J. P. (1994). *The State Shame and Guilt Scale*. 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*, 376–393. <http://dx.doi.org/10.1037/0012-1649.38.3.376>
- 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 and Family Psychology Review, 17*, 19–40. <http://dx.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*, e9. <http://dx.doi.org/10.1017/jrr.2013.9>
- Quiles, Z. N., & Bybee, J. (1997). Chronic and predispositional guilt: Relations to mental health, prosocial behavior, and religiosity. *Journal of Personality Assessment, 69*, 104–126. http://dx.doi.org/10.1207/s15327752jpa6901_6
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage.
- 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*, 751–772. <http://dx.doi.org/10.1177/0265407513503595>
- Silfver, M. (2007). Coping with guilt and shame: A narrative approach. *Journal of Moral Education, 36*, 169–183. <http://dx.doi.org/10.1080/03057240701325274>
- 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*, e59510. <http://dx.doi.org/10.1371/journal.pone.0059510>
- Smith, G. A. (1986). Observer drift: A drifting definition. *The Behavior Analyst, 9*, 127–128. <http://dx.doi.org/10.1007/BF03391937>
- Tangney, J. P., Stuewig, J., & Mashek, D. J. (2007). Moral emotions and moral behavior. *Annual Review of Psychology, 58*, 345–372. <http://dx.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*, 797–809. <http://dx.doi.org/10.1037/0022-3514.70.4.797>
- Thapar, A., Collishaw, S., Pine, D. S., & Thapar, A. K. (2012). Depression in adolescence. *The Lancet, 379*, 1056–1067. [http://dx.doi.org/10.1016/S0140-6736\(11\)60871-4](http://dx.doi.org/10.1016/S0140-6736(11)60871-4)
- 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*, 536–546. <http://dx.doi.org/10.1016/j.cpr.2010.03.007>
- 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*, 607–620. <http://dx.doi.org/10.1007/s10802-011-9591-6>
- Tracy, J. L., & Robins, R. W. (2004). Putting the self into self-conscious emotions: A theoretical model. *Psychological Inquiry, 15*, 103–125. http://dx.doi.org/10.1207/s15327965pli1502_01
- Vaish, A., Carpenter, M., & Tomasello, M. (2016). The early emergence of guilt-motivated prosocial behavior. *Child Development, 87*, 1772–1782. <http://dx.doi.org/10.1111/cdev.12628>
- West, B. T., Welch, K. B., & Galecki, A. T. (2014). *Linear mixed models: A practical guide using statistical software* (2nd ed.). Boca Raton, FL: CRC Press. <http://dx.doi.org/10.1201/b17198>
- Witvliet, C. v., 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*, 223–238.
- Witvliet, C. v., 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*, 219–233.

Received November 12, 2018

Revision received June 20, 2019

Accepted June 26, 2019 ■