Positive evaluative conditioning effects on harsh parenting behaviors and theorized antecedents: two experiments

Michael Francis Wagner
A social information-processing model of child physical abuse predicts that reducing negative evaluations about children should also reduce situational-based attributions of hostile intent, anger, and harsh and harmful parenting behaviors. Previous research provided evidence that use of a positive evaluative conditioning (EC) procedure is effective at reducing these situationally based variables. The purpose of the two studies reported in this dissertation is to explore possible mechanisms that can explain the effects that positive EC has on these variables.

Study 1 \((N = 77)\) employed three EC procedures (forward positive, backward positive, and forward pseudo) to examine how well three theories of EC (referential, propositional, and affective implicit misattribution) explain previous research results. Study 2 \((N = 50)\) tested whether previous research results could be explained by semantic learning (i.e., that children have positive traits). In contrast to Study 1 and previous research, Study 2 employed stimuli (emojis) that were relatively devoid of semantic content. If effects were observed in Study 2, semantic learning could be ruled out as the sole mechanism underlying effects of positive EC.
Results of Study 1 showed that backward positive EC (generally) produced many of the same effects that the forward positive EC produced. Based on Study 1 results, it is concluded that referential and propositional processes mediated EC effects in forward positive EC, whereas propositional and affective implicit misattribution processes mediated EC effects in the backward positive EC. Results of Study 2 (using emojis instead of trait words) showed that positive EC produced many of the same positive EC effects observed in Study 1 and previous research. Based on results from Study 2, it is concluded that semantic learning cannot fully account for the effects observed in Study 1 and previous research.
NORTHERN ILLINOIS UNIVERSITY
DE KALB, ILLINOIS

AUGUST 2018

POSITIVE EVALUATIVE CONDITIONING EFFECTS ON HARSH PARENTING BEHAVIORS AND THEORIZED ANTECEDENTS:
TWO EXPERIMENTS

BY
MICHAEL FRANCIS WAGNER
© 2018 Michael Francis Wagner

A DISSERTATION SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE
DOCTOR OF PHILOSOPHY

DEPARTMENT OF PSYCHOLOGY

Dissertation Co-Directors:
Joel S. Milner and John J. Skowronski
I am indebted to and grateful for the NIU Center for the Study of Family Violence and Sexual Assault, NIU Department of Psychology, and my helpful and caring committee: Drs. Joel S. Milner, John J. Skowronski, Julie L. Crouch, Randy J. McCarthy, Brad J. Sagarin, and Amanda M. Durik. I especially thank Joel and John for their extraordinary and dedicated mentorship and support throughout my tenure as a graduate student. I also wish to thank Shelby Malone, who provided helpful and professional data collection assistance for parts of my dissertation research.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF APPENDICES</th>
<th>v</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Interventions That Target Attributions of Hostile Intent to Reduce Aggression in Adolescents</td>
<td>3</td>
</tr>
<tr>
<td>Interventions That Target Beliefs About the Malleability of Personality in Order to Reduce Attributions of Hostile Intent in Adolescents</td>
<td>11</td>
</tr>
<tr>
<td>Social Cognitive Models of Harsh Parenting and Related Interventions That Target Attributions of Hostile Intent to Reduce Parental Harsh Discipline and Harm</td>
<td>17</td>
</tr>
<tr>
<td>The Social Information Processing Model of Child Physical Abuse</td>
<td>26</td>
</tr>
<tr>
<td>Applying Positive Evaluative Conditioning (EC) in the Context of the SIP Model of CPA</td>
<td>35</td>
</tr>
<tr>
<td>2. PRELIMINARY STUDY</td>
<td>47</td>
</tr>
<tr>
<td>Method</td>
<td>48</td>
</tr>
<tr>
<td>Results</td>
<td>55</td>
</tr>
<tr>
<td>Discussion</td>
<td>57</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td>64</td>
</tr>
<tr>
<td>Overview of Dissertation Studies</td>
<td>65</td>
</tr>
<tr>
<td>Chapter</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>3. STUDY 1</td>
<td>66</td>
</tr>
<tr>
<td>Method</td>
<td>69</td>
</tr>
<tr>
<td>Analytic Strategy</td>
<td>73</td>
</tr>
<tr>
<td>Results</td>
<td>82</td>
</tr>
<tr>
<td>Discussion</td>
<td>106</td>
</tr>
<tr>
<td>4. STUDY 2</td>
<td>111</td>
</tr>
<tr>
<td>Method</td>
<td>112</td>
</tr>
<tr>
<td>Analytic Strategy</td>
<td>114</td>
</tr>
<tr>
<td>Results</td>
<td>115</td>
</tr>
<tr>
<td>Discussion</td>
<td>133</td>
</tr>
<tr>
<td>5. GENERAL LIMITATIONS AND SUMMARY</td>
<td>141</td>
</tr>
<tr>
<td>Summary</td>
<td>143</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>144</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>157</td>
</tr>
</tbody>
</table>
# LIST OF APPENDICES

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. SAMPLE LAYOUT OF STUDY MEASURES</td>
<td>157</td>
</tr>
<tr>
<td>B. AMBIGUOUS CHILD PHOTOGRAPHS</td>
<td>163</td>
</tr>
<tr>
<td>C. AMBIGUOUS CHILD BEHAVIORS</td>
<td>165</td>
</tr>
<tr>
<td>D. SAMPLE VOODOO DOLL TASK</td>
<td>167</td>
</tr>
<tr>
<td>E. INFORMED CONSENT DOCUMENT USED IN PRELIMINARY STUDY</td>
<td>169</td>
</tr>
<tr>
<td>F. DEBRIEFING DOCUMENT USED IN PRELIMINARY STUDY</td>
<td>171</td>
</tr>
<tr>
<td>G. SAMPLE EMOJIS USED IN STUDY 2</td>
<td>173</td>
</tr>
<tr>
<td>H. STUDY 1 DEPENDENT VARIABLE MEANS (SDS) BY EXPERIMENTAL CELL</td>
<td>176</td>
</tr>
<tr>
<td>I. STUDY 1 DEPENDENT VARIABLE CHANGE SCORE MEANS (SDS) BY EVALUATIVE CONDITIONING (EC) CONDITION</td>
<td>178</td>
</tr>
<tr>
<td>J. LINEAR RELATIONSHIPS BETWEEN PRE-EC SCORES AND EC CHANGE SCORES WITHIN THE FORWARD POSITIVE EC CONDITION AND THE BACKWARD POSITIVE EC CONDITION</td>
<td>180</td>
</tr>
<tr>
<td>K. STUDY 1 RESULTS OF MULTI-LEVEL MODEL ANALYSES FOR POSITIVE CHILD EVALUATIONS</td>
<td>182</td>
</tr>
<tr>
<td>L. STUDY 1 RESULTS OF MULTI-LEVEL MODEL ANALYSES FOR NEGATIVE CHILD EVALUATIONS</td>
<td>184</td>
</tr>
<tr>
<td>Appendix</td>
<td>Study 1 Results of Multi-Level Model Analyses for Likings</td>
</tr>
<tr>
<td>----------</td>
<td>--------------------------------------------------------</td>
</tr>
<tr>
<td>M.</td>
<td>Study 1 Results of Multi-Level Model Analyses for Future Discipline Expectations</td>
</tr>
<tr>
<td>N.</td>
<td>Study 1 Results of Multi-Level Model Analyses for Attributions of Hostile Intent</td>
</tr>
<tr>
<td>O.</td>
<td>Study 1 Results of Multi-Level Model Analyses for Anger</td>
</tr>
<tr>
<td>P.</td>
<td>Study 1 Results of Multi-Level Model Analyses for Harsh Verbal Discipline</td>
</tr>
<tr>
<td>Q.</td>
<td>Study 1 Results of Multi-Level Model Analyses for Harsh Physical Discipline</td>
</tr>
<tr>
<td>R.</td>
<td>Study 1 Results of Multi-Level Model Analyses for Doing Nothing</td>
</tr>
<tr>
<td>S.</td>
<td>Study 1 Results of Multi-Level Model Analyses for Positive Child Evaluations</td>
</tr>
<tr>
<td>T.</td>
<td>Study 2 Dependent Variable Means (SDS) by Experimental Cell</td>
</tr>
<tr>
<td>U.</td>
<td>Study 2 Dependent Variable Change Score Means (SDS) by Evaluative Conditioning (EC) Condition</td>
</tr>
<tr>
<td>V.</td>
<td>Linear Relationships Between Pre-EC Scores and EC Change Scores Within the Positive EC Condition</td>
</tr>
<tr>
<td>W.</td>
<td>Study 2 Results of Multi-Level Model Analyses for Negative Child Evaluations</td>
</tr>
<tr>
<td>X.</td>
<td>Study 2 Results of Multi-Level Model Analyses for Likings</td>
</tr>
<tr>
<td>Y.</td>
<td>Study 2 Results of Multi-Level Model Analyses for Future Discipline Expectations</td>
</tr>
<tr>
<td>Z.</td>
<td>Study 2 Results of Multi-Level Model Analyses for Future Discipline Expectations</td>
</tr>
<tr>
<td>Appendix</td>
<td>STUDY 2 RESULTS OF MULTI-LEVEL MODEL ANALYSES FOR</td>
</tr>
<tr>
<td>----------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>AA.</td>
<td>ATTRIBUTIONS OF HOSTILE INTENT</td>
</tr>
<tr>
<td>BB.</td>
<td>ANGER</td>
</tr>
<tr>
<td>CC.</td>
<td>HARSH VERBAL DISCIPLINE</td>
</tr>
<tr>
<td>DD.</td>
<td>HARSH PHYSICAL DISCIPLINE</td>
</tr>
<tr>
<td>EE.</td>
<td>DOING NOTHING</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION

Theoretical models of human aggression and violence generally include cognitive, affective, and physiological variables as predictors of aggression. The General Aggression Model (Anderson & Bushman, 2002; DeWall, Anderson, & Bushman, 2011) posits that person factors (e.g., high, but unstable self-esteem; Bushman & Baumeister, 1998) and situational inputs (e.g., provocation, frustration, and aggressive cues such as the presence of guns) influence aggressive behavior. Person factors and situational inputs increase aggressive behavior by generating internal states. Internal states include hostile thoughts and scripts, affect (e.g., anger), and arousal. Finally, internal states influence the extent and type of appraisal and decision processes that may contribute to possible actions (i.e., thoughtful actions or relatively impulsive actions, and both types of action have potential to manifest as aggression).

In the domain of parent-child interactions, social-cognitive information processing models of child physical abuse (CPA), such as the Social Information Processing (SIP) model of CPA (Crouch & Milner, 2005; Milner, 1993, 2000, 2003) assert that parental attributions of hostile intent to child behaviors are associated with harsh discipline and aggressive parenting behaviors, including CPA. The SIP model of CPA also posits that negative child-related schemas, which include negative child-related evaluations (i.e., evaluations of a child or children as inherently negative or bad), predispose abusive and high-risk parents to attribute hostile intent
to child behaviors, especially when the child behaviors are ambiguous with regard to child intent. These attributions of hostile intent increase the likelihood that parents engage in harsh discipline, CPA, and aggressive parenting behaviors.

Interventions that target theorized causal antecedents of CPA and parental aggression toward children to reduce perpetration are desirable, given that CPA is associated with a myriad of negative outcomes for the child. Examples of negative outcomes of CPA include increased risk for poor health (Kempe et al., 1985; Zelenko et al., 2001), depression (Kazdin, Moser, Colbus, & Bell, 1985; Straus & Kantor, 1994), low self-esteem (e.g., Egeland, Sroufe, & Erickson, 1983), Posttraumatic Stress Disorder (e.g., Ackerman, Newton, McPherson, Jones, & Dykman, 1998; Widom, 1999; Ystgaard, Hestetun, Loeb, & Mehlum, 2004), emotion regulation difficulties (e.g., Slep & O’Leary, 2007), Borderline Personality Disorder (Hallquist, Hipwell, & Stepp, 2015; Herman, Perry, & van der Kolk, 1989; Zanarini et al., 1997), insecure attachment (Egeland & Sroufe, 1981), hostile attribution style (Dodge, Bates, & Pettit, 1990), aggression (George & Main, 1979; Main & George, 1985), and death (Daro & McCurdy, 1991; Kempe et al., 1985). Given these documented outcomes, collectively the literature suggests intervention on putative antecedents of CPA, such as negative child-related attitudes and tendency to attribute hostile intent to child behavior, is a worthwhile endeavor. To begin to support the claim that reducing attributions of hostile intent directed to children in parents will reduce parent-to-child aggression, including CPA, intervention studies that targeted attributions of hostile intent will be reviewed.
Interventions That Target Attributions of Hostile Intent to Reduce Aggression in Adolescents

One early body of research (e.g., Guerra & Slaby, 1990; Pepler, King, & Byrd, 1991) that documented effects of social-cognitive interventions designed to reduce general aggression (i.e., aggression not specific to the parenting domain) in response to provocation included studies that used samples of aggressive adolescents or children. Some interventions were multi-faceted (i.e., targeted multiple theorized social-cognitive antecedent variables to aggression; Guerra & Slaby, 1990; Pepler et al., 1991). For example, in multi-faceted social-cognitive interventions to reduce aggression in adolescents, beliefs about the legitimacy or appropriateness of aggression in response to perceived provocation were targeted along with attributions of hostile intent in response to ambiguous behaviors (e.g., Guerra & Slaby, 1990; Pepler et al., 1991; also see Wilson & Lipsey, 2006, for a meta-analysis of general social-cognitive intervention effects on aggression in school children).

In the Wilson and Lipsey (2006) meta-analyses, the overall weighted standardized mean difference ($d$) between the treatment and control groups on measures of aggressive or disruptive behavior in universal (i.e., applied to all students in the school) social-cognitive interventions to reduce aggressive behavior was $0.26$, $p < .001$, $k = 47$. Thus, overall, the treatment groups, relative to the control groups, exhibited $0.26$ fewer standard deviation units of aggressive behavior.

For example, Guerra and Slaby (1990) conducted an intervention study designed, in part, to reduce CPA. Guerra and Slaby randomly assigned (balanced by gender) a group ($N = 126$) of
incarcerated, antisocial adolescents to one of three treatment groups: a cognitive mediation training group, an “attention” control group, and a no-treatment control group. Prior to the study, the sample was screened to ensure that the participants (a) were 15-18 years old, (b) had a Grade 6 or higher reading level, (c) were not learning handicapped, (d) had a parole date that was more than six months away, and (e) were incarcerated for antisocial or violent offenses. Among other components (i.e., modification of beliefs that aggression is a legitimate and acceptable behavior; increasing self-control) that are not the foci of the present dissertation, the cognitive mediation training included a component that addressed social-cognitive skills that facilitate making benign attributions to others’ involved in a “social problem.”

Guerra and Slaby (1990) defined a social problem as an ambiguous scenario involving two individuals in which one individual performed an action that led to a negative outcome for another individual. In response to a hypothetical ambiguous social problem, adolescents might infer that the individual who performed the action that led to a negative outcome for another individual performed such action with either hostile intent, benign intent, or no/irrelevant intent (i.e., accidental). The social-cognitive skills in the cognitive mediation training included (a) attending to relevant, non-hostile cues when defining social problems and setting goals commensurate to the definition of the social problem (i.e., defining an ambiguous social situation [e.g., child bumping into you in a hallway] in benign instead of hostile terms and setting goals that are congruent with the definition of the ambiguous social problem as non-hostile); (b) seeking information that might serve as evidence that a social scenario did not involve hostile intent; (c) generating a variety of potential responses and consequences to the defined social
problem, some of which might include benign responses; and (d) prioritizing response options that increase the likelihood in producing a non-aggressive response.

The attention control group was matched to the cognitive mediation skills training group with respect to time in sessions but was taught basic skills (i.e., math and reading) that were thought to be relatively unrelated to social-cognitive antecedents to aggression. Both the cognitive mediation training group and the attentional control group were treated in small discussion groups each containing from 10 to 14 youths; four discussion groups were in each treatment condition. The discussion groups met for a 12-week period in one-hour sessions once per week. Finally, the no-treatment control group did not receive any form of treatment.

Results of the Guerra and Slaby (1990) study indicated that the cognitive mediation skills training, relative to the attention control group and no-treatment control group, significantly decreased the frequency with which participants defined ambiguous social problems in hostile terms and significantly decreased participant selection of goals to behave in a hostile manner in response to ambiguous social problems. The cognitive mediation skills training, relative to the attention control group and no-treatment control group, also significantly influenced participants’ responses in the following ways: (a) increased the number of situational facts that the adolescents asked to receive about the ambiguous social problem, (b) increased the number of generated potential solutions to the ambiguous social problem, (c) increased the number of generated potential consequences of the ambiguous social problem, (d) changed the best solution chosen by the adolescents so that the best solution was less aggressive, and (e) changed the second best solution chosen by the adolescents so that the second best solution was less aggressive. Moreover, within the cognitive mediation skills training group, aggressive behavior significantly
decreased from pre- to post-treatment. Aggressive behavior was measured pre- and post-treatment by computing a mean of ratings provided by the adolescents’ supervisors (who were blind to treatment group) using an 18-item behavior checklist. For each item on the checklist, supervisors were asked to indicate (on a five-point scale) how frequently, based on the respondents’ experience with the youth, each adolescent engaged in the behaviors described (e.g., “pushes or hits other youth”). Insufficient information was available in Guerra and Slaby to report effect sizes for analyses.

Although multi-faceted, social-cognitive interventions to reduce aggression have been shown to be effective, one limitation is inherent in the multi-faceted property of the interventions. Specifically, as other authors (e.g., Hudley & Graham, 1993) have pointed out, in studies of multi-faceted interventions it is difficult to ascertain which components of the intervention explain observed reductions in aggressive behavior. For example, although the intervention used in Guerra and Slaby (1990) was effective, it is not clear whether the changes observed in the study were due to the social-cognitive skills training, the beliefs-in-aggression-as-legitimate component, or the self-control component. Thus, in contrast to multi-faceted interventions, subsequent interventions were tailored to alter specific theorized antecedents to aggression. For example, Hudley and Graham (1993) designed an intervention to reduce aggression in African American elementary school boys only targeting the *hostile attribution bias* (i.e., the tendency to attribute hostile intent to the ambiguous actions of others).

Specifically, Hudley and Graham (1993) randomly assigned 72 aggressive African American boys and 36 non-aggressive African American boys to one of three treatment conditions: attributional retraining, attention training, and no treatment. To be classified as
aggressive, boys had to meet three requirements. One requirement was that the boy scored above the median on perceived aggression as measured by responses on the Teacher Checklist (Coie & Dodge, 1988). Specifically, teachers completed the eight-item Teacher Checklist aggression scale (example item: “This child starts fights”). Responses to the eight items were made on a scale ranging from 1 (*not at all like this child*) to 5 (*very much like this child*). Responses to the items were summed.

The second requirement to be classified as aggressive was being strongly disliked, according to reported social preferences, by their peers. The third requirement was that boys had to be nominated as aggressive by their peers. Measures to assess whether boys met the second and third requirements were administered during sociometric interviews in the schools. The sociometric interviews solicited responses to items that asked the boys “…to write down the names of the three students in their class whom they liked most, the three whom they liked least, and the three who best fit each of five behavior descriptions” (Hudley & Graham, 1993, p. 126). For each child, a social preference score was computed by subtracting the total number of “liked least” nominations from the total number of “liked most” nominations. A boy met the second requirement if the social preference score was less than zero.

Responses to the five behavior descriptions collected during the sociometric interview were used to determine whether boys met the third requirement. Specifically, three of the behavior descriptions were aggressive (i.e., starts fights, has a very short temper, disrupts the group), and two of the behavior descriptions were prosocial (i.e., works well with others, is helpful to other students). Peer-nominated aggression was determined by the ratio of the number of aggressive behavior descriptions to the number of prosocial behavior descriptions. If a boy’s
peer-nominated aggression score was higher than 2, the boy met the third requirement to be classified as aggressive.

In the first part of the attributional retraining intervention, boys were taught to accurately detect intentionality during ambiguous social interactions (criteria for accurate detection of intentionality were not described in Hudley & Graham, 1993). Techniques used included role play and discussion of personal experiences. During these techniques, the boys were taught “…to search for, interpret, and properly categorize the verbal and behavioral cues emitted by others in social dilemmas…” (Hudley & Graham, 1993, p. 127). Specific activities in the first part of the attributional intervention included training in interpreting intent based on facial expressions and producing videos that demonstrated ability to distinguish between prosocial, accidental, hostile, and ambiguous intent. In the second part of the intervention, boys were taught skills to increase the likelihood that the boys would attribute non-hostile intent to negative social situations. Specific activities in the second part of the intervention included brainstorming and categorizing a list of potential causes for the negative social situations. Finally, in the third part of the intervention boys were taught the meaning and implication of attributing hostile intent to ambiguous social situations and when to respond non-aggressively. Specific activities in the third part of the intervention included teaching the boys decision rules that would increase the likelihood of responding non-aggressively when not enough information is available to attribute intent to the peer in the social situation.

To assess the impact of their intervention, Hudley and Graham (1993) asked participants in each of their treatment groups to complete pre- and post-treatment questionnaires that contained five hypothetical social situations. Each questionnaire presented hypothetical social
situations that involved negative outcomes for the respondent (e.g., getting pushed down) that were caused by a hypothetical peer. The five hypothetical social situations varied by the type of cues that would lead to an attribution of different intents (i.e., prosocial, accidental, or hostile). After each situation, the boys completed items assessing the extent to which the hypothetical peer enacted the behavior with intent, the extent to which the boys blamed the hypothetical peer and felt angry, and the boys’ behavioral response to the hypothetical ambiguous social situation. With respect to the latter behavioral response item, boys indicated a behavioral response on a scale ranging from prosocial action (“do something nice for this other kid”) to directly hostile behavior (i.e., “have it out right then and there”). The pre-treatment questionnaire was administered approximately one month prior to the treatment, and the post-treatment questionnaire was administered within two weeks of completion of the treatment.

Approximately one month after treatment, aggressive boys participated in a laboratory analog task to measure aggression. The laboratory task involved a non-aggressive peer boy giving directions to the aggressive boy on how to reach a destination in a grid map depicting streets and buildings. The directions provided by the experimenter to the non-aggressive boy were incorrect, so the aggressive boy who received and followed the directions would never reach the destination. During the interaction, the aggressive boy’s aggressive verbal communications to the non-aggressive direction giver were recorded as a measure of verbal aggression. At the conclusion of the task, the aggressive boy completed items assessing (a) the aggressive boy’s attribution of the non-aggressive boy’s intent and (b) the aggressive boy’s anger. Finally, formal disciplinary referrals at the boys’ schools were examined at pre- and post-intervention.
Hudley and Graham (1993) found that the attributional intervention, but not the attention control or the no-treatment control treatments, significantly reduced aggressive boys’ judgments of the extent to which the hypothetical peer enacted the behaviors in the ambiguous social situation with hostile intent, the extent of the aggressive boys’ anger, and the aggressive boys’ aggressive behavior. As hypothesized, these results were qualified, such that the previously described pattern of results was observed for only hypothetical social situations in which cues of peer intent were ambiguous. Insufficient information in Hudley and Graham precluded calculating effect sizes for observed effects.

Although results of Hudley and Graham (1993) provided evidence supporting the effectiveness of the attributional retraining intervention, interventions to reduce attributions of hostile intent and aggressive behavior in schoolchildren have not been unequivocally successful. For example, a subsequent study (Hudley et al., 1998) used the attributional retraining intervention in Hudley and Graham re-labeled as “The BrainPower Program.” In Hudley et al. (1998), 256 aggressive boys and 128 non-aggressive boys were randomly assigned to receive an attributional retraining treatment (i.e., “The BrainPower Program”), an attention control treatment, or no treatment. Except for (a) the inclusion of 6-month and 12-month follow-up measurement points, and (b) the use of a different measure of aggression (described forthwith), the procedures and measures in Hudley et al. were the same as the procedures and measures used in Hudley and Graham. Pre- and post-treatment, Hudley et al. (1998) measured boys’ aggression using teachers’ responses in a questionnaire asking about students’ aggressive behaviors. Item examples were, “controls temper in conflict situations with peers,” “responds appropriately to teasing by peers,” and “responds appropriately when pushed or hit by other children.” Teachers’
responses to these items were made on a scale ranging from 0 (never) to 2 (very often). Summation of responses to these items served as a measure of the boys’ aggressive behaviors.

Results of Hudley et al. (1998) replicated Hudley and Graham’s (1993) reductions in attributions of hostile intent and aggression. That is, the attributional retraining treatment, but not the other treatment conditions, significantly reduced attributions of hostile intent and aggression from pre-treatment to post-treatment (i.e., at the conclusion of the 12-week program). However, the Hudley et al. analyses did not reveal significant differences between treatment conditions at the 12-month follow-up either for attributions of hostile intent or aggressive behavior. The absence of differences between treatment groups at the 12-month follow-up suggests that the benefits of the attributional retraining program were transient. Nonetheless, the previously described studies, which using adolescent or child samples, provide some empirical support for the contention that aggression can be reduced, at least in the short term, by reducing perceiver attributions of hostile intent in ambiguous social scenarios.

Interventions That Target Beliefs About the Malleability of Personality in Order to Reduce Attributions of Hostile Intent in Adolescents

Researchers have investigated the effectiveness of an intervention targeting adolescents’ beliefs about the malleability (or fixedness) of others’ basic personality traits to reduce attributions of hostile intent in response to hypothetical ambiguous peer scenarios (Yeager et al., 2013). Such interventions are based on the hypothesis that teaching that others’ basic personality traits are malleable (incremental mindset) rather than fixed (entity mindset) reduces attributions
of hostile intent. This hypothesis is derived from theory and research that suggests that having more of an entity, relative to incremental, mindset is associated with greater likelihood of inferring that behaviors are caused by underlying trait dispositions rather than situational factors (Chiu, Hong, & Dweck, 1997; Hong, Chiu, Dweck, & Sacks, 1997; Levy & Dweck, 1999).

To examine the empirical support for the idea that having an entity mindset is linked to attributions of hostile intent in adolescents, Yeager et al. (2013) meta-analyzed results of 11 studies that examined the relationship between entity mindset and attributions of hostile intent. In each study, entity mindset was measured using five items that had been developed in previous research (Yeager, Trzesniewski, Tirri, Nokelainen, & Dweck, 2011). The items targeted personality traits relevant to bullying and peer victimization in high school. For example, items included, “bullies and victims are types of people that really can’t be changed” and “everyone is either a winner or a loser in life.” Responses to the items were made on a scale ranging from 1 (strongly disagree) to 6 (strongly agree), and means of the items were used as a continuous predictor of attributions of hostile intent. In the 11 studies attributions of hostile intent were measured using one of two methods. One method was to examine attributions of hostile intent in response to a hypothetical peer scenario as described by Dodge (2006). The hypothetical peer scenario was as follows:

Imagine that you were walking in a crowded hallway at school and everybody was rushing to get to the next class so they wouldn’t be late. While you were looking the other way, you and another student bumped into each other (pretty hard), so it hurt your shoulder and you dropped the books that you were carrying. The other student paused briefly, looked at you quickly, and then turned away and hurried to class. (Yeager et al., 2013, p. 5)

After reading the scenario, students responded to items that measured attributions of hostile intent. Among the seven studies included in the meta-analysis, the wording and response
format for the items measuring attributions of hostile intent varied. The variants were minor (i.e., the items were face-valid, self-report measures of attributions of hostile intent). An example of an item used to measure attributions of hostile intent in the vignette studies was, “They were being mean to me on purpose,” which was followed by a response scale that ranged from 1 (*not at all*) to 7 (*an extreme amount*). Responses on the items that measured attributions of hostile intent were always averaged within each study to form indices of attributions of hostile intent.

In the studies included in the Yeager et al. (2013) meta-analysis, the second method used to measure attributions of hostile intent involved a laboratory task that commonly elicits reactive aggression. Specifically, some studies in the meta-analysis used a Cyberball paradigm (Williams, 2009) to simulate an experience of social exclusion. In the Cyberball paradigm, student participants were led to believe that they would be playing an online game of catch with two other participants; in reality the game was pre-programmed. At the outset, student participants were thrown the ball twice and then were not thrown the ball again for the remainder of the Cyberball paradigm. Yeager et al. did not report details regarding the duration or number of tosses in the Cyberball paradigm; however, Yeager et al. reported that a standardized version of the Cyberball paradigm was used. Meta-analyses (Hartgerink, van Beest, Wicherts, & Williams, 2015) have shown that effects of Cyberball include intrapersonal effects (e.g., higher self-reported anger, lower self-esteem, lower perceived control, and higher indices of sympathetic system arousal [e.g., skin conductance]) and interpersonal effects (e.g., reduced donations to charity, reduced helping behavior, reduced money allocations in economic games, and increased aggression). In the studies included in the meta-analysis that used the Cyberball
paradigm, two items measured attributions of hostile intent. An example of one of the two items used in the Cyberball studies is, “They were being mean to me on purpose.” Responses to the two hostile intent items were made on a scale that ranged from 1 (not at all) to 7 (an extreme amount). Responses were averaged to form indices of attribution of hostile intent.

Results of the meta-analysis in Yeager et al. (2013, Study 1) revealed that having an entity mindset is positively predictive (weighted mean $r = .18$, $p < .001$) of attributions of hostile intent in response to hypothetical ambiguous peer situations (using the Dodge [2006] vignette or the Cyberball paradigm). Further, results of the meta-analysis showed that attributions of hostile intent accounted for the relationship between entity mindset and students’ desire for vengeance in response to provocation. Thus, the Yeager et al. (2013) meta-analytic results supported the hypothesized association between entity mindset and attributions of hostile intent, and in turn, attributions of hostile intent are linked to aggressive behaviors. Identifying interventions that reduce attributions of hostile intent and associated aggressive behaviors is relevant to the present dissertation because such research findings buttress targeting attributions of hostile intent to reduce harsh parenting behaviors, which in some cases will reflect aggression (i.e., retribution in response to perceived transgression or threat to self).

Study 2 and Study 3 in Yeager et al. (2013) were conducted to test whether reducing entity mindset or increasing incremental mindset is associated with attributions of hostile intent and aggressive responses to ambiguous peer situations. Both Study 2 and Study 3 in Yeager et al. used identical interventions designed to reduce entity mindset in a sample of ninth-grade students at a suburban high school. In Study 2, students were randomly assigned to one of two treatments: an entity mindset intervention or a control treatment that taught that academic skills,
such as study skills, were malleable. Students were then asked to read the ambiguous peer situation from Dodge (2006) that was described previously.

After reading the ambiguous peer situation, students responded to items assessing attributions of hostile intent. The items used were, “They were being mean to me on purpose” and “They were trying to be mean to me.” Responses to these items were made on a scale ranging from 1 (not at all) to 7 (an extreme amount). Responses were averaged to form an index of attributions of hostile intent. Open-ended responses also were solicited to measure attributions of hostile intent using an item that asked, “How would you feel?” in response to the ambiguous peer provocation. Open-ended responses were classified as either positive (e.g., trying to understand or empathize with the transgressor), neutral (e.g., ignoring the transgression), negative (e.g., wanting to hurt the transgressor), or ambivalent (e.g., containing positive elements and negative elements). As hypothesized, Yeager et al. (2013) found that students who received the entity mindset intervention, relative to students who received the control treatment, scored lower on the hostile attribution index ($d = .71$, average across two samples). In addition, logistic regression analysis revealed that receiving the entity mindset intervention, relative to the control treatment, was associated with increased likelihood of having a positive, relative to negative, reaction to the ambiguous peer provocation.

In Study 3 of Yeager et al. (2013), a longitudinal, randomized controlled trial design was used to examine whether teaching that others’ personality traits are malleable (rather than fixed) reduces attributions of hostile intent and aggression in a sample of ninth-grade students. As in Study 2, students were randomly assigned to receive an entity mindset intervention or control treatment. The entity mindset intervention was identical to the one used in Study 2. The control
treatment involved teaching participants that athletic skills were malleable. Eight months after receipt of the entity mindset intervention or control treatment, the same ambiguous peer provocation used in Study 2 was administered to the students. Students then completed the items used in Study 2 to assess attributions of hostile intent, completed items that assessed desire for vengeance, and provided open-ended reactions to the ambiguous peer situation.

In Study 3, because of skew, Yeager et al. (2013) dichotomized the attributions of hostile intent index, such that the lowest value on the scale was equal to zero and any scores above the lowest value on the scale were recoded to 1. As hypothesized, Yeager et al. replicated the main findings of Study 2. At eight months post-treatment, students who received the entity mindset intervention, relative to students who received the control treatment, were significantly less likely to attribute hostile intent to the peer in the ambiguous situation. Specifically, 22% of students who received the entity mindset intervention attributed hostile intent to the peer, whereas 38% of students who received the control treatment attributed hostile intent to the peer (unstandardized logistic $b = -1.13, p = .047$). Moreover, students who received the entity mindset intervention, relative to students who received the control treatment, were significantly lower on desire for vengeance ($d = .62, p = .008$). Finally, Yeager et al. found that the entity mindset intervention effect on desire for vengeance was significantly partially mediated (about one-third of the effect) by reductions in attributions of hostile intent.
Social Cognitive Models of Harsh Parenting and Related Interventions That Target Attributions of Hostile Intent to Reduce Parental Harsh Discipline and Harm

Social cognitive models of CPA also incorporate the notion that attributions of hostile intent are a putative causal factor prompting harsh parental discipline or parental behavior that causes harm to the child. For purposes of this dissertation, the term “parental harsh discipline” will refer to verbal (e.g., yelling, shouting, screaming) or physical (e.g., spanking, slapping, hitting) behaviors enacted by a parent toward their child with the goal being to improve or correct the child’s behavior. In contrast, the term “harm” will refer to retributive damage or injury caused by the parent toward the child, regardless of the presence of any concurrent parental intent related to child behavior improvement or correction.

Bugental and colleagues (e.g., Bugental’s transactional model of CPA; e.g., Bugental, Blue, & Cruzcosa, 1989; Bugental & Happaney, 2000) have posited that parental attributions of hostile intent to child behaviors are caused by parental relationship schemas that frame the parent-child relationship or interaction as a power struggle. According to this theoretical framework, parents believe that children enact behaviors that are aversive to parents in order to demonstrate power over the parents. An implication is that parents who hold this belief are likely to perceive themselves as “victims” of aversive child behaviors. In turn, parental beliefs about being a “victim” of aversive child behaviors are thought to be related to parents’ inferences that the child performed such behaviors with hostile intent. One theorized consequence of such parental beliefs is the use of harsh physical discipline behaviors (e.g., hitting, spanking) to re-assert control in the parent-child relationship.
Another social cognitive model (e.g., Azar, Nix, & Makin-Byrd, 2005; Stern & Azar, 1998) of CPA differs from Bugental’s model in the description of features of parental cognitive schemas about parent-child relationships. In Azar’s framework, adaptive parenting schemas are schemas that facilitate parents’ flexible generation of adaptive response alternatives. In contrast, maladaptive parenting schemas are any cognitive structures (e.g., beliefs) that increase the risk of parental harsh discipline or CPA. For example, one component of a maladaptive schema might be the belief that very young children annoy parents on purpose. Despite the differences between the Azar model and the Bugental model, it is important to note that the models are compatible because they both predict reductions in attributions of hostile intent, so both inform the research proposed in the present dissertation.

Consistent with Bugental’s and Azar’s cognitive models of CPA, Bugental et al. (2002) designed a prevention program to change the extent to which parents believe that problems in childrearing are caused by children’s hostile intent. Bugental et al. (2002) screened parents who were expecting or had just given birth. Ninety-six parents who scored at “moderate risk” for CPA based on scores on a semi-structured interview (the Family Stress Checklist; Murphy, Orkow, & Nicola, 1985) accepted an invitation to participate in a treatment program. The parents were randomly assigned to one of three treatment conditions: a no-treatment control, an “unenhanced” home visitation program, or an “enhanced” home visitation program. In the no-treatment control condition, parents received no services. In the unenhanced home visitation program, parents were given parent education, were taught ways to increase social support, and were taught healthy methods of anger management. In the enhanced home visitation program, parents were offered the same services as the parents in the unenhanced home visitation
In addition, parents in the enhanced home visitation condition participated in a cognitive appraisal training component. During the cognitive appraisal training, parents were asked to list recently experienced problems in childrearing. Then the parents were asked to generate potential causes of the problem. If the first generated cause involved blame (self or child) or attributing hostile intent to the child, then the parents were asked to continue generating potential causes for the childrearing problem. When the parent generated a cause that did not involve blame or an attribution of hostile intent, the treatment administrator would stop the parent and focus on ways to solve the childrearing problem, given the benign cause. Next, parents were asked to generate ways they could reduce or eliminate the problem, and parents decided on a strategy that they would use to solve the childrearing problem. Finally, on a subsequent home visit, the treatment administrator discussed the selected strategy with the parent to ask whether the strategy was effective, how might the effective strategy be improved further, and the best ways to implement the effective strategy.

To assess the impact of the parenting interventions, Bugental et al. (2002) measured (among other outcomes such as child health) harsh parenting. Harsh parenting was measured using responses on the Conflict Tactics Scale (CTS; Straus, 1979). The CTS is a checklist of a variety of parenting behaviors. Each item assesses the frequency of occurrence of a parenting behavior. CTS items were placed into one of two categories: physically abusive behaviors (e.g., beating up, biting) and use of non-abusive force (e.g., spanking). Abuse was operationally defined with respect to (a) whether or not there was any use of abusive force and (b) frequency (i.e., how many times the parent performed the abusive behavior). The harsh parenting measure was computed by summing the frequencies of physically abusive behaviors and the use of non-
abusive force. In addition to harsh parenting, aspects of child health relevant to CPA were measured using maternal responses to questions asked during a health interview. More specifically, health interview items assessed frequency of child injuries, frequency of child illness, and frequency of child feeding problems. Responses were standardized and then averaged to form the measure of child health. Harsh parenting and child health were moderately and significantly positively correlated ($r = .32$). In addition to the outcome measures, measures of hypothesized cognitive mediators of the effects of the enhanced home visitation program were administered. Bugental et al. hypothesized that three mediating variables would be altered in response to the enhanced home visitation program: (1) parents’ inferences of power in the parent-child relationship (i.e., parent-child control imbalance), (2) affect (i.e., anxiety and depressive symptoms), and (3) satisfaction with social support. Bugental et al. hypothesized that parent-child control imbalance would mediate effects of the enhanced, but not the unenhanced, home visitation program because the cognitive retraining component that was unique to the enhanced home visitation condition was theorized to reduce parent-child control imbalance. In contrast, Bugental et al. hypothesized that only depressive symptoms, anxiety symptoms, and satisfaction with social support would mediate the effects of the unenhanced, relative to the no-treatment control group, on outcome measures. Measurement of the hypothesized mediating constructs (i.e., perceived parent-child control imbalance, depressive and anxiety symptoms, and satisfaction with social support) is described below.

Bugental et al. (2002) measured beliefs regarding parent-child control imbalance using the Parent Attribution Test (PAT; Bugental et al., 1989). The PAT contains items that assess four factors that reflect the extent to which negative childrearing situations are (a) self-caused,
(b) self-involved but not self-caused, (c) other-caused, and (d) other-involved but not other-caused. The relationships between these four variables and physically abusive parenting behavior and child health were examined. In addition, Bugental et al. measured parental beliefs about parent-child control imbalance by asking parents to draw the self and the child. Relative size differences between the parent and the child in the parents’ drawings were used as a marker of parental perceptions of power and control imbalance in the parent-child relationship.

Affective changes were another hypothesized mechanism of change for the effects of the enhanced home visitation program on CPA and child health. Affect was operationally defined as levels of state anxiety, trait anxiety, and depressive symptoms. State anxiety and trait anxiety were measured using the State-Trait Anxiety Inventory (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983). Depressive symptoms were measured using the Beck Depression Inventory (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). Finally, social support was assessed using the Social Provisions Scale (SPS; Cutrona & Russell, 1987). The SPS is a 24-item measure of satisfaction with social support in six domains: Guidance, Reliable Alliance, Reassurance of Worth, Opportunity for Nurturance, Attachment Feelings, and Social Integration. As reported by Bugental et al. (2002), only scores from the Guidance and Reliable Alliance subscales were analyzed. An example item in the Guidance scale is, “There is someone I could talk to about important decisions in my life.” An example item in the Reliable Alliance scale is, “There are people I can count on in an emergency.” The Reliable Alliance scale and the Guidance scale were selected because these scales were thought to reflect the quality (not amount) of social support. Bugental et al. did not provide additional theoretical rationale for the exclusion from analysis of the other scales of the SPS.
Results from Bugental et al. (2002) revealed that the enhanced home visitation program (i.e., the home visitation program that was augmented with a cognitive training component), relative to the unenhanced home visitation program and the control condition combined, produced significantly less frequent use of harsh parenting (i.e., physically abusive behavior and non-abusive force). Harsh parenting frequency did not differ between the unenhanced home visitation program and the control condition. Child health in the enhanced home visitation program was significantly higher than child health in the unenhanced home visitation program, which in turn was significantly higher than child health in the control condition. Bugental et al. concluded that the pattern of results supported the causal effect of attributions of hostile intent on harsh parenting. Data supporting this conclusion included (a) parents in the home visitation program were randomly assigned to the treatment conditions (i.e., enhanced home visitation program, unenhanced home visitation program, and the control condition); and (b) the cognitive reappraisal training, which specifically targeted attributions of hostile intent, was specific to the enhanced home visitation condition.

However, it should be noted that characteristics of the enhanced home visitation (e.g., the added time and attention per se to parents in the enhanced home visitation program) might have driven differences in harsh parenting and child health between the enhanced home visitation condition and the unenhanced home visitation condition. Further, in regression analyses, Bugental et al. (2002) found only partial support for the hypothesis that reductions in perceived parent-child control imbalance mediated the effect of the enhanced home visitation condition on harsh parenting and child health. Specifically, when controlling measures of perceived parent-child control imbalance in a series of regression analyses, the effect of the enhanced home
visitation condition was substantial and approached significance ($b^* = -.29$ without controlling for perceived parent-child control imbalance, $b^* = -.22$ while controlling for perceived parent-child control imbalance). If perceived parent-child control imbalance fully mediated the effect of the enhanced home visitation condition on harsh parenting and child health, the effect of the enhanced home visitation condition should have been eliminated when controlling for parent-child control imbalance. Instead, the results suggest that perceived parent-child control imbalance only partially mediated the effect of the enhanced home visitation program on harsh parenting and child health. Thus, it is possible that the enhanced home visitation program reduced attributions of hostile intent independent of perceived parent-child control imbalance.

Instead of targeting parent-child control imbalance (via cognitive appraisals of hypothetical parent-child scenarios) as a means of reducing hostile intent attributions, Rutledge et al. (2018) sought to reduce parents’ attributions of hostile intent and use of harsh discipline by reducing parents’ beliefs that their child’s personality is immutable. To do so, Rutledge et al. (2018) screened parents for high scores on an implicit personality theories measure (Dweck, Chiu, & Hong, 1995). High-entity mindset parents were randomly assigned to receive either (a) an experimental treatment that taught parents that children’s personality, children’s temperament, children’s behavior, and parenting ability are malleable attributes (i.e., *incremental mindset condition*); or (b) a control condition that involved watching a portion of a documentary film (*March of the Penguins*). The experimental treatment in Rutledge et al. (2018) was delivered to parents using an audio-enhanced slideshow presentation. As part of the experimental treatment, parents were presented with testimonials made by community parents who expressed incremental beliefs about parenting or their child’s personality. After the slideshow presentation,
parents wrote (and then audio-recorded) a letter to a new parent, encouraging them to adopt an incremental view of their children and their role as a new parent.

Results reported by Rutledge et al. (2018) revealed that, contrary to hypotheses, overall (i.e., for all three types of child transgressions that were used), participants in the incremental mindset condition ($M = 2.79$, $SD = 1.01$, $n = 32$), compared to participants in the control condition ($M = 3.20$, $SD = 1.14$, $n = 31$), exhibited (non-significantly) lower mean attributions of hostile intent, $d = .38$, 95% lower end-point [-0.04, $\infty$], one-tailed $p = .072$. Similarly contrary to hypotheses, overall, participants in the incremental mindset condition ($M = 2.46$, $SD = 0.90$), compared to participants in the control condition ($M = 2.83$, $SD = 1.15$), self-reported (non-significantly) lower mean likelihood to use harsh parenting (i.e., yell/shout/slap/hit) on the child, $d = .36$, 95% lower end-point [-0.06, $\infty$], one-tailed $p = .115$.

In sum, evidence for the effectiveness of interventions to reduce parental attributions of hostile intent and harsh parenting is mixed. Specifically, Rutledge et al. (2018), did not observe the hypothesized differences between the incremental mindset intervention and the control condition in attributions of hostile intent or harsh parenting. (However, it is true that differences were in the predicted directions (see Rutledge, 2016, for a discussion of potential future directions for research on the potential effects of incremental mindset interventions on attributions of hostile intent and harsh parenting).

In contrast, evidence reported in Bugental et al. (2002) supports the hypothesis that teaching parents to generate benign solutions to ambiguous child behaviors reduces parents’ attributions of hostile intent and harsh parenting. Specifically, the enhanced home visitation program ($M = .06$) was associated with the least frequent use of harsh parenting, whereas the
unenhanced home visitation program and control condition (combined $M = .25$, overall $SD_{pooled} = .26$) exhibited more frequent harsh parenting, $d = .73$, 95% lower end-point [.31, $\infty$], one-tailed $p = .011$. Thus, the Bugental et al. (2002) results suggest program effectiveness.

However, a pragmatic limitation potentially will be introduced when such parent education programs are implemented to service families at high risk for CPA. According to the SIP model of CPA, high-risk parents hold negative child-related schemata (e.g., negative child-related attitudes, beliefs that children are inherently bad, beliefs advocating the use of corporal punishment, and beliefs that children do bad behaviors to annoy parents) that parents bring to parent-child interactions. Some scholars (Milner & Crouch, 2013; Milner et al., 2017) have speculated that high-risk parents are motivated to maintain consistency with prior views (e.g., children are inherently bad) when faced with schema-incongruent information. For example, when high-risk parents are faced with schema-incongruent information to maintain cognitive consistency high-risk parents may disbelieve or discount the validity of schema-incongruent information while accepting only schema-irrelevant or schema-congruent information (i.e., an example of “motivated reasoning”; Kunda, 1990).

One consequence of such “motivated reasoning” may be that high-risk parents attrite. Two potential mechanisms that might explain the link between “motivated reasoning” and high-risk parents’ attrition are (1) high-risk parents lose interest in the program and/or the parent educational content because of perceived invalidity of the parent educational content or (2) high-risk parents avoid encountering schema-incongruent information so that cognitive dissonance never occurs (i.e., “motivated avoidance”).
Given these issues, Milner and Crouch (2013) suggest that a procedure is needed that eliminates mechanisms (e.g., “motivated reasoning” processes, “motivated avoidance”) that putatively prevent parent education programs from effectively reducing attributions of hostile intent and harsh parenting. One such solution may lie in the procedure of positive evaluative conditioning (EC). Because of its relatively non-threatening nature, positive evaluative conditioning might avoid the theorized psychological causes of high-risk parents’ attrition from parenting education programs. At the same time, the EC procedure has the potential to reduce CPA risk by changing components of high-risk or abusive parents’ child-related schemata.

This latter prediction is derived from the Social Information Processing (SIP) model of child physical abuse (Crouch & Milner, 2005; Milner, 2000). Thus, the prediction suggesting that EC can reduce CPA risk will be better understood after a brief presentation of the SIP model. This is accomplished in the section that follows. After describing the SIP model of CPA, in a subsequent section the EC procedure is described both in general terms and in the specific context of the studies that were conducted for the present dissertation.

The Social Information Processing Model of Child Physical Abuse

The SIP model of CPA (Crouch & Milner, 2005; Milner, 2000) posits that abusive and high-risk parents hold negative child-related schemata (e.g., negative child-related attitudes, beliefs that children are inherently bad, beliefs advocating the use of corporal punishment, and beliefs that children do bad behaviors to annoy parents) that parents bring to parent-child interactions. These pre-existing negative child-related schemas are hypothesized to influence
parental social information processing during a parent-child interaction at three social-cognitive phases (perceptions, interpretations and evaluations that give meaning to social behavior, and information integration and response selection) and a fourth behavioral stage (response implementation and monitoring).

Consider the following example to illustrate the components of the SIP model of CPA: A mother or father at high risk for CPA is combing the child’s hair and the child grimaces (in response to the pain) when the parent unintentionally yanks knotted hair. The high-risk parent’s negative pre-existing attitudes about children might cause the parent to misperceive the grimace to be a look of anger instead of pain (perceptions of social behavior). The high-risk parent might then erroneously misattribute hostile intent to the child’s grimace; the parent might infer that the child’s intent when grimacing was to annoy or irritate the parent and evaluate the behavior as negative or bad (interpretations and evaluations). In response to the high-risk parent’s erroneous attribution of hostile intent, the high-risk parent might feel anger and frustration. Possible responses to the situation may be considered, but given the situational information processed, the parent may consider mostly punitive responses (information integration and response selection). Thus, the high-risk parent might choose to use harsh discipline (e.g., spanking) to attempt to correct the child’s behavior; alternatively, the high-risk parent might strike the child to inflict harm (response implementation). The outcome in either hypothetical scenario might escalate to, or reflect, CPA.

Consistent with and relevant to propositions of the SIP model of CPA, person memory research has shown how accessed or activated knowledge structures (e.g., “aggressiveness” or “kindness”) can influence judgments about individuals (e.g., Bargh & Thein, 1985; Higgins,
King, & Mavin, 1982; Higgins, Rholes, & Jones, 1977; Srull & Wyer, 1979). In two classic experiments, Srull and Wyer tested the hypothesis that activation of a knowledge structure (i.e., the traits hostile [Experiment 1] and kind [Experiment 2]) would cause perceivers to form impressions about an actor in a manner consistent with the direction of the activated knowledge structure. In Experiment 1, undergraduate students completed a scrambled sentence task designed to activate the concept hostile. In scrambled sentence tasks (Costin, 1975), participants encounter a set of four words in random order that could be re-ordered to form at least two complete sentences. In Experiment 1, the scrambled sentences (e.g., “leg break his arm”) could be re-ordered to make the hostile sentence “break his arm” or “break his leg.” Participants were instructed to underline, as quickly as possible, three of the four words in each item that could be re-ordered to form a complete sentence. Hostile priming items always formed hostile sentences, while control items formed filler sentences (e.g. “the hug boy kiss” could be re-ordered to read “hug the boy” or “kiss the boy”).

After the scrambled sentence task, participants read descriptions of behaviors performed by a hypothetical target person. The behavioral descriptions were pre-tested to be ambiguous with respect to the trait underlying the behaviors. Specifically, a sample of participants who did not take part in the Srull and Wyer experiments rated individual behaviors in a large pool along a scale assessing hostility from 0 (not at all hostile) to 10 (extremely hostile). From this pool, five behaviors were selected that conveyed high hostility ($M = 8.08$) and five behaviors were selected that conveyed low hostility ($M = 0.58$). In addition, five ambiguous behaviors were selected. To be selected as a behavior that was ambiguous with respect to hostility, the behavior was required to meet two criteria: (1) have a mean hostility rating that was middling between the mean rating
of low-hostility behaviors and the mean rating of high-hostility behaviors and (2) have a standard deviation of hostility ratings that was greater than 2.76, which was equal to the largest standard deviation for either low-hostility behaviors or high-hostility behaviors.

Srull and Wyer (1979) hypothesized that depending on the activated trait concept, the ambiguous behaviors would be interpreted in terms of the activated trait concept (e.g., hostility in Experiment 1). Accordingly, it was hypothesized that the number of hostile items in the scrambled sentence task would be positively related to the extent to which participants evaluated the hypothetical target person along dimensions that relate to hostility. To measure hostility-related evaluations of the target person, Srull and Wyer asked participants to rate the target on adjective descriptors. Some of the descriptors related to hostility/kindness (hostile, unfriendly, dislikable, kind, considerate, and thoughtful). The other descriptors were semantically unrelated to hostility/kindness but were selected to be evaluatively-laden (i.e., boring, selfish, narrow-minded, dependable, interesting, and intelligent). Responses to the descriptors were made on a scale from 1 (Not at all) to 10 (Extremely). After making the target ratings, participants rated the hostility of each of the 20 pre-tested behaviors on a response scale from 1 (Not at all hostile) to 10 (Extremely hostile). Participants concluded the study session by providing estimates of the co-occurrence of the trait hostile with each of the other 11 descriptors used to measure impressions of the target person. Estimates of traits co-occurring with trait hostility were measured using 11 items that followed the form, “If a person is hostile, how likely is it that he is ______?”

To manipulate the activation of the trait concept hostility, Srull and Wyer (1979) administered different versions of the scrambled sentence task. The versions differed in (a) the
total number of items (60 or 30) and (b) the percentage of the total number of items that were hostile priming items (20% or 80%). Thus, the absolute number of hostile priming items ranged from 6 to 48. Further, Srull and Wyer manipulated time lag between priming task and encoding of ambiguous behavioral descriptions (no delay, 1-hour delay, 24-hour delay). The time lag manipulation was designed to test the hypothesis that priming effects decrease as a function of time lag, either because intervening activated constructs during the time lag decrease the likelihood that the initially activated trait is used in the encoding process (Wyer & Srull, 1980) or because the time lag per se decreases “excitation” of the initially activated trait concept (Collins & Loftus, 1975).

Results of the Srull and Wyer (1979) experiments revealed support for the hypotheses. The total number of items on the scrambled sentence task and the proportion of the items that reflected hostility additively increased the mean hostile rating of the target person. In addition, as the time lag between the scrambled sentence task and the trait rating task increased, the mean hostility ratings decreased. The time lag effect is consistent with Srull and Wyer’s theorized mechanisms (i.e., the construct-interference hypothesis [Wyer & Srull, 1980] and the excitation reduction hypothesis [Collins & Loftus, 1975]). Thus, Srull and Wyer concluded that priming the trait concept hostility influences trait judgments, at least when the target behaviors from which trait inferences/judgments were generated are ambiguous. However, note that the effect size in the Srull and Wyer study may substantially overestimate the actual effect size (McCarthy et al., 2018).

Relatedly, other research (Bargh & Maguire, 1985; Higgins et al., 1982) has shown that trait constructs that are chronically or highly accessible in memory influence judgments of
behaviors that are ambiguous. For example, Higgins et al. (1982) reasoned that chronically thinking about others in terms of certain traits reflects, and increases, trait accessibility.

Operating under this assumption, in Higgins et al. solicited 16 undergraduate participants’ highly accessible traits. To solicit the participants’ highly accessible traits, participants were asked to think about two male friends, two female friends, and themselves. For each of these target people, the participants provided up to ten traits that described the target person. For each participant, a trait was considered accessible if (a) it appeared in the participant’s description of both him or herself and at least one friend or (b) it appeared in the participant’s description of three or more friends. Traits that were not classified as accessible were deemed inaccessible.

Participants in a first experiment conducted by Higgins et al. (1982) returned approximately two weeks after the initial experimental session. During the return visit, participants read an essay that was constructed to be idiosyncratic to each participant. Each participant’s essay contained behaviors that reflected traits that were accessible to the participant and behaviors that reflected traits that were deemed inaccessible to the participant. After reading the essay and after a 10-minute filler task designed to clear short-term memory, participants were asked to free recall (i.e., reproduce word for word) the essay. In addition, participants were asked to write down the “sort of person they thought the target person was” (Higgins et al., 1982, p. 38). Finally, two weeks later participants returned to produce another free recall of the essay and to rate their impressions of the target person.

As hypothesized, Higgins et al. (1982) found that accessible traits were significantly more likely than inaccessible traits to appear in participants’ free recall of the essays about the target person and in participants’ impressions formed about the target person. In other words,
inaccessible traits were more likely than accessible traits to be omitted from participants’ reproductions of the essay and participants’ impressions of the target person. Thus, the results of Higgins et al. support the notion that trait concepts that are highly accessible influence people’s social judgments about others to be consistent with the trait concepts that are highly accessible to the perceiver.

As discussed previously, according to the SIP model of CPA, high-risk or abusive parents hold negative child-related schemata that influence social information processing during parent-child interactions. Given that high-risk or abusive parents more often report having experienced violent victimization or perpetration (Milner, 1993, 2000, 2003), trait concepts related to hostility and aggression are putatively highly accessible in high-risk or abusive parents. Thus, highly or chronically accessible trait concepts related to hostility and aggression are considered a component of high-risk or abusive parents’ negative child-related schemata (e.g., Crouch et al., 2009). In contrast, in low-risk or non-abusive parents, trait concepts related to hostility and aggression are putatively relatively inaccessible.

Evidence for the existence of negative, child-related schemata in high-risk or abusive parents is mixed. To summarize the literature, a recent meta-analysis (McCarthy, Wagner, Basham, & Jones, 2016) was conducted on studies comparing high-risk or abusive parents and low-risk or non-abusive parents on measures of “positive impressions” and “negative impressions.” In the McCarthy et al. (2016) meta-analysis, parental impressions of children were defined as “parents’ judgments, inferences, or evaluations about a child’s disposition or personality” (p. 4). Twenty-four independent studies were identified that met inclusion criteria, i.e., that the study (a) included parent participants, (b) provided a way to compute a standardized
mean difference comparing high-risk or abusive parents on positive/negative impressions about
children, and (c) ensured that child stimuli were held constant across participants. Separate
random-effects meta-analyses for parents’ positive impressions and parents’ negative
impressions revealed the following results: Overall, high-risk/abusive parents, relative to low-
risk/non-abusive parents, formed less positive impressions of children, $d = 0.17$, 95% CI [0.00,
0.33], and more negative impressions of children, $d = .35$, 95% CI [.20, .51]. In summary,
results of McCarthy et al. are consistent with the SIP model hypothesis that high-risk or abusive
parents, relative to low-risk or non-abusive parents, have more negative child-related schemas.

Another proposition in the SIP model of CPA is that negative child-related schemata in
high-risk or abusive parents influence the interpretation of social information in the context of
parent-child interactions. Negative child-related schemata putatively have their largest
influences on social information processing when aspects of the parent-child interaction or child
behavior are ambiguous (Milner, 2000). For example, studies (i.e., Crouch et al., 2010; Farc et
al., 2008) have shown that high-risk, relative to low-risk, parents judge ambiguous child faces
depicted in photographs as especially hostile.

Farc et al. (2008) conducted two experiments. In both experiments, general population
parents provided evaluations of ambiguous child faces (see Farc et al., 2008, for additional
details on pretesting procedures). In addition, parents were randomly assigned to a hostile
priming condition or a neutral control condition before rating the ambiguous child photographs.
In Experiment 1, the hostile priming task was a scrambled sentence task that has been used in
prior research (Srull & Wyer, 1979) to prime hostility. In Experiment 2, the hostility concept
was primed using a subliminal vigilance task designed to prime hostility despite participants’
inability to consciously perceive the hostile primes. Specifically, for the vigilance task parents in Experiment 2 were instructed to focus on a central focal point on the computer screen. Priming words were presented on the screen in the parafoveal visual region where participants would be unable to cognitively process the priming words (Bargh & Chartrand, 2000). The hostile word set included the following 12 words: uncooperative, aggressive, defiantly, irritable, mean, oppositional, unfriendly, cold, violent, hostile, difficult, and negative. The neutral word set comprised the following 16 words: water, long, number, people, what, little, many, something, together, different, between, said, every, another, always, and there.

In both experiments conducted by Farc et al. (2008), after completing the priming task parents were asked to rate ambiguous child pictures on trait descriptor adjectives: hostile, negative, difficult, friendly, cooperative, sweet, content, lively, and attached. Responses for each trait descriptor adjective were made on a scale from 1 (not at all) to 10 (extremely likely), then parents completed the Child Abuse Potential Inventory (CAP; Milner, 1986) so that parents could be classified as low-risk or high-risk for CPA. Consistent with the SIP model of CPA, the CPA risk classification was assumed to be a proxy for the chronic accessibility of negative or hostile child-related schemata.

Responses to the positive trait descriptors and negative trait descriptors were compared between CPA risk groups to test the hypothesis that high-risk, relative to low-risk, parents judge ambiguous child stimuli as more hostile. Consistent with hypotheses, high-risk, relative to low-risk parents, rated ambiguous children as more positive and less negative. Overall, hostile, relative to neutral, priming produced significantly higher ratings on the trait descriptor “hostile.” There was no interaction between CPA risk group and priming condition, indicating additive (not
interactive) effects of chronic accessibility (i.e., child-related schemata differences between high-risk and low-risk parents) and acute accessibility (i.e., impact of hostile priming) of hostile or negative child-related schemata on parents’ ratings of ambiguous children.

Associative or linking theories of priming and chronic accessibility effects posit that priming activates a construct that, in turn, putatively influences judgments or behavioral responses (Wyer, 1974). Moreover, such theories suggest that constructs associated with the primed construct may become activated via spreading activation (i.e., the hypothesis that priming one construct, such as hostility, also primes related constructs, such as anger).

Associative or linking theories also suggest that by eliminating putative associations between children and negativity in individuals, social-cognitive processes putatively influenced by the negative child-related schemata will be less likely to occur. Therefore, according to the SIP model, eliminating or reducing the association between children and hostility or negativity will result in reductions in harsh parenting.

Such ideas influenced the research described in the present dissertation. In my dissertation research, variants of a positive EC procedure were employed in an attempt to reduce negative child evaluations. The expectation was that because of links between negative child evaluations and other mental constructs, such reductions in negative child evaluations would co-occur with reductions in attributions of hostile intent, anger, and harsh parenting behaviors.

**Applying Positive Evaluative Conditioning (EC) in the Context of the SIP Model of CPA**

EC is functionally defined as the changing of liking or valence of a stimulus that is a result of repeated pairings of the stimulus with a second, valenced stimulus. In an EC procedure,
the stimulus for which the valence is expected to change is called the conditional or conditioned stimulus (CS), and the stimulus that is paired with the CS to change the CS’s valence is called the unconditional or unconditioned CS. Evaluative conditioning is a type of Pavlovian conditioning in which liking or evaluations are the properties being changed; Pavlovian conditioning generally refers to any change in a stimulus as a function of pairing with the US (De Houwer, 2012). As in the proposed research, when the goal is to change positive evaluations of the CS, the EC procedure can be labeled positive EC.

Characteristics of the CS and US vary between EC studies, and many types of EC characteristics have been identified and shown to exert effects on the magnitude of change following EC. According to meta-analyses by Hofmann et al. (2010), changing pre-existing evaluations, relative to forming new evaluations, of the CS is typically associated with smaller EC effects. However, studies published after Hofmann et al. (and thus not included in the meta-analyses) provide additional positive evidence for EC effects when the CS has pre-existing valence. For example, Hollands, Prestwich, and Marteau (2011) demonstrated that an EC procedure that used pictures of unhealthy snacks as CSs and aversive pictures of negative potential health consequences of unhealthy eating as USs significantly increased the choice to consume fruit rather than unhealthy snacks in a behavioral choice task; moreover, EC-induced changes in implicit attitudes toward unhealthy eating mediated the EC effects on behavioral choice. Other studies published after Hofmann et al. showed similar EC effects in other basic, applied, and clinical domains: social anxiety (Clerkin & Teachman, 2010; Schnabel & Asendorpf, 2015, Study 1), alcohol liking and use (Houben, Havermans, & Wiers, 2010; Houben, Schoenmakers, & Wiers, 2010), recycling (Geng, Liu, Xu, Zhou, & Fang, 2013), and
homelessness (Balas & Sweklej, 2013). Further, like Hollands et al. (2011), most studies that measured behavior after EC found that EC-induced changes in evaluations relate to short-term changes in behavior.

Although EC is associated with the functional changes described above, the EC literature is also rife with controversy concerning hypothesized cognitive mechanisms underlying effects obtained in EC research. Understanding cognitive mechanisms underlying EC effects is important because experimental conditions can constrain the underlying cognitive mechanisms that mediate the observed EC effects observed in the manipulated experimental conditions. To the extent that cognitive mechanisms required for EC are not present or operating typically in realized experimental conditions, one may observe boundary conditions to the EC effect. Boundary conditions are conditions in which the hypothesized effects are expected not to occur because the necessary cognitive mechanisms (a) are not operating in manipulated experimental conditions or (b) are manipulated via an experimental condition so that the cognitive mechanisms are “turned off” or operate in a way to produce effects that may mandate a label consistent with a functional effect distinct from EC.

According to the SIP model of CPA, EC-induced reductions in negative child-related evaluations will remove the influences of negative child-related evaluations on social information processing in high-risk or abusive parents. For example, negative child-related evaluations are thought to increase the likelihood that high-risk or abusive parents, relative to low-risk or non-abusive parents, automatically encode ambiguous child behaviors in hostile terms and attribute hostile intent to ambiguous child behaviors (e.g., Crouch & Milner, 2005). Empirical results mostly have been consistent with these notions (Crouch et al., 2009; Crouch et
al., 2010; Farc et al., 2008; cf. Risser, Skowronski, & Crouch, 2011, for research suggesting that implicit attitudes toward children are unrelated to CPA risk).

The SIP model of CPA also suggests that reducing negative child-related evaluations, a component of negative child-related schemata in high-risk or abusive parents, will reduce the likelihood of parental use of harsh parenting (e.g., Crouch & Milner, 2005). Two types of harsh parenting can be defined: *harsh discipline* and *harm*. Harsh discipline refers to parental behaviors directed toward the child with the goal to improve future child behavior. Harsh discipline can be delivered in verbal modality (*harsh verbal discipline*, e.g., yelling, shouting, or screaming) or in physical modality (*harsh physical discipline*, e.g., slapping, hitting, or spanking). In contrast to harsh discipline, *harm* is done regardless of whether or not the motive of the harm included shaping the child’s future behavior.

Milner, Wagner, and Crouch (2017, Experiment 6) generated evidence supporting the hypothesis that positive EC procedure alters parents’ attributions of hostile intent, harsh verbal discipline, harsh physical discipline, and harm. More specifically, in Milner et al., parents in a convenience sample (*N* = 107) were run in one of two positive EC dose conditions: (1) one EC presentation or (2) two EC presentations in which the second EC presentation was viewed 3 min after the first EC presentation. Parents in both EC dose conditions first were shown three ambiguous child photographs pre-tested and used in prior research (Farc et al., 2008). For each of the three ambiguous child photographs, parents were shown a list of six adjectives (i.e., sweet, friendly, cooperative, hostile, negative, difficult) that might describe the child depicted in the photograph. For each adjective, parents were asked to indicate the extent that the adjective described the child depicted in the photograph. Parents indicated their responses to each
adjective by selecting the appropriate value on a response scale ranging from 1 to 10. The low end of the scale was labeled Not at all, the middle range of the scale was labeled Somewhat, and the high end of the scale was labeled Extremely. Subsequently, for the same ambiguous child photograph, participants were asked to indicate (using the same response scale described above) the likelihood that the child depicted in the photograph was likely to require discipline in the future. Across the ambiguous child photographs, responses to three adjectives (i.e., sweet, friendly, cooperative) were averaged to form a measure of positive child-related evaluations. Across the ambiguous child photographs, responses to the other three adjectives (i.e., hostile, negative, difficult) were averaged to form a measure of negative child-related evaluations. Across the ambiguous child photographs, responses to the item assessing future discipline expectations were averaged to form a measure of expectations of future discipline. In the next phase, parents were exposed to ambiguous child scenarios.

Ambiguous child scenarios were created by pairing each of the three ambiguous child photographs with each of six ambiguous child behaviors. Ambiguous child behaviors were created and pre-tested in a prior study (Crouch et al., 2010). Specifically, an initial pool of 20 sentences was generated. Each sentence in the pool contained (a) a child’s name (e.g., Riley), (b) a child’s action (e.g., “kicked his legs”), and (c) a care-giving context (e.g., “as his mother changed his diaper”). The children’s names were selected from a list of the 100 most popular names in 2002. Children’s actions were selected from a pool of cues that a child younger than 3 years of age might display to indicate disengagement. Care-giving contexts were contexts that occur frequently (e.g., feeding, bathing) when providing care for young children.
To determine which of the initial 20 child sentences reflected child behaviors that were ambiguous with respect to hostile intent, a sample of 66 female undergraduates was recruited. After reading each behavior, participants were shown 24 adjectives. Participants were asked “to rate the extent to which each word below describes the child’s behavior.” Twelve adjectives were negative descriptors: uncooperative, aggressive, defiant, difficult, irritable, mean, negative, oppositional, hostile, unfriendly, cold, and violent. The other 12 adjectives were positive descriptors: peaceful, animated, loving, attached, playful, content, happy, friendly, lively, sweet, accepting, and endearing. Participants rated each descriptor on a scale from 1 (not at all) to 5 (very much). For each child behavior, a measure of positive child-related evaluations was computed by averaging the ratings for the positive descriptors, and a measure of negative child-related evaluations was computed by averaging the ratings for the positive descriptors.

Below each ambiguous child scenario were items that solicited data used as measures of attributions of hostile, anger, harsh verbal discipline, and harsh physical discipline. For each of the items, responses were averaged across the eighteen ambiguous child scenarios. The specific items were (a) attributions of hostile intent: “To what extent did the child do this behavior to intentionally annoy/irritate you?”; (b) anger: “To what extent did the child do this behavior to intentionally annoy/irritate you?”; (c) harsh verbal discipline: “What is the likelihood that you would yell/shout/scream at your child because of this behavior?”; (d) harsh physical discipline: “What is the likelihood that you would slap/hit/spank your child because of this behavior?” For each ambiguous child scenario, measures (a) through (d) were presented in a fixed order.

After the set of items (i.e., [a] through [d] described above), participants proceeded to a separate page in the survey where the ambiguous child scenario on the previous page was
displayed again. This page contained a modified version of the Voodoo Doll Task validated for use in parent-to-child aggression research (McCarthy, Crouch, Basham, Milner, & Skowrons, 2016). Specifically, participants encountered the following prompt:

After the child performs this behavior, you might want to harm the child. Below is a picture of an outline of a child. Click on the blank square areas if you wish to harm the child at all at the place you clicked. After you have made your decision about harming the child, click NEXT.

A schematic outline of a child image was below the prompt. Consistent with the information provided in the prompt above, parents in Milner et al. (2017, Experiment 6) could click on researcher-defined square areas on the child outline. Parents could only click each area once. Parents could click on up to 23 unique areas on the child outline. Each of these clicks putatively represented an instance of “symbolic harm,” which rests on an assumption the “people easily project properties of known individuals onto symbolic representations of these individuals” (McCarthy et al., 2016; p. 135).

Symbolic representations of an individual can be generated by the perceiver via “magical thinking.” Specifically, DeWall et al. (2013) reviewed research that suggests that external agents (e.g., individuals) can influence another person’s behavior in the absence of any physical explanation for the influence (e.g., in physical space the external agent is located far away from the target of the agent’s influence). Putatively, magical thinking is one of the key psychological mechanisms that underpins both (a) the set of evidence used to support DeWall et al.’s argument and (b) each instance of symbolic harm during the VDT (e.g., pins inserted or “clicks” that represent pins inserted).

DeWall et al. (2013) generated evidence supporting the construct validity of pins used during the VDT as a measure of symbolic harm or aggressive inclinations (against an intimate
partner or a close friend). In DeWall et al., some participants were told to think of a close friend while others were told to think of their current intimate/romantic partner. For example, dispositional tendencies toward physical aggressiveness (assessed using the Physical Aggressiveness subscale of the Buss-Perry Aggression Questionnaire; Buss & Perry, 1992) was positively related to the number of pins used during the VDT. DeWall et al. (2013) observed parallel results for a measure of “intimate partner violence inclinations” (p. 424).

Specifically, participants responded to a series of 20 vignettes in which each vignette described a brief scenario involving a negative social interaction involving one’s partner (i.e., either close friend or romantic partner). Example vignettes were, “My partner ridicules or makes fun of me.” For each vignette participants responded on a scale from 1 (Not at all likely that I would be physically aggressive) to 9 (Extremely likely that I would be physically aggressive). Obtaining hypothesized experimental results across dependent measures of putatively related constructs (i.e., self-reported physical aggressiveness, “intimate partner violence inclinations”) provides evidence for the VDT’s construct validity. However, given that the participants in DeWall et al. (2013) were instructed to imagine their close friend (or intimate partner if applicable), the results of DeWall et al. do not directly test whether a modified VDT may be validly applied in parent-to-child aggression research.

To provide evidence that the VDT is useful measure in the context of parent-to-child aggression research, McCarthy et al. (2016) conducted a series of six studies. All six studies provided evidence for the valid application of a modified VDT in parent-to-child aggression research. An example of such evidence was that across three studies (Study 1, Study 2, and Study 6), parents’ level of trait aggression (assessed using the brief or full-length version of the
Buss-Perry Aggression Questionnaire; Bryant & Smith, 2001; Buss & Perry; 1992; Webster et al., 2014) positively related to pin usage during the VDT. Especially relevant to the present studies is that in Study 6, parents’ CPA risk positively related to pin usage during the VDT. Instructional prompts in the VDT varied across the studies, yet all variations yielded (hypothesized) parallel results. McCarthy et al. concluded that, on the basis of being sensitive to individual differences (e.g., trait aggression, Studies 1, 2, and 6; risk for CPA, Study 6), experimental manipulations (e.g., following hypothetical child misbehavior, imagining child non-compliance, relative to imagining child compliance, Study 5), and demographic variables (e.g., fathers’ pin usage > mothers’ pin usage; parent education level negatively related to pin usage), the modified VDT can be validly applied to research on parenting-to-child interactions.

After completing items (a) through (f) described above and the VDT for all ambiguous child scenarios, participants in Milner et al. (2017, Experiment 6) proceeded to the positive EC phase of the experiment. Instructional screens indicated to participants that their job during the task was to focus on the center of the screen while image-word pairs are presented on the screen. The image-word pairs were the positive EC trials. In each positive EC trial, a plus sign (+) was displayed for 1000 ms. Following the +, one of three ambiguous child photographs (i.e., the CS, programmed to be displayed for 20 ms) was displayed. One of three positive descriptor UCS (i.e., sweet, friendly, and cooperative) appeared immediately after the ambiguous child photograph and was programmed to display for 20 ms. Thus, nine unique positive EC trials were created by pairing each unique CS with a unique UCS. Twenty replicates of each positive EC trial (presented in a random order unique to each participant) constituted the 180-trial positive EC procedure. The positive EC procedures to be used in the proposed research will be described
in detail later. Subsequent to the positive EC procedure, participants in Milner et al. (2017, Experiment 6) participants completed pre-EC items again post-EC.

As hypothesized, results of Milner et al. (2017, Experiment 6) revealed, regardless of EC dose, significant mean gains in positive child evaluations pre-EC, relative to post-EC. As hypothesized, regardless of EC dose, significant mean reductions were observed for negative child evaluations, the likelihood that the child would require discipline in the future, and attributions of hostile intent, anger, harsh verbal discipline, harsh physical discipline, and harm.

A planned exploratory analysis removed participants who could not change in the hypothesized direction (because their pre-EC score was on the floor or ceiling of the scale). Relative to the analysis on the full sample, this exploratory analysis revealed larger mean differences in dependent measures when comparing pre-EC and post-EC scores (albeit mean differences that were estimated with less precision due to fewer participants). Moreover, exploratory analyses (Milner et al., 2017, Experiment 6) revealed that among the participants who gave at least one pin during the pre-EC VDT ($n = 25$), 11 or 44% gave no pins during the post-EC VDT. In contrast, among the participants who gave no pins during the pre-EC VDT ($n = 71$), only one or 1.4% gave at least one pin during the post-EC VDT. These last two findings may indicate that not only does positive EC have potential utility, but it may also carry minimal additional risk (e.g., increasing harm to children).

Other studies reported in Milner et al. (2017) provide evidence to counter two (undesired) explanations for observed positive EC effects: (a) positive EC effects are caused by demand characteristics (e.g., participants having insight into the study purpose or hypotheses and producing responses consistent with this insight) and (b) positive EC effects are caused by non-
specific elements of the procedure (e.g., merely participating in the research produced positive EC effects). One such study was Experiment 3. The design of Experiment 3 rendered such undesired explanations for observed positive EC effects (across all experiments reported in Milner et al., 2017) highly implausible for three reasons: (1) the study assessed evaluations of exemplar ambiguous child photographs that were not included in the positive EC procedure, (2) the study used a control “pseudo-EC” condition in which the US had no valence (i.e., non-sense letter strings), and (3) EC effects for positive child evaluations and negative evaluations were maintained approximately one week later.

In Milner et al. (2017) Experiment 3, community parents ($N = 80$ before exclusions) were randomly assigned to the positive EC condition ($n = 44$ after exclusions) or the pseudo-EC condition ($n = 32$ after exclusions). In Experiment 3, three ambiguous child photographs (from the same pool of ambiguous child photographs pre-tested and reported in Farc et al., 2008) were used as “exemplars.” The exemplar ambiguous child photographs were not used as CSs (i.e., they were never shown during the EC procedures). A subset of the same (previously described) measures (i.e., positive child-related evaluations, negative child-related evaluations, and expectations in the need for future discipline) administered pre-EC and post-EC for each of the initial ambiguous child photographs (i.e., CS) were administered for the exemplar ambiguous child photographs. Where possible, participants returned approximately one week later and repeated the initial session’s protocol (overall retention rate = 71/76 = 93.4%).

A summary of results for Experiment 3 (Milner et al., 2017, pp. 52-53) indicated that positive EC (but not the pseudo-EC) increased positive child-related evaluations and decreased negative child-related evaluations and expectations in the need for future discipline. For the CSs,
the observed increases and observed reductions were maintained at 1 week follow-up. For the exemplar ambiguous child photographs, only positive child-related evaluations increased pre-EC to post-EC, but positive child-related evaluations significantly decreased post-EC to 1 week follow-up.

Thus, the collective evidence (Milner et al., 2017) sets a strong foundation on which to base the proposed studies. Nevertheless, despite the strengths of the Milner et al. (2017) experiments, a number of unanswered research questions remain. Some of these questions concern the mental mechanisms that prompt the changes observed in the Milner et al. (2017) EC paradigm. These mechanisms will be described in detail in later sections of the present dissertation.

However, to set the stage for the studies that explore such questions, a preliminary study was conducted. One goal of the study was to confirm (i.e., replicate) the high reliability for measures used in Milner et al. (2017). A second purpose of the preliminary study was to test hypothesized relationships among positive child evaluations, negative child evaluations, future discipline expectations, attributions of hostile intent, anger, harsh verbal discipline, harsh physical discipline, doing nothing, and harm. Finding evidence for such relationships would provide evidence for the study measures’ construct validity.
CHAPTER 2
PRELIMINARY STUDY

The following hypotheses were tested in the preliminary study:

Hypothesis 1: Positive child evaluations will be negatively correlated with the following responses to ambiguous child scenarios: attributions of hostile intent, anger, harsh verbal discipline, harsh physical discipline, doing nothing, and harming the child.

Hypothesis 2: Negative child evaluations will be positively correlated with the following reactions to ambiguous child behaviors: attributions of hostile intent, anger, harsh verbal discipline, harsh physical discipline, doing nothing, and harming the child.

Hypothesis 3: The following responses to ambiguous child behaviors will be positively correlated: attributions of hostile intent, anger, verbal discipline, physical discipline, and harm.

Hypothesis 4: Parents’ selection of doing nothing in response to the ambiguous child behaviors will be negatively correlated with attributions of hostile intent, anger, use of harsh verbal discipline, use of harsh physical discipline, and harming the child.

If hypotheses are supported, the results will constitute evidence for the construct validity of proposed measures.
Method

Participants

Inclusion criteria for the study were that participants (a) were parents, (b) had at least one child under the age of 18 living in the home, and (c) spoke English as their primary language. Two-hundred and fifty participants were recruited online via Mechanical Turk (www.amazon/mturk.com) and were forwarded to the study on Qualtrics (www.qualtrics.com). Of the 250 participants recruited, seven were excluded for not meeting the inclusion criteria, leaving a final sample of 243 parents. With respect to gender, the final sample was 48% male and 52% female. With respect to race, the final sample was 5% Asian, 10% Black, 83% White, 1% American Indian, and < 1% other race not listed. With respect to marital status, the final sample was 14% single, 57% married, 20% cohabitating, 7% separated or divorced, 1% widowed, and 1% did not wish to respond. The mean age of the final sample was 35.5 (SD = 8.9) years. The mean educational level was 14.8 (SD = 2.0) years. The mean number of children in the home was 1.5 (SD = .77).

Stimulus Materials and Measures

Refer to Appendix A for a sample battery of measures developed for the proposed study in a format similar to how participants viewed the measures on a computer. All measures in the preliminary study were described in the previous review of Milner et al.’s (2017) experiments.
Demographics

On a demographics questionnaire, parents reported their gender, race, marital status, age in years, educational level in years, and number of children in the home.

Ambiguous Child Photographs

The ambiguous child photographs used in the preliminary study were used in previous research (Farc et al., 2008; Milner et al., 2017). Ambiguous child photographs were desired because according to the SIP model of CPA, negative child-related schemata have their greatest influence on social information processing when child-related stimuli are ambiguous. To establish ambiguity, Farc et al. (2008) asked 56 general-population parents to rate child photographs (including the photographs used in the present study). In Farc et al.’s study, half of the sample of parents rated the likelihood that the child depicted in each photograph was being hostile, and half of the sample of parents rated the likelihood that the child depicted in each photograph was being cooperative. For each child photograph, difference scores between hostility ratings and cooperativeness ratings were computed. The three child photographs that evinced difference scores closest to zero (mean difference score across the three selected child photographs = .01) were defined as ambiguous. One photograph depicted a girl, and the other two photographs each depicted a boy. See Appendix B for the three pictures used in the preliminary study.

Milner et al. (2017) further assessed the ambiguity of the selected child photographs using the Similarity-Intensity Model (SIM; Thompson, Zanna, & Griffin, 1995). The SIM considers both the similarity and intensity of the evaluations to compute estimates of attitudinal
ambivalence. According to the SIM model, attitudinal ambivalence = \((\text{Att}_{\text{pos}} + \text{Att}_{\text{neg}})/2 - |\text{Att}_{\text{pos}} - \text{Att}_{\text{neg}}|\). Attitudinal ambivalence scores for the three pictures used in the Preliminary Study ranged from 1.40 to 2.14 (\(M = 1.74, \text{SD} = 2.01\)).

Positive Child Evaluations

For each of the three ambiguous child photographs, three positively valenced descriptors (i.e., sweet, friendly, cooperative) were rated on a response scale ranging from 0 through 9. The low end of the scale was described using the term “Not at all,” the middle of the scale was described using the term “Somewhat,” and the high end of the scale was described using the term “Extremely.” Responses to the three positive descriptors were averaged across the three ambiguous child photographs to form an index of positive child evaluations (overall internal consistency = .88 for the present sample).

Negative Child Evaluations

For each of the three ambiguous child photographs, three negatively valenced descriptors (i.e., hostile, negative, difficult) were rated on a response scale ranging from 0 through 9. The numerically-low end of the scale was described using the term “Not at all,” the middle of the scale was described using the term “Somewhat,” and the numerically-high end of the scale was described using the term “Extremely.” Responses to the three negative descriptors were averaged across the three ambiguous child photographs to form an index of negative child evaluations (overall internal consistency = .84 for the present sample).
**Ambiguous Child Behaviors**

The ambiguous child behaviors were developed and pre-tested in prior research (Farc et al., 2008) as described earlier. The ambiguous child behaviors were the same used by Milner et al. (2017, Experiment 5 and Experiment 6).

**Attributions of Hostile Intent**

For each of the 18 ambiguous child scenarios, parents were asked, “How likely was his/her behavior done intentionally to annoy/irritate you?” Half of the items used “annoy” and half of the items used “irritate.” Responses were made on a 10-point scale. The numerically low end of the response scale was labeled “Not at all,” the middle of the scale was labeled “Somewhat,” and the numerically high end of the scale was labeled “Extremely.” Responses to the 18 items were averaged to form an index of attributions of hostile intent (for the present sample overall internal consistency = .93).

**Anger**

For each of the 18 ambiguous child scenarios, parents were asked, “To what extent would this behavior make you feel angry?” Responses were made on a 10-point scale. The numerically low end of the response scale was labeled “Not at all,” the middle of the scale was labeled “Somewhat,” and the numerically high end of the scale was labeled “Extremely.” Responses to the 18 items were averaged to form an index of anger (for the present sample overall internal consistency = .95).
Harsh Verbal Discipline

For each of the 18 ambiguous child scenarios, parents were asked, “How likely is it that you would yell/shout/scream at the child because of this behavior?” Responses were made on a 10-point scale. The numerically low end of the response scale was labeled “Not at all,” the middle of the scale was labeled “Somewhat,” and the numerically high end of the scale was labeled “Extremely.” Responses to the 18 items were averaged to form an index of harsh verbal discipline (for the present sample overall internal consistency = .95).

Harsh Physical Discipline

For each of the 18 ambiguous child scenarios, parents were asked, “How likely is it that you would slap/hit/spank the child because of this behavior?” Responses were made on a 10-point scale. The numerically low end of the response scale was labeled “Not at all,” the middle of the scale was labeled “Somewhat,” and the numerically high end of the scale was labeled “Extremely.” Responses to the 18 items were averaged to form an index of harsh physical discipline (for the present sample overall internal consistency = .95).

“Do Nothing” Responses

For each of the 18 ambiguous child scenarios, parents were asked, “How likely is it that you would do nothing because of this behavior?” Responses were made on a 10-point scale. The numerically low end of the response scale was labeled “Not at all,” the middle of the scale was labeled “Somewhat,” and the numerically high end of the scale was labeled “Extremely.”
Responses to the 18 items were averaged to form an index of likelihood to do nothing (for the present sample overall internal consistency = .95).

**Harm**

Harm was measured using a modified version of the Voodoo Doll Task validated for use in parent-to-child aggression research (McCarthy et al., 2016). Specifically, for each of the 18 ambiguous child scenarios, parents were shown a schematic outline of a child image. The instructions above the child outline stated that, if they wished, respondents could inflict harm on the child by clicking on researcher-defined square areas on the child outline. Parents could only click each area once. If parents clicked on an area, the area was highlighted indicating that the parent had selected the area. If parents clicked on an area that had already been selected, the area was unhighlighted, indicating that the parent had unselected the area. Parents could click on up to 23 unique areas on the outline. If parents did not wish to harm the child, parents could click on no area and proceed to the next question by clicking the “Next” button at the bottom-right corner of the screen. See Appendix D for the modified version of the Voodoo Doll Task used in the preliminary study. For each parent, the numbers of clicked (selected) areas on the child outlines for the 18 child-behavior combinations were summed to form an index of harm. Thus, the harm index possible values ranged from zero to 414.
On Amazon’s Mechanical Turk (https://www.mturk.com/mturk/welcome), parents first clicked on a Human Information Task, which contained information about the present study’s procedure. The procedural information on Mechanical Turk was: “In this 30-minute study, we will ask that you rate child images on trait words. Then, you will be asked questions about how you would feel and behave in response to specific behaviors performed by the children depicted in the child images.”

After clicking on a link in the Human Information Task, participants were forwarded to an online survey on Qualtrics (http://www.Qualtrics.com) containing the preliminary study’s measures. Participants were shown an informed consent page (see Appendix E). After participants read the informed consent page, and if they chose to participate, parents proceeded to the online survey. Parents entered a CAPTCHA code to ensure that they were not random “bot” responders and completed questions that assessed inclusion criteria (i.e., that they were parents, had at least one child under the age of 18 living in the home, and spoke English as their primary language). Then, in a random order unique to each participant, parents were shown each of the three ambiguous child photographs. Beneath each photograph, parents responded to items that were used to compute the indices of positive child-related evaluations, negative child-related evaluations, and future discipline expectations. Next, in a random order unique to each participant, participants were shown the child pictures each paired with one of six ambiguous child behaviors (one at a time), for a total of 18 child photograph-behavior combinations. After each child photograph-behavior combination, participants responded to the items that were used to form indices of hostile intent attributions, anger, harsh verbal discipline, harsh physical
discipline, likelihood of doing nothing, and harm. Next, parents completed a measure (i.e., the 
Child Abuse Potential [CAP] Inventory) that was not included in the preliminary study analyses 
but was included as a part of a larger project. After completing the CAP Inventory, parents (a) 
were given a code to enter on Mechanical Turk to receive $4 in compensation, (b) were given a 
space to input any comments or feedback about the study, and (c) were shown debriefing 
information.

Results

Analysis Strategy

To test hypotheses regarding relationships between variables, except for pairs of variables 
involving the harm index, Pearson correlations were computed. For pairs of variables involving 
the harm index, negative binomial simple regression analysis was conducted in which the harm 
index was the dependent variable and in which another variable in the matrix was the predictor 
variable. Consistent with prior research (e.g., McCarthy et al., 2016), negative binomial 
regression was conducted for harm because harm was a count (i.e., discrete integer) variable for 
which many subjects had a value of 0. All hypothesis tests were directional (one-tailed). No 
correction to alpha or \( p \)-values were made to hypothesis tests because (a) the analyses were 
planned and (b) to avoid losing statistical power.

Hypothesis Testing

Contrary to Hypothesis 1, positive child-related evaluations were not significantly 
inversely correlated with attributions of hostile intent, \( r = .12, p = .054 \), 95% CI [-0.24, 0.00];
anger, $r = -.06, p = .388, 95\% \text{ CI } [-0.18, 0.07]$; harsh verbal discipline, $r = .02, p = .796, 95\% \text{ CI } [-0.10, 0.11]$; harsh physical discipline, $r = .03, p = .622, 95\% \text{ CI } [-0.11, 0.15]$; or “do nothing” responses, $r = .01, p = .887, 95\% \text{ CI } [-0.12, 0.14]$. Additionally, positive child-related evaluations did not significantly predict harm, $b = .14, p = .496, RR = 1.15, 95\% \text{ CI } [0.77, 1.73]$.

However, results confirmed the other hypotheses. As predicted in Hypothesis 2, negative child-related evaluations were positively correlated with attributions of hostile intent, $r = .30, p < .001, 95\% \text{ CI } [.18, .41]$; anger, $r = .23, p < .001, 95\% \text{ CI } [.11, .36]$; harsh verbal discipline, $r = .19, p = .004, 95\% \text{ CI } [.07, .32]$; and harsh physical discipline, $r = .13, p = .038, 95\% \text{ CI } [.00, .25]$. Negative child-related evaluations positively predicted harm, $b = .13, p = .025, RR = 1.14, 95\% \text{ CI } [1.02, 1.27]$; for every unit increase on negative child-related evaluations, the increase in the rate of pin usage was equal to 14%.

As predicted in Hypothesis 3, attributions of hostile intent were positively correlated with anger, $r = .74, p < .001, 95\% \text{ CI } [.68, .79]$; harsh verbal discipline, $r = .67, p < .001, 95\% \text{ CI } [.59, .73]$; and harsh physical discipline, $r = .45, p < .001, 95\% \text{ CI } [.34, .54]$. Attritutions of hostile intent positively predicted pin usage, $b = .83, p < .001, RR = 2.29, 95\% \text{ CI } [1.96, 2.67]$; thus, for every one unit increase on attributions of hostile intent, the increase in the rate of pin usage was 129%. Anger was positively correlated with harsh verbal discipline, $r = .88, p < .001, 95\% \text{ CI } [.85, .91]$, and harsh physical discipline, $r = .76, p < .001, 95\% \text{ CI } [.70, .81]$. Anger positively predicted pin usage, $b = 1.00, p < .001, RR = 2.72, 95\% \text{ CI } [2.25, 3.28]$; for every unit increase on anger, the increase in the rate of pin usage was 172%. Harsh verbal discipline was positively correlated with harsh physical discipline, $r = .81, p < .001, 95\% \text{ CI } [.76, .85]$. Harsh verbal discipline positively predicted pin usage, $b = 1.30, p < .001, RR = 3.66, 95\% \text{ CI } [2.95,
for every one unit increase on harsh verbal discipline, the increase in the rate of pin usage was 266%. Harsh physical discipline positively predicted pin usage, $b = 1.14, p < .001, RR = 3.14, 95\% CI [2.49, 3.96]$; for every one unit increase on harsh physical discipline, the increase in the rate of pin usage was 214%.

As predicted in Hypothesis 4, do nothing responses were negatively correlated with attributions of hostile intent, $r = -.40, p < .001, 95\% CI [-.50, -.29]$; anger, $r = -.40, p < .001, 95\% CI [-.50, -.29]$; harsh verbal discipline, $r = -.42, p < .001, 95\% CI [-.52, -.31]$, and harsh physical discipline, $r = -.29, p < .001, 95\% CI [-.40, -.17]$. Also as predicted, do nothing responses negatively predicted pin usage, $b = -.43, p < .001, RR = .65, 95\% CI [.59, .72]$; for every one unit increase on likelihood to do nothing, the decrease in the rate of pin usage was 33%.

Discussion

Evidence obtained from the preliminary study suggests that the measures used were internally consistent and demonstrated the hypothesized cross-sectional associations in a sample of general-population parents. As hypothesized, the following variables were significantly positively correlated: (a) negative child-related evaluations, (b) attributions of hostile intent, (c) anger, (d) use of harsh verbal discipline, (e) use of harsh physical discipline, and (f) harm. Compared to negative child-related evaluations, positive child-related evaluations evinced numerically smaller correlations with attributions of hostile intent, anger, use of harsh verbal discipline, use of harsh physical discipline, and harm.
Although the absence of significant correlations between positive child-related evaluations and other study variables runs contrary to hypotheses, post-hoc this result is consistent with theory and research that standard units of negativity, relative to standard units of positivity, are stronger predictors of trait judgments, evaluations, and behavior (the positive-negative asymmetry; for reviews, see Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Rozin & Royzman, 2001). Rozin and Royzman (2001) described a specific type of positive-negative asymmetry called negativity dominance, a principle that asserts that the “…holistic perception and appraisal of integrated negative and positive events (or objects, individuals, hedonic episodes, personality traits, etc.) is more negative than the algebraic sum of the subjective values of those individual entities” (pp. 298-299). Thus, the absence of correlations between positive child-related evaluations and other study variables may be a manifestation of negativity dominance. See Milner et al. (2017) for regression analyses that are also consistent with this idea. In addition, results from the preliminary study replicated prior research results showing that in general-population parents attributions of hostile intent (e.g., Dix, Ruble, & Zambarano, 1989; Nix et al., 1999; Smith & O’Leary, 1995) and anger (Leung & Slep, 2006) are associated with use of harsh discipline.

However, a potential limitation in the preliminary study with respect to the measures of positive child-related evaluations and negative child-related evaluations is that these measures are not evaluatively pure. Specifically, the trait descriptors that were used to construct the measures contain additional meaning above and beyond mere evaluation (i.e., good or bad; like or dislike). Consider, for example, the trait friendly. Friendly means “kind” or “supportive.” Research has demonstrated that college students who described themselves as more talkative,
humble, trustful, warm, cooperative, or courteous and less dominant and conceited were most liked by others (Wortman & Wood, 2011). Other research has shown that, nomothetically, the trait of friendliness (and traits that are synonyms of friendliness, such as kind) are highly positive and central to the self-concept (Sedikides & Green, 2000). Nonetheless, in principle, even highly friendly or kind individuals can be highly disliked.

For example, ethnographer Elijah Anderson (1994) noted that in poor urban areas, the notion of “respect” is important to individuals because the more “respect” individuals have, the less likely it is that someone on the streets will victimize the individual. According to Anderson, indicators of “respect” include “bling” (i.e., jewelry), stoicism, and vigilantism. Thus, someone who appeared especially gregarious and friendly might be likely to be a target of violence in “the streets” because the individual is naïve to and behaving inconsistently with the “Street Code.”

Another potential limitation is that during the positive EC procedure used in the preliminary study, semantic components of the measures of positive child-related evaluations and negative child-related evaluations are activated. Thus, instead of putatively increasing positive evaluations through various standard cognitive mechanisms that are postulated to underlie EC effects (e.g., the formation or modification of evaluative associations in memory; see Hofmann et al., 2010, for a review), it is possible that the EC procedure used in the preliminary study simply activated cognitive representations of the traits implied by the descriptor CS, which then influenced responses on the post-EC outcome measures (i.e., conditioning effects were produced by a trait learning or semantic priming mechanism).

If priming is a plausible explanation for observed EC effects in the studies proposed in my dissertation, then EC’s pragmatic utility to reduce use of harsh discipline and harm in parents
at high risk for CPA may be reduced. As mentioned earlier, priming effects dissipate across time (Srull & Wyer, 1979). Such fleeting effects are barriers to goals for maintaining reductions in high-risk parents’ negative child-related evaluations and subsequent use of harsh discipline and harm.

One might think that this problem could be bypassed by directly assessing the effect of the positive EC procedure on the extent to which the CS is liked. I did so in Study 1 of the present dissertation. This allows for testing the possibility that EC primes constructs related to the UCSs (i.e., the concepts sweet, friendly, and cooperative) without increasing parental liking toward children. However, this idea is not foolproof. It is plausible that positive EC (a) primes the concepts sweet, friendly, and cooperative, and (b) conscious introspection when asked to respond to the parental liking item may result in parents inferring that the CS is also highly likable because the parents have learned that the child is sweet, friendly, and cooperative.

However, results from Milner et al. (2017, Experiment 3) mitigate the plausibility that such priming effects would influence a liking measure in this way. Specifically, priming theory posits that after activation of the relevant construct, activation decreases with time. Thus, assuming that one week is sufficient for any positive EC-induced semantic priming (i.e., construct activation as a result of viewing positive EC) to dissipate, priming theory cannot easily account for the positive EC effects observed by Milner et al. at 1 week follow-up (for the CS, if not for the exemplar ambiguous child photographs). Instead, the positive EC effects found after a week’s lag are consistent with (and thus plausibly explained by) the putative standard mechanisms thought to underlie EC, such as the formation or modification of evaluative associations in memory. Such associations are thought to be temporally stable.
Despite the liking measure’s mentioned limitations, another reason to include a measure of parental liking toward children is that dislike of children is theorized to be a component of negative child-related schemata in high-risk parents. Finally, parental liking toward children is highly congruent with definitions of EC as a change in liking for the CS consistent with the valence of the US (e.g., De Houwer, 2007). Thus, including a measure of parental liking toward children contributes to my dissertation research in at least two ways: (1) Dislike of children may be a component of negative child-related schemata in high-risk parents; (2) parental liking toward children closely maps onto the definition of what EC is theorized to change (i.e., liking).

Although priming may be one cognitive mediator of observed EC effects, as discussed previously, several theories postulate other alternative mechanisms for EC effects. One of the major aims of the Hofmann et al. (2010) meta-analyses was to investigate which theory or theories were most supported by or consistent with the meta-analyzed EC research results. Hofmann et al. provided a cursory review of major theories of EC (see Hofmann et al., 2010, pp. 391-394). A single experiment (or set of experiments) will be unable to determine, among all existing EC theories, any theory that is “true” or “most correct.” Nonetheless, Study 1 tested hypotheses derived from three competing EC theories: (1) the referential account of EC (e.g., Baeyens, Eelen, & Van den Bergh, 1990), (2) the propositional account of EC (e.g., Mitchell, De Houwer, & Lovibond, 2009a; Mitchell et al., 2009b), and (3) the affective implicit misattribution account of EC (Jones, Fazio, & Olson, 2009).

According to the referential account of EC, EC is mediated by associations formed in memory that link two constructs. Hofmann et al. (2010) noted that, unlike the type of associations that are thought to underlie classical Pavlovian conditioning or expectancy learning
(i.e., CS elicits the prediction or expectation of the US), the referential account of EC posits that the associations between the linked constructs are merely referential (i.e., CS merely refers to the US). According to Hofmann et al., primary hypotheses derived from the referential account of EC include: (a) EC increases with number of CS-US co-occurrences; (b) EC occurs regardless of contingency awareness (i.e., “…explicit awareness of CS-US contingencies is not necessary for EC to occur,” Hofmann et al., 2010, p. 393); (c) EC is resistant to extinction (i.e., after EC, presentations of CS in the absence of US do not attenuate EC-induced evaluations of the CS).

According to Hofmann et al., the main distinction between the associative mechanisms underlying EC and classical (Pavlovian) conditioning is that EC is the formation of referential relations in memory, whereas classical conditioning is the formation of predictive associations (i.e., signal learning) in memory.

According to the propositional account of EC, EC is mediated by the formation of propositions about the CS-US relation, which in turn “can function as a justification for determining liking of the CS.” The propositional account of EC assumes that participants must consciously recognize, at minimum, that the CS is related to the US (i.e., participants must form a proposition relating the US to the CS). The propositional account of EC also posits that the content of propositional knowledge relating the CS to US can determine the direction of valence change after pairing regardless of the valence of the US.

Evidence has accumulated (e.g., Hofmann et al., 2010; Moran, Bar-Anan, & Nosek, 2016) that suggests that both the referential account of EC and the propositional account of EC are prima facie plausible. Moran et al. (2016) noted when people view two affective stimuli that frequently co-occur, these affective stimuli may be related explicitly by propositions. For
example, viewers of media that depict the fictional superhero Batman cannot deny that Batman frequently co-occurs with crime. Yet, do consumers of Batman media dislike Batman as a result of the frequent co-occurrences of Batman and crime? Presumably the answer is “no”, because we know that Batman fights crime. It is the proposition, “Batman fights crime” that is theoretically relevant to the debate about whether the referential or the propositional account of EC is more consistent with EC research results. Results from five experiments (Moran et al., 2016; total N = 505) suggest that co-occurrence produces EC effects even when explicit relational information suggests that the evaluation of the CS should be the opposite valence of the US. One implication is that both (a) mechanisms related to the referential account of EC and (b) mechanisms related to the propositional account of EC are consistent with EC research results.

The affective implicit misattribution account (Jones et al., 2009) posits that the valence elicited by a stimulus during EC may “transfer” to the other stimulus via the perceiver’s incorrect and implicit (i.e., out of conscious awareness) attribution of the valenced stimulus’s affect to the other stimulus. Jones et al. identified three broad conditions that increase the likelihood that affect from one stimulus will be implicitly misattributed to the other stimulus in EC: (1) the affect evoked by a stimulus could plausibly have come from the other stimulus, (2) the source of the elicited affect is ambiguous, and (3) close spatial-temporal continuity or proximity between stimuli. Researchers (e.g., Hofmann et al., 2010) point out that an implication of the implicit misattribution account is that a “backward conditioning” procedure in which the US temporally/spatially precedes the CS should produce EC effects (via implicit misattribution).
Statement of the Problem

Milner et al. (2017) provided substantial evidence to support the hypothesis that positive EC increases positive child evaluations and reduces parents’ negative child evaluations, child-related attributions of hostile intent, anger, harsh discipline, and harm. Several mechanistic theories (e.g., the previously described referential account, the propositional account, and the implicit affective misattribution account) strive to explain EC effects observed in the research literature. A corpus of research (e.g., Hofmann et al., 2010; see also in a special 2016 issue of Social Cognition, De Houwer & Hughes, 2016; Gawronski, Balas, & Hu, 2016; Halbeisen & Walther, 2016; Hüttler & Fiedler, 2016; Moran, Bar-Anan, & Nosek, 2016; Stahl & Heycke, 2016; Unkelbach & Fiedler, 2016) has emerged testing these theories. More research is needed to investigate mechanisms driving EC effects because understanding underlying mechanisms is important both (a) for testing theory and (b) to understand those functional aspects of EC that might be changed to optimize its effectiveness and pragmatic utility.

Overview of Dissertation Studies

The present dissertation describes two studies that tested potential cognitive mechanisms underlying EC effects observed by Milner et al. (2017). In the usual EC paradigm, the CS preceded the UCS (this ordering is called forward conditioning). Study 1 used this procedure but also used an alternative procedure in which the UCS preceded the CS (this ordering is called backward conditioning). There are three “standard” theories (reviewed in more detail below) that might account for the EC effects (the referential account mechanism, the propositional
account mechanisms, and the implicit misattribution account mechanism) that emerged from the forward conditioning procedures used in Milner et al. One of these accounts (referential) is incompatible with EC effects using the backward conditioning procedure used in Study 1. Hence, the emergence of an EC effect in a backward conditioning EC paradigm would suggest that the results reported by Milner et al. (2017) could not be fully accounted for in the referential view of EC. To minimize the possibility that semantic (rather than evaluative) learning may have played a role in producing the previous EC research results reported by Milner et al. (2017), Study 2 employed an EC procedure in which the unconditioned stimuli that are relatively free of semantic content (positive emojis and “neutral” emojis). The emergence of an EC effect in such circumstances would argue against an interpretation of the Milner et al. results that relies on semantic learning or semantic priming.
CHAPTER 3

STUDY 1

Study 1 sought information about whether the EC effect observed by Milner et al. (2017) depends on the temporal ordering of the CS and the US. Accordingly, Study 1 employed three conditions: forward positive EC condition, backward positive EC condition, and a forward pseudo-EC (control) condition. Employing these experimental conditions afforded the ability to test hypotheses derived from competing theories regarding cognitive mechanisms that underpin any observed EC effects in the present study.

In Study 1, competing EC theories predicted divergent results. The EC theories are the referential account, propositional account, and the affective implicit misattribution account. According to the referential account (e.g., Baeyens, Eelen, Crombez, & Van den Bergh, 1992), EC effects are mediated by forming referential relations between the CS and US, such that the CS becomes a signal (or a referent) for the US. The referential account implies that the temporal order of the EC stimuli matter: forward conditioning is required to obtain EC effects. The propositional account (Mitchell, De Houwer, & Lovibond, 2009a; Mitchell, De Houwer, & Lovibond, 2009b) posits that EC effects are mediated (a) by the retrieval of propositions relating the CS to US (e.g., “Children are good” or “Children are positive”) and (b) subsequent truth evaluation of the propositions.
More specifically, Mitchell et al. (2009a, 2009b) conjecture that propositions may be stored in long-term memory as representations and retrieved from long-term storage by automatic processes (i.e., processes that operate in the absence of conscious awareness or control). Repeated CS-US pairings activate stored propositions relating the two stimuli, and subsequent truth valuation of the activated propositions manifest as EC effects. Thus, the propositional account of EC predicts that, with the exception of the forward pseudo-EC condition, the forward positive EC condition and the backward positive EC condition will show hypothesized EC effects because the EC procedure in these conditions activate the components of stored propositions relating the CS to US. According to this view, subsequent truth evaluation of these retrieved and constructed propositions based on the EC pairings will manifest in the hypothesized EC effects. The affective implicit misattribution account (Jones et al., 2009) posits that the valence elicited by a stimulus during EC may be misattributed to the other stimulus via the perceiver’s incorrect and implicit (i.e., out of conscious awareness) misattribution of the valenced stimulus’s affect to the other stimulus.

First it is hypothesized that, pre-EC, positive child-related evaluations will be negatively correlated with attributions of hostile intent, anger, harsh verbal discipline, harsh physical discipline, harm, and doing nothing.

It is hypothesized that, pre-EC, liking will be positively correlated with positive child evaluations and negatively correlated with negative child evaluations.

It is hypothesized that, pre-EC, liking will be negatively correlated with attributions of hostile intent, anger, harsh verbal discipline, harsh physical discipline, harm, and doing nothing.
It is hypothesized that, pre-EC, negative child-related evaluations will be positively correlated with attributions of hostile intent, anger, harsh verbal discipline, harsh physical discipline, harm, and doing nothing.

It is hypothesized that, pre-EC, the following responses to ambiguous child behaviors will be positively correlated: attributions of hostile intent, anger, harsh verbal discipline, harsh physical discipline, and harm.

It is hypothesized that, pre-EC, “doing nothing” in response to the ambiguous child behaviors will be negatively correlated with attributions of hostile intent, anger, harsh verbal discipline, harsh physical discipline, and harming the child.

Based on the results of Milner et al. (2017), it was hypothesized that, in the forward positive EC condition, change scores representing hypothesized (directional) EC effects will be significantly greater than (a) the change scores in the forward pseudo-EC condition and (b) zero. Based on the competing EC theories described above, two competing hypotheses were tested. Data from Study 1 will favor the referential account over the propositional account if EC effects are present in the forward positive EC condition and absent in both the backward positive EC condition and the forward pseudo-EC condition. Data will be inconsistent with both the implicit misattribution account of EC and the propositional account of EC if only the forward conditioning condition is effective at changing dependent measures in the hypothesized directions.

When constructing the backward conditioning procedure, the inter-trial interval must be sufficiently long so that forward conditioning is highly implausible as an alternative explanation to potential EC effects in the backward conditioning condition. Specifically, the inter-trial
interval for Study 1, programmed to be 500 ms, might not provide enough temporal distance between trials to rule out the possibility that forward conditioning effects might occur in the backward conditioning condition. Thus, observing a significant EC effect in the backward conditioning condition would be ambiguous with respect to whether backward conditioning or forward conditioning mediated the EC effect. Increasing the inter-trial interval to 12 s makes it highly implausible that a significant EC effect in the backward conditioning condition can be caused by forward conditioning processes. However, with an inter-trial interval of 12 s, keeping all other procedural details of the EC procedures similar to the procedures in Milner et al. (2017) became impractical (i.e., participants would have been watching EC for over 25 minutes). Instead, in all EC conditions, the number of trials (18) was set to keep the EC procedure at a similar duration (i.e., just under 4 minutes) as the EC procedures in the Milner et al. studies.

Method

Participants

Participants were a convenience group of 77 undergraduates at Northern Illinois University. Participants were randomly assigned to the three conditions (i.e., forward positive EC, backward positive EC, or forward pseudo-EC) using a computer program that allowed for a “blind” procedure: The experimenter and participant did not know the EC condition to which the participant was assigned until after the study.

The sample was 49% male and 51% female. The sample ethnicity was 69% White, 18% Black, 6% Asian, 3% Multi-racial and 4% “Other race not listed.” The sample marital status was 91% single, 3% married, 3% cohabiting, and 3% “not listed/do not wish to respond.” Twelve
percent of the sample identified as having Hispanic origin. The mean age was 20.1 ($SD = 2.9$) years. The mean educational level was $13.4$ ($SD = 1.6$) years. Six participants (8%) were parents (three participants had one child and three participants had two children).

**Power Analysis and Sample Size Planning**

Given the multiple experimental conditions and hypotheses tested in Study 1, insufficient information was available to reasonably estimate the effect size for power analyses. Instead, for simplicity and practicality, 25 participants were included in each of the experimental conditions. Specific contrasts (described later) were specified and tested in Study 1. Given these and other constraints, (e.g., the desired Type I/Type II error rate ratio), sensitivity power analyses were conducted. These calculations suggested that Study 1 was sufficiently powered to detect a $d_z$ (or differences in $d_z$) that were .28 and larger.

**Measures and Stimuli**

Except for the addition of the liking measure and the nonsensical USs in the forward pseudo-EC condition (described below), the measures and stimuli were the same as the measures and stimuli used in the Preliminary Study. In Study 1, Cronbach’s $\alpha$s for each dependent measure (pre-EC) were computed and are as follows: positive child evaluations $\alpha = .83$, negative child evaluations $\alpha = .71$, future discipline expectations $\alpha = .73$, attributions of hostile intent $\alpha = .91$, anger $\alpha = .94$, harsh verbal discipline $\alpha = .95$, harsh physical discipline $\alpha = .97$, doing nothing $\alpha = .96$. 
Liking toward children. Liking toward children was measured by responses to the item, “If you met this child, to what extent do you think you would like the child?” The low end of the response scale was labeled “Not at all,” the middle of the scale was labeled “Somewhat,” and the high end of the scale was labeled “Extremely.” Responses across the three ambiguous child images were averaged to form the measure of liking toward children. Cronbach’s α for the pre-EC liking measure was .83.

Procedure

In a random order unique to each participant, participants encountered each of the three ambiguous child photographs. While viewing each ambiguous child photograph, participants were asked to complete measures of positive child evaluations, negative child evaluations, and liking toward children (this last measure was not used in the preliminary study). Next, in a random order unique to each participant, participants saw the ambiguous child photographs each paired with one of six ambiguous child behaviors (one at a time), for a total of 18 ambiguous child behaviors. While viewing each ambiguous child scenario, participants were asked to complete the measures of attributions of hostile intent, anger, harsh verbal discipline, harsh physical discipline, ignoring, and harm.

Next, participants were randomly assigned to one of three conditions: forward positive EC, backward positive EC, or forward pseudo-EC. In the positive EC condition, each trial consisted of a focal point (i.e., plus sign) on a computer screen followed by one of the three ambiguous child photographs (CS). One of three positive descriptors (USs: sweet, friendly, cooperative) followed each ambiguous child photograph. Thus, nine pairings were presented as trials in the EC procedure. Each of the nine pairings appeared twice, so 18 pairings were shown
to each participant. Computer software (DirectRT, 2016) was programmed to present the focal point for 1000 ms. The child photograph and the positive descriptors each were programmed to be displayed for 20 ms each.

The plus sign and the positive descriptors presented on the computer screen (diagonal 48.5 cm) were displayed in black 24-point Times New Roman font (all capital letters). The computer screen resolution was 1024 × 768 pixels, and DirectRT was set to display at the computer screen’s resolution. Ambiguous child photographs were 489 × 606 pixels. Thus, the ambiguous child photographs consumed 48.8% of the computer screen’s horizontal axis and 78.9% of the computer screen’s vertical axis. Instead of 180 trials used in Milner et al. (2017), Study 1’s EC conditions contained only 18 trials because the Study 1 inter-trial interval was programmed to be 12 s. This inter-trial interval was necessary to prevent inadvertent forward EC effects from occurring in the backward conditioning condition.

Except for the replacement of the USs with nonsensical letter strings matched in number of letters, the pseudo-EC condition was identical to the forward positive EC condition. Except for switching temporal order of CS and US in the backward positive EC condition, the forward positive EC condition was identical to the backward positive EC condition.

Post-EC, participants completed the same measures that they completed pre-EC. Finally participants (a) were compensated with 2 research credits (1 credit per 30 min; estimated study completion time = 45 min), (b) were thanked and debriefed, and (c) were asked for feedback about the study (feedback from participants was not reported in the present study).
Analytic Strategy

Hypothesis Testing

To test hypotheses in Study 1, change scores were computed. These change scores were computed such that positive change scores indicated EC effects in the hypothesized direction. For example, for positive child evaluations the change scores were computed by subtracting pre-EC scores from post-EC scores, whereas for negative child evaluations the change scores were computed by subtracting post-EC scores from pre-EC scores. EC within-condition mean change scores were subjected to two sets (where a set consists of the same contrast separately for the seven dependent variables measured pre- and post-EC) of a priori (non-orthogonal) tests. One set of contrasts (C1) compared the forward conditioning condition to the forward pseudo-conditioning condition (hypothesis: forward conditioning condition > forward pseudo-conditioning condition). The second set of a priori contrasts (C2) compared the backward conditioning condition to the forward pseudo conditioning condition (hypothesis: backward conditioning condition > forward pseudo conditioning condition). If neither C1 or C2 were significant, data would be inconclusive with respect to which theory (referential, affective implicit misattribution, propositional) explains EC effects because no EC effects would be obtained. If C1 was significant whereas C2 was not significant, the data would be more consistent with the referential account than the propositional account and misattribution account. If both C1 and C2 were significant, data will be more consistent with the propositional account and misattribution account than the referential account.
To test hypotheses, two datasets were constructed for Study 1: a “full” dataset and a “reduced” dataset. The full dataset included all cases having usable data and included into the final sample based on inclusion/exclusion criteria. The reduced dataset was constructed (a) by recoding as “missing” floor values on pre-EC measures that were hypothesized to decrease pre-to post-EC and (b) recoding as “missing” ceiling values on pre-EC measures that were hypothesized to increase pre- to post-EC. Participants with missing values in the dataset were excluded from analyses performed on the reduced dataset. Thus, analyses performed on the reduced dataset were not contaminated by floor or ceiling effects but were smaller in sample size than the full dataset. Accordingly, assuming participants were not exhibiting an invalid response set (e.g., acquiescence) it was expected that relative to analyses on the full dataset, analyses on the reduced dataset will have lower estimation precision. However, it is expected that the reduced data set (vs. the full data set) will yield larger EC effect estimates (because the expected increase in hypothesized mean differences were expected to be larger in magnitude than the expected decrease in precision of estimation).

This analytic strategy used to test planned contrasts was applied to both the full dataset and the reduced dataset. All planned comparisons applied to each data set were unprotected against multiple comparison-induced Type I error inflations (to optimize statistical power for the reduced dataset analyses, given the design and analyzed $n$).
Planned Exploratory Analyses

Analysis of Experimental Design Using Repeated-Measures ANOVA

To examine fully the pattern of data, each dependent variable was subjected to separate EC Condition (forward positive EC, backward positive EC, forward pseudo-EC) × Time (pre-EC, post-EC) ANOVAs with repeated measures on the second variable. Where there were significant effects of EC condition, Student-Neuman-Keuls (SNK) follow-up tests (α = .05) were conducted to explore differences among means. Where there were significant effects of time (pre-EC, post-EC), paired-samples $t$ tests were conducted.

I conducted ANOVAs to examine all possible effects in the experimental design for complete reporting and exploratory purposes. However, the $F$ tests will not be as informative or statistically powerful as the directional $t$ tests used for hypothesis testing because $F$ tests are inherently non-directional; variances cannot be negative, and there is only one tail of the $F$ distribution given $df_1, df_2$ (see The 20% Statistician, 2016). Use of the $F$ ratio as a variance ratio test allows the use of the $F$ distribution to make inferences about cases where the variance in one group is less than the variance in another group, but the $F$ tests in this study are not variance ratio tests in this sense.

With the exceptions that (1) within the forward positive EC condition and backward positive EC condition I used directional $t$ tests to gauge the magnitude of within-condition “EC effects” and precision of their estimation; and (b) SNK tests were used to follow up significant EC Condition simple effects, follow-up diagnostic tests between means were unprotected and non-directional.
One alternative to the analysis of difference scores is to conduct analyses in which the each participant’s post-EC scores are adjusted for the pre-EC scores. Accordingly, for each dependent variable, I used analysis of covariance (ANCOVA) to explore the post-EC condition differences estimated at the mean of pre-EC scores. The ANCOVAs were “main effects” models; i.e., interactions between the covariate and EC condition were not modeled (in part because none were expected and in part because additional exploratory analyses [described later] examined the linear associations between covariates and EC change scores within the forward positive EC condition and backward positive EC condition). I would gain additional confidence in the results of my difference score analyses if they were confirmed by the results from the ANCOVA analyses.

In previous research examining EC effects in general population parents (Milner et al., 2017), for each dependent variable the linear relationship between pre-EC scores and EC change scores were examined within the positive EC condition. Pre-EC scores on dependent variables might be positively related to EC change scores (computed so that positive scores indicate change in the hypothesized direction). A positive relationship between pre-EC scores and EC change scores may reflect the fact that higher scores have more room on the variable scale to change.
Alternatively, a positive relationship between pre-EC scores and EC change scores might indicate that a number of participants endorsed values on dependent variables that cannot change (i.e., “floor” or “ceiling” effects, depending on the dependent variable). A negative relationship between pre-EC scores and EC change scores might indicate, for example, that participant thoughts, feelings, and behaviors support and reinforce the processes that contribute to the extreme scores. However, these possible explanations for correlations between pre-EC scores and EC change scores were untestable in the present study.

**Linear Relationships Between Demographic Characteristics and EC Change Scores Within the Forward Positive EC Condition and the Backward Positive EC Condition**

In previous research examining EC effects in general population parents (Milner et al., 2017), EC change scores for dependent variables (computed by subtracting pre-EC scores from post-EC scores) were not normally distributed. Some parents evinced large EC effects, whereas other parents evinced no EC effects. This pattern resulted in bimodal distributions for EC change scores. In an attempt to explain the bimodal distributions Milner et al. (2017) explored whether demographics (i.e., age, educational level, race, gender, number of children, marital status) were related to EC effect magnitude. Milner et al. found that demographic characteristics were not associated with EC change scores. Nonetheless, in the present study the associations between demographic characteristics and EC change scores were again explored. Age in years, educational level in years, and number of children were continuous variables, whereas gender, marital status, and race were categorical variables. Correlations were computed to examine the linear relationships between the continuous demographic characteristics (i.e., age, educational
level, and number of children) and EC change scores. Using separate one-way ANOVAs, I explored mean differences in EC effects (i.e., EC change scores) between genders, between ethnic categories, and between marital status categories.

**Binomial Tests for Harm**

It was desirable to establish that EC did not have the undesired effect of increasing harm pre- to post-EC for individuals who did not evince harm pre-EC. To explore this idea, a binomial test was conducted in which the proportion of individuals who used at least one pin post-EC out of the number of individuals who used no harm pre-EC was tested against an expected value of zero. A significant effect would indicate that for individuals who did not harm pre-EC, EC increases harm post-EC. Such a result would raise concerns regarding EC’s utility in reducing harm. Given that the absence of a within-condition EC effect suggests that the EC procedure does not produce alterations in participants’ cognitions about children and evaluations of children, I will conduct the binomial test only within those conditions that evince significant EC effects.

**Multi-Level Model Analyses**

Subsidiary multi-level model (MLM) analyses were conducted so that effects could be examined and interpreted at the item level (as opposed to aggregate, between-subjects level). It was hoped that results from such an analysis would establish that effects occurred across all stimuli used in the study.
To perform such analyses, the data was structured so that, separately for each dependent measure (i.e., positive child evaluations, liking, negative child evaluations, expectations for the future discipline, attributions of hostile intent, anger, harsh verbal behavior, harsh physical behavior, doing nothing, and harm), the scores for the 18 ambiguous child scenarios each served as units of analysis nested within person. In each level-1 equation (i.e., the trial-level equation), the predicted variable was each of the dependent measures, and the predictor variable(s) differed as a function of the models (described below).

For each dependent variable, to determine the appropriate model from which to estimate parameters of theoretical interest I used a sequential model-building approach. In all models, the level-1 residual variance-covariance matrix was estimated by the data, which eliminated many statistical assumptions. However, doing so overfits the model to data, which limits generalizability of the model. To build a model that provided the best fit to the data with the fewest number of estimated parameters (relative to a model in which all possible parameters are estimated, including random variances and covariances), nested models of increasing complexity (and decreasing parsimony) were tested hierarchically in the following way:

Model 1 (i.e., the level-2 intercept-only model) was the most parsimonious model tested:

\[
\text{Level 1 Equation: } Y_{ij} = \beta_{0j} + \epsilon_{ij} \\
\text{Level 2 Equation: } \beta_{0j} = \gamma_{00} + \tau_{0j}
\]

where \(Y_i\) represents the score on the dependent measure (e.g., positive evaluations) at time of measure \(i\) for person \(j\); \(\beta_{0i}\), the intercept in the level-1 equation, is predicted (at level 2) by the
overall grand mean for the dependent measure across persons (i.e., the fixed effect of the intercept, $\gamma_{00}$) plus between-person deviations ($\tau_{0j}$) from the overall grand mean.

Model 2 (i.e., time main-effect model) was

Level 1 Equation:  
$$Y_{ij} = \beta_{0j} + \beta_{time}(\text{Time, coded pre-EC = 0, post-EC = 1}) + \epsilon_{ij}$$

Level 2 Equations:  
$$\beta_{0j} = \gamma_{00} + \tau_{0j}$$
$$\beta_{timej} = \gamma_{timej} + \tau_{timej}$$

The level-1 predictor of time was dummy-coded (pre-EC = 0, post-EC = 1) so that the level-1 intercept predicted at level 2 can be interpreted as unadjusted person means on the dependent measure pre-EC. At level 2, the intercept at level 1 was predicted by the intercept at level 2 (i.e., fixed effect for the intercept denoted $\gamma_{00}$, which will be interpreted as the unadjusted person mean for the dependent measure pre-EC) plus a random effect (denoted $\tau_{00}$, which can be interpreted as the variance attributed to sampling from different distributions of persons). The level-1 coefficient for time was predicted at level 2 by the fixed effect of time (denoted $\gamma_{timej}$) plus a random effect (denoted $\tau_{timej}$).

Model 3 (i.e., conditional time-effect model) is

Level 1 Equation:  
$$Y_{ij} = \beta_{0j} + \beta_{time}(\text{Time, coded pre-EC = 0, post-EC = 1}) + \epsilon_{ij}$$

Level 2 Equations:  
$$\beta_{0j} = \gamma_{00} + \tau_{0j}$$
$$\beta_{timej} = \gamma_{timej} + \gamma_{\text{forward conditioning positive EC condition}} + \gamma_{\text{backward conditioning condition}} + \tau_{timej}$$
The level-1 predictor for time was dummy-coded (pre-EC = 0, post-EC = 1) so that the level-1 intercept predicted at level 2 can be interpreted as unadjusted person means on each of the dependent measures pre-EC. At level 2, the intercept at level 1 was predicted by the intercept at level 2 plus a random effect (denoted $\tau_{00}$, which can be interpreted as the variance attributed to sampling from different distributions of persons). The level-1 coefficient for time was predicted at level 2 by the fixed effect of time (denoted $\gamma_{timej}$), plus the fixed effect of forward positive EC, plus the fixed effect of backward positive EC, plus a random effect (denoted $\tau_{timej}$). EC condition variables were not modeled to predict the intercept at level 1 because random assignment was used to assign level-2 units (i.e., persons) to EC condition (expected parameter between EC condition and other predictors = 0).

To minimize the number of sequences during the sequential model-building, Model 2 estimated random effects for the intercept and for the effect of time (i.e., both the pre-EC scores for persons and the difference in means between pre-EC and post-EC on the dependent measure are modeled to vary between-persons). Similarly, the decisions to include the parameters as described in Model 2 and Model 3 were motivated by similar reasoning (i.e., minimizing sequences in the model-building approach while building meaningfully-different models of theoretical interest). These decisions were guided, in part, by desires to maintain model parsimony; in addition, see Bates, Kleigl, Vasishth, and Baayen (2015) for arguments and evidence that overfitting MLM models to data reduces statistical power (given a desired Type I error rate).

To test whether Model 2 is a significantly better fit to the data than Model 1, likelihood ratio tests were conducted. Because parameters were estimated with restricted maximum
likelihood procedures, to test the hypothesis that Model 2 provides a significantly better fit to the data than Model 1, a difference test of the log likelihood of the two models was conducted.

The MLM analyses were conducted for each dependent measure (i.e., positive child evaluations, negative child evaluations, liking, future discipline expectations, attributions of hostile intent, anger, harsh verbal discipline, harsh physical discipline, doing nothing, and harm). For harm, to model the discrete count data (i.e., all zero and positive integers up to and including 23, the number of possible pins for each ambiguous child scenario), the negative binomial distribution (instead of Gaussian normal) was assumed and linked by a logit function. This approach mirrors the approach used by McCarthy et al. (2016) to analyze “pins” given during their versions of the VDT. However, analyses on the harm variable were untenable due to complications arising from the use of the LaPlace approximation handler. Many parameter estimates for this variable were outside the parameter space.

Results

See Appendix H for descriptive statistics for raw data by experimental cell.

Hypothesis Testing

In the hypothesis testing reported below, for each dependent variable the change scores were computed so that positive scores indicate change in the direction hypothesized for the (forward and backward) positive EC conditions. The change scores were computed in this manner for all participants, not just the participants in the forward positive EC condition or
participants in the backward positive EC condition. See Appendix I for descriptive statistics for change score data by experimental cell.

All hypothesis tests involving pre-EC correlations, C1, C2, and within-positive EC-condition one-sample \( t \) tests were directional (one-tailed \( p \)-values). For pre-EC correlational hypotheses, I report the \( r, p \), and one-sided 95% confidence interval. For each hypothesis test involving C1 and C2, I report the results of the \( t \)-test (i.e., \( t, p \)), mean difference of mean change scores between compared EC conditions (abbreviated \( M_{\text{diff}} \)), one-sided 95% confidence interval (which corresponds to the information in the directional hypothesis test). In addition, for the forward positive EC condition and backward positive EC condition, I report the result of the directional one-sample \( t \)-test, mean and standard deviation of change scores, and one-sided 95% confidence interval. For the forward pseudo-EC condition, I report the result of the non-directional (two-tailed) one-sample \( t \)-test, mean and standard deviation of change scores, and 95% confidence interval.

**Pre-EC Correlational Hypotheses**

Contrary to hypotheses, pre-EC, positive child evaluations were not significantly negatively correlated with attributions of hostile intent, \( r = -.10, p = .18, 95\% \text{ CI}[-1.0, .09]; \) anger, \( r = -.02, p = .285, 95\% \text{ CI}[-1.0, .12]; \) harsh verbal discipline, \( r = -.02, p = .425, 95\% \text{ CI}[-1.0, .17]; \) harsh physical discipline, \( r = .04, p = .37, 95\% \text{ CI}[-1.0, .23]; \) harm, \( b = -.00, p = .99, RR = 1.00, 95\% \text{ CI} [.64, 1.60]; \) and doing nothing, \( r = -.05, p = .34, 95\% \text{ CI}[-1.0, .14]. \)
As hypothesized, pre-EC, liking was significantly positively correlated with positive child evaluations, $r = .75, p < .001, 95\% \text{ CI}[.66, 1.0]$. As hypothesized, pre-EC, liking was significantly negatively correlated with negative child evaluations, $r = -.23, p = .0215, 95\% \text{ CI}[-1.0, .04]$; attributions of hostile intent, $r = -.19, p = .047, 95\% \text{ CI}[-1.0, .00]$. Contrary to hypotheses, pre-EC liking was not significantly negatively correlated with anger, $r = -.17, p = .07, 95\% \text{ CI}[-1.0, .02]$; harsh verbal discipline, $r = -.09, p = .22, 95\% \text{ CI}[-1.0, .10]$; harsh physical discipline, $r = -.08, p = .24, 95\% \text{ CI}[-1.0, .11]$; doing nothing, $r = -.01, p = .475, 95\% \text{ CI}[-1.0, .18]$; and harm, $b = -.31, p = .065, RR = .73, 95\% \text{ CI}[,50, 1.07]$.

As hypothesized, pre-EC, negative child evaluations were significantly positively correlated with attributions of hostile intent, $r = .39, p < .001, 95\% \text{ CI}[.22, 1.0]$; anger, $r = .39, p < .001, 95\% \text{ CI}[.22 , 1.0]$; harsh verbal discipline, $r = .34, p = .001, 95\% \text{ CI}[.17, 1.0]$; harsh physical discipline, $r = .22, p = .027, 95\% \text{ CI}[.03, 1.0]$. Contrary to hypotheses, pre-EC, negative child evaluations were not significantly positively correlated with doing nothing, $r = -.04, p = .38, 95\% \text{ CI}[-.22, 1.0]$; and harm, $b = -.05, p = .41, RR = .94 95\% \text{ CI}[.56, 1.57]$.

As hypothesized, pre-EC, attributions of hostile intent were significantly positively correlated with anger, $r = .60, p < .001, 95\% \text{ CI}[.46, 1.0]$; harsh verbal discipline, $r = .54, p < .001, 95\% \text{ CI}[.39, 1.0]$; harsh physical discipline, $r = .26, p = .012, 95\% \text{ CI}[.07, 1.0]$. Contrary to hypotheses, pre-EC, attributions of hostile intent were not significantly positively predictive of harm, $b = .17, RR = 1.18, p = .220, 95\% \text{ CI}[,76, 1.78]$.

As hypothesized, pre-EC, anger was significantly positively correlated with harsh verbal discipline, $r = .79, p < .001, 95\% \text{ CI}[.71, 1.0]$; and harsh physical discipline, $r = .51, p < .001,
95% CI[.35, 1.0]. Contrary to hypotheses, pre-EC, anger was not positively significantly predictive of harm, $b = .26, p = .26$, $RR = 1.30$, 95% CI[.78, 2.0].

As hypothesized, pre-EC, harsh verbal discipline was significantly positively correlated with harsh physical discipline, $r = .62, p < .001$, 95% CI[.49, 1.0]; and significantly positively predictive of harm, $b = .37, p = .04$, $RR = 1.44$, 95% CI[.90, 2.09].

As hypothesized, pre-EC, harsh physical discipline was significantly positively predictive of harm, $b = .52, p = .005$, $RR = 1.67$, 95% CI[1.01, 2.34].

**Positive Child Evaluations**

**Full dataset.** C1 compared positive child evaluation change scores between the forward positive EC condition and the forward pseudo-EC condition. As hypothesized, C1 was significant, $t(50) = 3.00, p = .002$, $M_{\text{diff}} = 1.33, SE = 0.44; 95\% \text{ CI}[0.59, \infty]$. As hypothesized, in the forward positive EC condition the mean increase in positive child evaluations was significant, $t(25) = 3.40, p = .001; M = 1.28, SD = 1.92, 95\% \text{ CI}[0.64, \infty]$, $dz = .67$, post hoc power = 95.3%.

C2 compared change scores for positive child evaluations between the backward positive EC condition and the forward pseudo-EC condition. C2 was significant, $t(49) = 3.07, p = .002$, $M_{\text{diff}} = 1.18, SE = 0.38; 95\% \text{ CI}[0.54, \infty]$. In the backward positive EC condition the mean increase in positive child evaluations was significant, $t(24) = 3.69, p = .001; M = 1.13, SD = 1.53, 95\% \text{ CI}[0.61, \infty]$, $dz = .74$, post hoc power = 97.4%.

In the pseudo-EC condition, the mean change in positive child evaluations was not significant, $t(26) = -0.22, p = .830; M = -0.05, SD = 1.20, 95\% \text{ CI}[-.54, .43]$, $dz = -.04$, post hoc power = 5.4%. 
Reduced dataset. Zero participants had the maximum response value (9) for positive child evaluations pre-EC. Thus, the full sample was not reduced.

Negative Child Evaluations

Full dataset. C1 compared negative child evaluation change scores between the forward positive EC condition and the forward pseudo-EC condition. As hypothesized, C1 was significant, $t(50) = 1.93, p = .030, M_{diff} = 0.59, SE = 0.31; 95\% CI[.08, \infty]$. As hypothesized, in the forward positive EC condition the mean reduction in negative child evaluations was significant, $t(25) = 3.44, p = .001; M = 1.28, SD = 1.92, 95\% CI[.64, \infty], dz = .67, \text{post hoc power} = 95.3\%$.

C2 compared negative child evaluation change scores between the backward positive EC condition and the forward pseudo-EC condition. C2 was significant, $t(49) = 3.62, p < .001, M_{diff} = 1.03, SE = 0.285; 95\% CI[0.55, \infty]$. In the backward positive EC condition the mean reduction in negative child evaluations was significant, $t(24) = 6.01, p < .001; M = 1.26, SD = 1.05, 95\% CI[.90, \infty], dz = 1.20, \text{post hoc power} = 99.9\%$.

In the pseudo-EC condition, the mean change in negative child evaluations was not significant, $t(25) = 1.17, p = .254; M = .23, SD = .99, 95\% CI[-.17, .63], dz = .23, \text{post hoc power} = 20.4\%$.

Reduced dataset. Zero participants had the minimum response value (0) for negative child evaluations pre-EC. Thus, the full sample was not reduced.
Liking

Full dataset. C1 compared liking change scores between the forward positive EC condition and the forward pseudo-EC condition. Contrary to hypotheses, C1 was not significant, $t(50) = 1.15, p = .127, M_{diff} = 0.40, SE = 0.34; 95\% CI[-0.18, \infty]$. Contrary to hypotheses, in the forward positive EC condition the mean increase in liking was not significant, $t(25) = 1.42, p = .084; M = .37, SD = 1.34, 95\% CI[-.07, \infty], d_z = .28$, post hoc power = 39.9%.

C2 compared liking change scores between the backward positive EC condition and the forward pseudo-EC condition. C2 was significant, $t(49) = 1.84, p = .035, M_{diff} = 0.559, SE = 0.30; 95\% CI[0.05, \infty]$. In the backward positive EC condition the mean increase in liking was significant, $t(24) = 2.62, p = .008; M = .53, SD = 1.02, 95\% CI[.18, \infty], d_z = .52$, post hoc power = 81.1%.

In the pseudo-EC condition, the mean change in liking was not significant, $t(26) = -0.11, p = .910; M = -.02, SD = 1.14, 95\% CI[-.49, .43], d_z = -.02$, post hoc power = 5.1%.

Reduced dataset. The reduced sample for C1 contained 50 participants. Three participants (one from the backward positive EC condition and two in the forward pseudo-EC condition) selected the maximum response value (9) for liking pre-EC. C1 compared liking change scores between the forward positive EC condition and the forward pseudo-EC condition. Contrary to hypotheses, C1 was not significant, $t(48) = 1.20, p = .119, M_{diff} = 0.43, 95\% CI[-0.17, \infty]$. Contrary to hypotheses, in the forward positive EC condition the mean increase in liking was not significant, $t(23) = 1.42, p = .084; M = .40, SD = 1.39, 95\% CI[-.08, \infty], d_z = .72$, post hoc power = 97.3%.
The reduced sample for C2 contained 50 participants. C2 compared change scores for liking between the backward positive EC condition and the forward pseudo-EC condition. C2 was significant, $t(48) = 1.88, p = .033, M_{diff} = 0.58, 95\% \text{ CI}[0.06, \infty]$. In the backward positive EC condition the mean increase in liking was significant, $t(23) = 2.63, p = .007; M = .56, SD = 1.03, 95\% \text{ CI}[0.18, \infty]$, $dz = .54$, post hoc power = 83.6%.

No data were removed from the forward pseudo-EC condition.

Future Discipline Expectations

Full dataset. C1 compared future discipline expectations change scores between the forward positive EC condition and the forward pseudo-EC condition. Contrary to hypotheses, C1 was not significant, $t(50) = 1.23, p = .112, M_{diff} = 0.41, SE = 0.33; 95\% \text{ CI}[-0.15, \infty]$. As hypothesized, in the forward positive EC condition the mean reduction in future discipline expectations was significant, $t(25) = 2.67, p = .007; M = .69, SD = 1.32, 95\% \text{ CI}[,25, \infty], dz = .52$, post hoc power = 82.5%.

C2 compared change scores for future discipline expectations between the backward positive EC condition and the forward pseudo-EC condition. C2 was not significant, $t(49) = 1.12, p = .135, M_{diff} = 0.38, SE = 0.34; 95\% \text{ CI}[-0.19, \infty]$. In the backward positive EC condition the mean reduction in future discipline expectations was significant, $t(24) = 2.41, p = .012; M = .67, SD = 1.38, 95\% \text{ CI}[,19, \infty], dz = .49$, post hoc power = 76.9%.

In the pseudo-EC condition, the mean change in future discipline expectations was not significant, $t(25) = 1.36, p = .188; M = .23, SD = .99, 95\% \text{ CI}[-.15, .71], dz = .23$, post hoc power = 20.4%.
**Reduced dataset.** One participant from the backward positive EC condition had the minimum response value (0) for future discipline expectations pre-EC. The reduced sample for C2 contained 50 participants. C2 compared change scores for future discipline expectations between the backward positive EC condition and the forward pseudo-EC condition. C2 was not significant, $t(48) = 1.18, p = .122, M_{diff} = 0.41, SE = 0.34; 95\% \text{ CI}[-0.19, \infty]$. As hypothesized, in the backward positive EC condition the mean reduction in future discipline expectations was significant, $t(24) = 2.42, p = .012; M = .69, SD = 1.40, 95\% \text{ CI} [.20, \infty]$, $d_{z} = .49$, post hoc power = 76.7%.

No data were removed for the forward positive EC condition and the forward pseudo-EC condition.

**Attributions of Hostile Intent**

**Full dataset.** C1 compared change scores for attributions of hostile intent between the forward positive EC condition and the forward pseudo-EC condition. Contrary to hypotheses, C1 was not significant, $t(50) = 0.15, p = .559, M_{diff} = -0.04, 95\% \text{ CI}[-0.52, \infty]$. Contrary to hypotheses, in the forward positive EC condition the mean reduction in attributions of hostile intent was not significant, $t(25) = 0.95, p = .175; M = .19, SD = 1.02, 95\% \text{ CI}[-1.15, \infty]$, $d_{z} = .19$, post hoc power = 24.1%.

C2 compared change scores for attributions of hostile intent between the backward positive EC condition and the forward pseudo-EC condition. C2 was not significant, $t(49) = 1.13, p = .132, M_{diff} = 0.28, 95\% \text{ CI}[-0.14, \infty]$. In the backward positive EC condition the mean
reduction in attributions of hostile intent was significant, \( t(24) = 3.80, p < .001; M = .51, SD = .67, 95\% \text{ CI} [.28, \infty], dz = .76, \text{post hoc power} = 98.0\% \).

In the pseudo-EC condition, the mean change in attributions of hostile intent was not significant, \( t(25) = 1.15, p = .261; M = .23, SD = 1.03, 95\% \text{ CI} [-.18, .65], dz = .22, \text{post hoc power} = 19.0\% \).

**Reduced dataset.** Zero participants had the minimum response value (0) for attributions of hostile intent pre-EC. Thus, the full sample was not reduced.

**Anger**

**Full dataset.** C1 compared anger change scores between the forward positive EC condition and the forward pseudo-EC condition. Contrary to hypotheses, C1 was not significant, \( t(50) = 1.25, p = .109, M_{\text{diff}} = 0.27, SE = 0.22; 95\% \text{ CI} [-0.09, \infty] \). However, as hypothesized, in the forward positive EC condition the mean reduction in anger was significant, \( t(25) = 2.04, p = .026; M = .30, SD = .75, 95\% \text{ CI} [.05, \infty], dz = .40, \text{post hoc power} = 63.3\% \).

C2 compared anger change scores between the backward positive EC condition and the forward pseudo-EC condition. C2 was significant, \( t(49) = 1.97, p = .027, M_{\text{diff}} = 0.39, SE = 0.20; 95\% \text{ CI} [0.06, \infty] \). In the backward positive EC condition the mean reduction in anger was significant, \( t(24) = 3.65, p = .001; M = .42, SD = .57, 95\% \text{ CI} [.22, \infty], dz = .74, \text{post hoc power} = 97.4\% \).

In the pseudo-EC condition, the mean change in anger was not significant, \( t(25) = 0.20, p = .841; M = .03, SD = .80, 95\% \text{ CI} [-.29, .36], dz = .04, \text{post hoc power} = 5.4\% \).
Reduced dataset. The reduced sample for C1 contained 50 participants. Three participants (one from the forward positive EC condition, one from the backward positive EC condition, and one from the forward pseudo-EC condition) had the minimum response value (0) for anger pre-EC. C1 compared anger change scores between the forward positive EC condition and the forward pseudo-EC condition. Contrary to hypotheses, C1 was not significant, \( t(48) = 1.35, p = .092\), \( M_{\text{diff}} = .30\), 95% CI\([-0.07, \infty]\). As hypothesized, in the forward positive EC condition the mean reduction in anger was significant, \( t(24) = 2.22, p = .018\); \( M = .33\), \( SD = .75\), 95% CI\([.08, \infty]\), \( d_z = .44\), post hoc power = 70.4%.

The reduced sample for C2 contained 49 participants. C2 compared anger change scores between the backward positive EC condition and the forward pseudo-EC condition. C2 was significant, \( t(47) = 1.98, p = .027\), \( M_{\text{diff}} = .40\), 95% CI\([0.06, \infty]\). As hypothesized, in the backward positive EC condition the mean reduction in anger was significant, \( t(23) = 3.69, p = .001\); \( M = .44\), \( SD = .58\), 95% CI\([.23, \infty]\), \( d_z = .76\), post hoc power = 98.0%.

In the pseudo-EC condition, the mean change in anger was not significant, \( t(24) = 0.20, p = .841\); \( M = .03\), \( SD = .82\), 95% CI\([-0.31, .36]\), \( d_z = .04\), post hoc power = 5.4%.

Harsh Verbal Discipline

Full dataset. C1 compared harsh verbal discipline change scores between the forward positive EC condition and the forward pseudo-EC condition. Contrary to hypotheses, C1 was not significant, \( t(50) = 0.77, p = .222\), \( M_{\text{diff}} = 0.12\), \( SE = 0.15\); 95% CI\([-0.14, \infty]\). Contrary to hypotheses, in the forward positive EC condition the mean reduction in harsh verbal discipline
was not significant, \( t(25) = 1.30, p = .103; M = .13, SD = .51, 95\% CI[-.04, \infty], dz = .25, \) post hoc power = 34.3%.

C2 compared harsh verbal discipline change scores between the backward positive EC condition and the forward pseudo-EC condition. C2 was not significant, \( t(49) = 1.59, p = .059, M_{diff} = 0.23, SE = 0.145; 95\% CI[-0.01, \infty]. \) As hypothesized, in the backward positive EC condition the mean reduction in harsh verbal discipline was significant, \( t(24) = 2.88, p = .004; M = .24, SD = .42, 95\% CI[.10, \infty], dz = .57, \) post hoc power = 86.9%.

In the pseudo-EC condition, the mean change in harsh verbal discipline was not significant, \( t(25) = 0.09, p = .928; M = .01, SD = .80, 95\% CI[-.23, .25], dz = .01, \) post hoc power = 5.0%.

Reduced dataset. The reduced sample for C1 contained 40 participants. Eighteen participants (eight from the forward positive EC condition, six from the backward positive EC condition, and four from the forward pseudo-EC condition) had the minimum response value (0) for harsh verbal discipline pre-EC. C1 compared harsh verbal discipline change scores between the forward positive EC condition and the forward pseudo-EC condition. Contrary to hypotheses, C1 was not significant, \( t(38) = 0.80, p = .212, M_{diff} = .16, 95\% CI[-.17, \infty]. \) Contrary to hypotheses, in the forward positive EC condition the mean reduction in harsh verbal discipline was not significant, \( t(17) = 1.30, p = .105; M = .19, SD = .61, 95\% CI[-.06, \infty], dz = .31, \) post hoc power = 35.1%.

The reduced sample for C2 contained 41 participants. C2 compared harsh verbal discipline change scores between the backward positive EC condition and the forward pseudo-EC condition. C2 was not significant, \( t(39) = 1.63, p = .056, M_{diff} = 0.29, 95\% CI[-.01, \infty]. \) As
hypothesized, in the backward positive EC condition the mean reduction in harsh verbal discipline was significant, \( t(18) = 3.03, p = .004; M = .32, SD = .46, 95\% \text{ CI}[.14, \infty], dz = .70, \) post hoc power = 90.1%.

In the pseudo-EC condition, the mean change in harsh verbal discipline was not significant, \( t(21) = 0.20, p = .843; M = .03, SD = .65, 95\% \text{ CI}[-.26, .32], dz = .05, \) post hoc power = 5.6%.

**Harsh Physical Discipline**

**Full dataset.** C1 compared harsh physical discipline change scores between the forward positive EC condition and the forward pseudo-EC condition. Contrary to hypotheses, C1 was not significant, \( t(50) = 1.12, p = .133, M_{\text{diff}} = 0.08, SE = 0.07; 95\% \text{ CI}[-0.06, \infty]. \) Contrary to hypotheses, in the forward positive EC condition the mean reduction in harsh physical discipline was not significant, \( t(25) = 0.37, p = .356; M = .02, SD = .29, 95\% \text{ CI}[-.08, \infty], dz = .07, \) post hoc power = 9.7%.

C2 compared harsh physical discipline change scores between the backward positive EC condition and the forward pseudo-EC condition. C2 was not significant, \( t(49) = 1.59, p = .059, M_{\text{diff}} = 0.06, SE = 0.04; 95\% \text{ CI}[-0.00, \infty]. \) Contrary to hypotheses, in the backward positive EC condition the mean reduction in harsh physical discipline was not significant, \( t(24) = 0.83, p = .208; M = .01, SD = .04, 95\% \text{ CI}[-.01, \infty]. dz = .02, \) post hoc power = 6.1%.

In the pseudo-EC condition, the mean change in harsh physical discipline was not significant, \( t(25) = -1.48, p = .151; M = -.06, SD = .19, 95\% \text{ CI}[-.13, .02], dz = -.32, \) post hoc power = 34.8%. 
Reduced dataset. The reduced sample for C1 contained 16 participants. Fifty-seven participants (19 from the forward positive EC condition, 21 from the backward positive EC condition, and 17 from the forward pseudo-EC condition) had the minimum response value (0) for harsh physical discipline pre-EC. C1 compared harsh physical discipline change scores between the forward positive EC condition and the forward pseudo-EC condition. Contrary to hypotheses, C1 was not significant, \( t(14) = 0.97, p = .175, M_{\text{diff}} = .22, 95\% \text{ CI}[-.17, \infty] \). Contrary to hypotheses, in the forward positive EC condition the mean reduction in harsh physical discipline was not significant, \( t(6) = 0.35, p = .368; M = .08, SD = .59, 95\% \text{ CI}[-.36, \infty], dz = .14, \) post hoc power = 9.4%.

The reduced sample for C2 contained 13 participants. C2 compared harsh physical discipline change scores between the backward positive EC condition and the forward pseudo-EC condition. C2 was not significant, \( t(11) = 1.31, p = .108, M_{\text{diff}} = 0.21, 95\% \text{ CI}[-.08, \infty] \). As hypothesized, in the backward positive EC condition the mean reduction in harsh physical discipline was significant, \( t(3) = 2.61, p = .040; M = .07, SD = .05, 95\% \text{ CI}[.01, \infty], dz = 1.40, \) post hoc power = 68.6%.

In the pseudo-EC condition, the mean change in harsh physical discipline was not significant, \( t(8) = -1.36, p = .211; M = -.14, SD = .31, 95\% \text{ CI}[-.38, .10], dz = -.45, \) post hoc power = 22.2%.

Doing Nothing

Full dataset. C1 compared doing nothing change scores between the forward positive EC condition and the forward pseudo-EC condition. Contrary to hypotheses, C1 was not significant,
Contrary to hypotheses, in the forward positive EC condition the mean increase in doing nothing was not significant, $t(25) = 1.52, p = .070$; $M = .52, SD = 1.73, 95\% CI[-.06, \infty]$, $dz = .30$, post hoc power = 43.8%.

C2 compared doing nothing change scores between the backward positive EC condition and the forward pseudo-EC condition. C2 was not significant, $t(49) = 0.67, p = .253$, $M_{\text{diff}} = 0.245, SE = 0.37; 95\% CI[-0.37, \infty]$. In the backward positive EC condition the mean increase in doing nothing was not significant, $t(24) = 1.33, p = .098$; $M = .45, SD = 1.68, 95\% CI[-.13, \infty]$, $dz = .27$, post hoc power = 37.0%.

In the pseudo-EC condition, the mean change in doing nothing was not significant, $t(25) = 1.28, p = .210$; $M = .20, SD = .80, 95\% CI[-.12, .52]$, $dz = .25$, post hoc power = 23.2%.

**Reduced dataset.** Zero participants had the maximum response value (9) for doing nothing pre-EC. Thus, the full sample was not reduced.

**Harm**

**Full dataset.** C1 compared harm change scores between the forward positive EC condition and the forward pseudo-EC condition. Contrary to hypotheses, C1 was not significant, $t(50) = 1.09, p = .141$, $M_{\text{diff}} = 0.31, SE = 0.28; 95\% CI[-0.17, \infty]$. As hypothesized, in the forward positive EC condition the mean reduction in harm was significant, $t(25) = 1.73, p = .048$; $M = .38, SD = 1.13, 95\% CI[.00, \infty]$, $dz = .34$, post hoc power = 51.7%.

C2 compared harm change scores between the backward positive EC condition and the forward pseudo-EC condition. C2 was significant, $t(49) = 1.75, p = .043$, $M_{\text{diff}} = 2.60, SE = 1.49$;
95% CI[0.11, ∞]. In the backward positive EC condition the reduction in harm was significant, \( t(24) = 1.78, p = .044; M = 2.68, SD = 7.53, 95\% \text{ CI}[.10, \infty], dz = .36 \), post hoc power = 54.1%.

In the pseudo-EC condition, the mean change in harm was not significant, \( t(25) = 0.44, p = .664; M = .08, SD = .89, 95\% \text{ CI}[-.28, .44], dz = .09 \), post hoc power = 7.3%.

**Reduced dataset.** The reduced sample for C1 contained 16 participants. Sixty participants (23 from the forward positive EC condition, 17 from the backward positive EC condition, and 20 from the forward pseudo-EC condition) had the minimum response value (0) for harm pre-EC. C1 compared harm change scores between the forward positive EC condition and the forward pseudo-EC condition. Contrary to hypotheses, C1 was not significant, \( t(14) = 0.97, p = .175, M_{\text{diff}} = .22, 95\% \text{ CI}[-.17, \infty] \). Contrary to hypotheses, in the forward positive EC condition the mean reduction in harm was not significant, \( t(6) = 0.35, p = .368; M = .08, SD = .59, 95\% \text{ CI}[-.36, \infty], dz = .14 \), post hoc power = 9.4%.

The reduced sample for C2 contained 13 participants. C2 compared harm change scores between the backward positive EC condition and the forward pseudo-EC condition. C2 was not significant, \( t(11) = 1.31, p = .108, M_{\text{diff}} = 0.21, 95\% \text{ CI}[-.08, \infty] \). As hypothesized, in the backward positive EC condition the mean reduction in harm was significant, \( t(3) = 2.61, p = .040; M = .07, SD = .05, 95\% \text{ CI}[.01, \infty], dz = 1.40 \), post hoc power = 68.6%.

In the pseudo-EC condition, the mean change in harm was not significant, \( t(8) = -1.36, p = .211; M = -.14, SD = .31, 95\% \text{ CI}[-.38, .10], dz = -.45 \), post hoc power = 22.2%. 

Planned Exploratory Analyses

Analysis of Experimental Design Using Repeated-Measures ANOVA

While the pre-planned comparisons were focused on the exact hypotheses proposed, additional information about Study 1 can be obtained by examining results for separate EC Condition (forward positive EC, backward positive EC, forward pseudo-EC) × Time (pre-EC, post-EC) ANOVAs with repeated measures on the second variable performed on each dependent measure.

Positive child evaluations. The EC Condition × Time interaction was significant, \( F(2, 74) = 5.52, p = .006, \eta_p^2 = .13, 95\% \text{ CI}[0.02, 0.24] \). The interaction was probed by examining all simple effects. In the forward positive EC condition, when pre-EC positive child evaluations (\( M = 4.50, SD = 1.40 \)) and post-EC positive child evaluations (\( M = 5.79, SD = 1.94 \)) were compared, the difference was significant, \( t(25) = 3.40, p = .001, 95\% \text{ CI} [.64, \infty] \). In the backward positive EC condition, when pre-EC positive child evaluations (\( M = 5.31, SD = 1.33 \)) and post-EC positive child evaluations (\( M = 6.44, SD = 1.50 \)) were compared, the difference was significant, \( t(24) = 3.70, p = .001, 95\% \text{ CI} [.61, \infty] \). In the forward pseudo-EC condition, when pre-EC positive child evaluations (\( M = 4.50, SD = 1.40 \)) and post-EC positive child evaluations (\( M = 5.79, SD = 1.94 \)) were compared, the difference was not significant, \( t(25) = -0.22, p = .800, 95\% \text{ CI} [-.54, .43] \).

Pre-EC positive child evaluations did not significantly vary between EC condition, \( F(2, 74) = 2.50, p = .089, \eta_p^2 = .06, 95\% \text{ CI} [.00, .15] \); forward positive EC \( M = 4.50, SD = 1.40 \); backward positive EC \( M = 5.31, SD = 1.33 \); forward pseudo-EC \( M = 5.10, SD = 1.285 \).
Post-EC positive child evaluations significantly varied between EC conditions, \( F(2, 74) = 4.58, p = .013, \eta^2_p = .11, 95\% \text{ CI}[.01, .21] \); forward positive EC \( M = 4.50, SD = 1.40 \); backward positive EC \( M = 5.31, SD = 1.33 \); forward pseudo-EC \( M = 5.10, SD = 1.28 \). SNK test results showed the following significant differences: backward positive EC = forward positive EC; forward positive EC = forward pseudo-EC; backward positive EC > forward pseudo-EC.

The main effect of EC condition was not significant, \( F(2, 74) = 3.05, p = .053, \eta^2_p = .08, 95\% \text{ CI}[0.00, 0.17] \); forward positive EC \( M = 5.145, SD = 1.39 \); backward positive EC \( M = 5.87, SD = 1.19 \); forward pseudo-EC \( M = 5.08, SD = 1.21 \).

The main effect of time was significant, \( F(1, 74) = 19.06, p < .001, \eta^2_p = .20, 95\% \text{ CI}[0.08, 0.33] \); pre-EC \( M = 4.97, SD = 1.36 \); post-EC \( M = 5.75, SD = 1.71 \).

**Negative child evaluations.** The EC Condition × Time interaction was significant, \( F(2, 74) = 5.80, p = .005, \eta^2_p = .13, 95\% \text{ CI}[0.02, 0.24] \). The interaction was probed by examining all simple effects. In the forward positive EC condition, when pre-EC negative child evaluations (\( M = 2.98, SD = 1.28 \)) and post-EC negative child evaluations (\( M = 2.16, SD = 1.64 \)) were compared, the difference was significant, \( t(25) = 3.40, p = .001, 95\% \text{ CI} [.41, \infty] \). In the backward positive EC condition, when pre-EC negative child evaluations (\( M = 3.02, SD = 1.38 \)) and post-EC negative child evaluations (\( M = 1.76, SD = 1.53 \)) were compared, the difference was significant, \( t(24) = 6.00, p < .001, 95\% \text{ CI} [1.28, \infty] \). In the forward pseudo-EC condition, when pre-EC negative child evaluations (\( M = 3.03, SD = 1.11 \)) and post-EC negative child evaluations (\( M = 2.80, SD = 1.37 \)) were compared, the difference was not significant, \( t(25) = 1.20, p = .300, 95\% \text{ CI} [-.63, .17] \).
Pre-EC negative child evaluations did not significantly vary between EC condition, $F(2, 74) = 0.01, p = .990, \eta^2_p = .00$, 95% CI[.00, 1.00]; forward positive EC $M = 2.98, SD = 1.285$; backward positive EC $M = 3.02, SD = 1.38$; forward pseudo-EC $M = 3.03, SD = 1.11$.

Post-EC negative child evaluations did not significantly vary between EC conditions, $F(2, 74) = 3.05, p = .053, \eta^2_p = .08$, 95% CI[.00, .17]; forward positive EC $M = 4.50, SD = 1.40$; backward positive EC $M = 5.31, SD = 1.33$; forward pseudo-EC $M = 5.10, SD = 1.285$.

The main effect of EC condition was not significant, $F(2, 74) = 1.10, p = .339, \eta^2_p = .03$, 95% CI[0.00, 0.10]; forward positive EC $M = 2.57, SD = 1.34$; backward positive EC $M = 2.39, SD = 1.36$; forward pseudo-EC $M = 2.91, SD = 1.14$.

The main effect of time was significant, $F(1, 74) = 38.40, p < .001, \eta^2_p = .34$, 95% CI[0.20, 0.46]; pre-EC $M = 3.01, SD = 1.24$; post-EC $M = 2.25, SD = 1.56$.

**Liking.** The EC Condition × Time interaction was not significant, $F(2, 74) = 1.54, p = .221, \eta^2_p = .04$, 95% CI[0.00, 0.12].

The main effect of EC condition was not significant, $F(2, 74) = 1.08, p = .345, \eta^2_p = .03$, 95% CI[0.00, 0.10]; forward positive EC $M = 5.955, SD = 1.515$; backward positive EC $M = 6.41, SD = 1.49$; forward pseudo-EC $M = 5.85, SD = 1.37$.

The main effect of time was significant, $F(1, 74) = 4.80, p = .032, \eta^2_p = .06$, 95% CI[0.00, 0.16]; pre-EC $M = 5.92, SD = 1.55$; post-EC $M = 6.21, SD = 1.60$.

**Future discipline expectations.** The EC Condition × Time interaction was not significant, $F(2, 74) = 0.86, p = .428, \eta^2_p = .00$, 95% CI[0.0, 1.0].
The main effect of EC condition was not significant, $F(2, 74) = 0.01, p = .990, \eta^2_p = .03$, 95% CI[0.00, 0.10]; forward positive EC $M = 4.22, SD = 1.60$; backward positive EC $M = 4.16, SD = 1.70$; forward pseudo-EC $M = 4.205, SD = 1.38$.

The main effect of time was significant, $F(1, 74) = 14.815, p < .001, \eta^2_p = .02$, 95% CI[0.00, 0.08]; pre-EC $M = 4.47, SD = 1.58$; post-EC $M = 3.92, SD = 1.75$.

**Attributions of hostile intent.** The EC Condition × Time interaction was not significant, $F(2, 74) = 0.89, p = .417, \eta^2_p = .02$, 95% CI[0.00, 0.09].

The main effect of EC condition was not significant, $F(2, 74) = 0.80, p = .455, \eta^2_p = .02$, 95% CI[0.00, 0.08]; forward positive EC $M = 2.26, SD = 1.65$; backward positive EC $M = 1.79, SD = 1.18$; forward pseudo-EC $M = 2.18, SD = 1.42$.

The main effect of time was significant, $F(1, 74) = 8.68, p = .004, \eta^2_p = .10$, 95% CI[0.02, 0.22]; pre-EC $M = 2.23, SD = 1.39$; post-EC $M = 1.93, SD = 1.61$.

**Anger.** The EC Condition × Time interaction was not significant, $F(2, 74) = 1.96, p = .149, \eta^2_p = .05$, 95% CI[0.00, 0.13].

The main effect of EC condition was not significant, $F(2, 74) = 0.01, p = .992, \eta^2_p = .00$, 95% CI[0.00, 1.00]; forward positive EC $M = 1.17, SD = 1.01$; backward positive EC $M = 1.14, SD = 1.11$; forward pseudo-EC $M = 1.15, SD = 1.13$.

The main effect of time was significant, $F(1, 74) = 9.37, p = .003, \eta^2_p = .11$, 95% CI[0.02, 0.23]; pre-EC $M = 1.28, SD = 1.12$; post-EC $M = 1.03, SD = 1.13$.

**Harsh verbal discipline.** The EC Condition × Time interaction was not significant, $F(2, 74) = 1.28, p = .285, \eta^2_p = .03$, 95% CI[0.00, 0.105].
The main effect of EC condition was not significant, $F(2, 74) = 0.985, p = .378, \eta^2_p = .02$, 95% CI[.00, .09]; forward positive EC $M = 0.79, SD = 1.05$; backward positive EC $M = 0.51, SD = 0.72$; forward pseudo-EC $M = 0.90, SD = 1.05$.

The main effect of time was significant, $F(1, 74) = 4.69, p = .0335, \eta^2_p = .11$, 95% CI[0.00, 0.16]; pre-EC $M = 0.80, SD = 1.03$; post-EC $M = 67, SD = 1.06$.

**Harsh physical discipline.** The EC Condition × Time interaction was not significant, $F(2, 74) = 1.04, p = .360, \eta^2_p = .03$, 95% CI[.00, .09].

The main effect of EC condition was not significant, $F(2, 74) = 1.06, p = .353, \eta^2_p = .03$, 95% CI[.00, .095]; forward positive EC $M = 0.19, SD = 0.45$; backward positive EC $M = 0.08, SD = 0.26$; forward pseudo-EC $M = 0.34, SD = 0.98$.

The main effect of time was not significant, $F(1, 74) = 0.155, p = .695, \eta^2_p = .00$, 95% CI[0.00, 0.05]; pre-EC $M = 0.20, SD = 0.63$; post-EC $M = 0.21, SD = 0.68$.

**Doing nothing.** The EC Condition × Time interaction was not significant, $F(2, 74) = 0.33, p = .720, \eta^2_p = .01$, 95% CI[.00, .05].

The main effect of EC condition was significant, $F(2, 74) = 8.05, p = .001, \eta^2_p = .18$, 95% CI[.05, .29]; forward positive EC $M = 5.81, SD = 2.32$; backward positive EC $M = 4.10, SD = 2.43$; forward pseudo-EC $M = 6.545, SD = 1.91$. SNK test results showed the following significant differences: forward pseudo-EC = forward positive EC > backward positive EC.

The main effect of time was significant, $F(1, 74) = 5.39, p = .023, \eta^2_p = .07$, 95% CI[0.01, 0.17]; pre-EC $M = 5.31, SD = 2.425$; post-EC $M = 5.70, SD = 2.63$.

**Harm.** The EC Condition × Time interaction was not significant, $F(2, 74) = 2.69, p = .075, \eta^2_p = .07$, 95% CI[.00, .16].
The main effect of EC condition was not significant, $F(2, 74) = 0.44, p = .644, \eta_p^2 = .01$, 95% CI[0.00, .06]; forward positive EC $M = 1.385$, $SD = 5.37$; backward positive EC $M = 3.02$, $SD = 7.88$; forward pseudo-EC $M = 2.385$, $SD = 5.29$.

The main effect of time was significant, $F(1, 74) = 4.425, p = .039, \eta_p^2 = .06$, 95% CI[0.00, 0.16]; pre-EC $M = 2.77$, $SD = 7.67$; post-EC $M = 1.74$, $SD = 5.35$.

Post-EC Differences Between EC Conditions for Each Dependent Variable, Estimated at the Mean of Pre-EC Scores (ANCOVA)

It can be desirable to confirm the results of analyses that use difference scores by using alternative statistical methods to explore the data. This was done using an ANCOVA method. As noted below, the results from the ANCOVA generally duplicated results from the difference score analysis.

Post-EC positive child evaluations. After adjusting for pre-EC positive child evaluations ($M = 4.97$), the effect of EC condition on post-EC positive child evaluations was significant, $F(2, 73) = 5.53, p = .006$; forward positive EC least squares mean ($LSM$) = 6.04, $SE = 0.29$; backward positive EC $LSM = 6.25$, $SE = 0.30$; forward pseudo-EC $LSM = 4.98$, $SE = 0.29$. $LSM$ pairwise contrasts showed the following significant differences: backward positive EC condition > forward pseudo-EC condition, $LSM$ diff. = 1.27, $SE = 0.41$, $t(73) = 3.08, p = .003$; forward positive EC condition > forward pseudo-EC condition, $LSM$ diff. = 1.06, $SE = .415$, $t(73) = 2.56, p = .0125$; forward positive EC = backward positive EC, $LSM$ diff = 0.21, $SE = .42$, $t(73) = 0.49, p = .625$. 
Post-EC negative child evaluations. After adjusting for pre-EC negative child evaluations ($M = 3.01$), the effect of EC condition on post-EC negative child evaluations was significant, $F(2, 73) = 5.92, p = .004$; forward positive EC $LSM = 2.19, SE = 0.21$; backward positive EC $LSM = 1.75, SE = 0.215$; forward pseudo-EC $LSM = 2.78, SE = 0.21$. $LSM$ pairwise contrasts showed the following significant differences: backward positive EC condition $>$ forward pseudo-EC condition, $LSM$ diff. = 1.03, $SE = 0.30$, $t(73) = 3.42, p = .001$; forward positive EC condition $>$ forward pseudo-EC condition $LSM$ diff. = 0.60, $SE = 0.30$, $t(73) = 2.00$, $p = .0495$; forward positive EC = backward positive EC $LSM$ diff. = 0.44, $SE = .30$, $t(73) = 1.44$, $p = .153$.

Post-EC liking. After adjusting for pre-EC liking ($M = 5.92$), the effect of EC condition on post-EC liking was not significant, $F(2, 73) = 2.13, p = .123$; forward positive EC $LSM = 6.25, SE = 0.22$; backward positive EC $LSM = 6.52, SE = 0.22$; forward pseudo-EC $LSM = 5.88, SE = 0.22$.

Post-EC future discipline expectations. After adjusting for pre-EC future discipline expectations ($M = 4.47$), the effect of EC condition on post-EC future discipline expectations was not significant, $F(2, 73) = 0.97, p = .38$; forward positive EC $LSM = 3.79, SE = 0.24$; backward positive EC $LSM = 3.81, SE = 0.245$; forward pseudo-EC $LSM = 4.16, SE = 0.24$.

Post-EC attributions of hostile intent. After adjusting for pre-EC attributions of hostile intent ($M = 2.23$), the effect of EC condition on post-EC attributions of hostile intent was not significant, $F(2, 73) = 1.07, p = .35$; forward positive EC $LSM = 2.05, SE = 0.18$; backward positive EC $LSM = 1.71, SE = 0.19$; forward pseudo-EC $LSM = 2.00, SE = 0.18$. 
Post-EC anger. After adjusting for pre-EC anger ($M = 1.28$), the effect of EC condition on post-EC anger was not significant, $F(2, 73) = 1.74$, $p = .18$; forward positive EC $LSM = 0.99$, $SE = 0.135$; backward positive EC $LSM = 0.87$, $SE = 0.14$; forward pseudo-EC $LSM = 1.23$, $SE = 0.135$.

Post-EC harsh verbal discipline. After adjusting for pre-EC harsh verbal discipline ($M = 0.80$), the effect of EC condition on post-EC harsh verbal discipline was not significant, $F(2, 73) = 1.65$, $p = .20$; forward positive EC $LSM = 0.67$, $SE = 0.10$; backward positive EC $LSM = 0.54$, $SE = 0.10$; forward pseudo-EC $LSM = 0.80$, $SE = 0.10$.

Post-EC harsh physical discipline. After adjusting for pre-EC harsh physical discipline ($M = 0.20$), the effect of EC condition on post-EC harsh physical discipline was not significant, $F(2, 73) = 0.92$, $p = .400$; forward positive EC $LSM = 0.18$, $SE = 0.04$; backward positive EC $LSM = 0.20$, $SE = 0.04$; forward pseudo-EC $LSM = 0.25$, $SE = 0.04$.

Post-EC doing nothing. After adjusting for pre-EC “doing nothing” ($M = 5.31$), the effect of EC condition on post-EC “doing nothing” was not significant, $F(2, 73) = 0.21$, $p = .812$; forward positive EC $LSM = 5.85$, $SE = 0.29$; backward positive EC $LSM = 5.63$, $SE = 0.31$; forward pseudo-EC $LSM = 5.62$, $SE = 0.30$.

Post-EC harm. After adjusting for pre-EC harm ($M = 2.77$), the effect of EC condition on post-EC harm was not significant, $F(2, 73) = 2.29$, $p = .110$; forward positive EC $LSM = 1.89$, $SE = 0.59$; backward positive EC $LSM = 0.75$, $SE = 0.61$; forward pseudo-EC $LSM = 2.55$, $SE = 0.59$. 
Linear Relationships Between Pre-EC Scores and EC Change Scores Within the Forward Positive EC Condition and the Backward Positive EC Condition

Appendix J presents Study 1 correlations between pre-EC scores for each dependent variable and corresponding change scores within the forward positive EC condition and the backward positive EC condition (combined).

Linear Relationships Between Continuous Demographic Characteristics and EC Change Scores Within the Forward Positive EC Condition and the Backward Positive EC Condition

After applying Bonferroni-Holm correction for multiple testing (22 correlations, corrected $\alpha = .002$), children, age, and education were not significantly correlated with EC change scores for any dependent variable.

EC Change Scores as a Function of Demographic Characteristics Within the Forward Positive EC Condition and the Backward Positive EC Condition

After applying Bonferroni-Holm correction for multiple testing (10 tests, corrected $\alpha = .005$), no significant differences in EC change scores emerged for all nominal demographic variables (i.e., gender, race, and marital status).
**Binomial Test for Harm**

Among the 60 participants who gave no harm EC, one participant gave harm post-EC (binomial test: $p < .001$, null hypothesis: true probability = 0).

**Multi-Level Model Analyses**

Results of MLMs for Study 1 are presented in Appendices K-R. These results are presented for informational purposes only; no inferences were made with respect to generalizing estimates of forward positive EC condition and backward positive EC condition predictors of pre- to post-EC change estimates (reference group = forward pseudo-EC condition).

**Discussion**

Contrary to hypotheses and the results of the preliminary study, pre-EC positive child evaluations did not significantly correlate with attributions of hostile intent, anger, harsh verbal discipline, harsh physical discipline, or harm. Prima facie, this suggests that the construct validity of positive child evaluations requires further research to verify. However, that the preliminary study found that positive child evaluations significantly correlated with other pre-EC dependent measures precludes the firm conclusion that no construct validity evidence was obtained. Additionally, positive child evaluations significantly correlated with liking ($r = .75$), and liking was significantly negatively correlated with both negative child evaluations and attributions of hostile intent. Thus, the present study provides initial evidence for the construct validity of the liking measure.
Three theories, which were the referential account of EC, the propositional account of EC, and the affective implicit misattribution account of EC, predicted divergent results regarding the effects of the different EC conditions in Study 1.

For positive child evaluations, both the forward positive EC and the backward positive EC induced pre- to post-EC reductions, whereas the forward pseudo-EC condition did not. The forward positive EC results are consistent with both the referential account and propositional account, whereas the backward positive EC results are consistent with the propositional account and affective implicit misattribution account. Both findings are consistent with EC research (e.g., see Hofmann et al., 2010, for meta-analyses of EC effects), showing that broadly, CS-US pairings involved in EC procedures cause CS valence to “take on” the valence of the US. Hofmann et al. found that backward EC procedures also show expected EC effects, although, as described earlier, depending on operating processes during the EC procedure, different EC theories explain such results. In addition, the forward positive EC effect on positive child evaluations conceptually replicates (using fewer trials) prior research (Milner et al., 2017).

Only the backward positive EC condition increased liking pre- to post-EC. Liking is a less semantically laden dependent variable than the measures of positive child evaluations and negative child evaluations. Although post-hoc, given the non-semantic attribute of the liking measure (i.e., it is relatively “evaluatively pure”), that only the backward positive EC condition significantly increased liking provides buttressing evidence for the idea that the backward positive EC condition essentially is an affective misattribution procedure. However, with respect to the hypothesis that backward positive EC condition produced EC effects on liking via the affective implicit misattribution mechanism, rigorous confirmatory research is needed.
Although both the positive EC conditions (forward and backward) reduced future discipline expectations, the reductions were not significantly greater than the reductions observed in the forward pseudo-EC condition. This finding was unexpected. Given that the forward pseudo-EC condition did not change other dependent variables (e.g., positive child evaluations, negative child evaluations), explanations for these effects that derive from demand effects and other method artifacts seem relatively implausible. Anecdotal evidence from participant feedback regarding the EC studies suggest that there is wide variability in how individuals interpret the items measuring future discipline expectations. Some individuals interpret the question with the reasonable assumption that all children require future discipline. Yet, this does not explain why future discipline expectations would decrease in the forward pseudo discipline condition. These observations suggest that future psychometric work is needed to measure future discipline expectations. Adding a timeframe to the item, e.g., “To what extent will this child require discipline in the following seven days,” may add utility to the measure.

Attributions of hostile intent were not significantly reduced pre-to-post EC in either the forward positive EC condition or the backward positive EC condition. Given the number of EC trials in this study was small (18 trials) compared to the number of trials in prior positive EC conditions (Milner et al., 2017; 180 trials), it is possible that there were insufficient EC trials to influence attributions of hostile intent despite significant effects on positive child evaluations and negative child evaluations. A related study limitation is that although positive EC increased positive child evaluations and reduced negative child evaluations in both the forward positive EC condition and backward positive EC condition, it is impossible to determine the extent to which specific mechanisms were operating in each of the positive EC conditions. For example, it is
possible that forward positive EC affected child evaluations via only the referential mechanism, but it is also possible that both the referential mechanism and the propositional mechanism caused changes in child evaluations. Regardless of the extent to which mechanisms prompted effects on child evaluations, the SIP model of CPA was not supported by the data. Changes in child evaluations should have corresponded to changes in attributions of hostile intent, but this was not so.

Only backward positive EC significantly reduced anger pre- to post-EC. Although affective implicit misattribution may have prompted positive EC changes in child evaluations and liking in the backward positive EC condition, it is unknown why only backward positive EC reduced anger. Although unexpected, it is possible that changes in child evaluations and liking reduce anger reactions (despite attributions of hostile intent). However, no research exists to explain why reductions in anger would occur despite stability in attributions of hostile intent pre to post-EC.

One speculation is that increases in positive child evaluations and reductions in negative child evaluations are protective against anger reactions after attributions of hostile intent. In other words, if an individual likes a child, the individual may be more likely to use cognitive appraisal or other emotion regulation strategies to regulate anger following attributions of hostile intent. Use of emotion regulation strategies was not studied. However, given a cognitive interpretation of behavior (i.e., attributions of hostile intent), emotion regulation is a plausible route through which attributions of hostile intent may not lead to anger. Future research is needed to explore and test this notion.
With respect to harsh verbal discipline, harsh physical discipline, and harm, neither the forward positive EC condition nor the backward positive EC condition reduced these variables pre- to post-EC. These findings were unexpected because prior research (Milner et al., 2017) found that forward positive EC significantly reduced harsh verbal discipline, harsh physical discipline, and harm pre- to post-EC. A limitation for this study’s ability to detect pre- to post EC reductions for harsh verbal discipline, harsh physical discipline, and harm is that pre-EC means for harsh verbal discipline and harsh physical discipline were very low. Similarly, only 17 (22%) participants gave any harm pre-EC. The cause of the especially high frequencies of minimum values on these dependent variables is unknown.

Finally, although Study 1 provided evidence that positive EC affects changes in child evaluations via multiple mechanisms (e.g., referential, propositional, or affect misattribution mechanisms), Study 1 did not address the extent to which semantic (vs. evaluative) learning mediated positive EC effects observed in Study 1 and in previous research (Milner et al., 2017). Hence, Study 2 was conducted to test whether using positive evaluative (not semantic) CS in positive EC replicates prior findings.
Study 2 explored the extent to which the EC effects reported by Milner et al. (2017) were exclusively related to the use of trait words as US. In the Milner et al. (2017) procedure, trait words (i.e., positive, sweet, friendly) were presented in such a way that they were predicted by the appearance of child photos. It is possible that what was learned in these presentations were associations between these traits and the cognitive representation of the “child” construct, and that the effects observed on the dependent measures used in Milner et al. (2017) reflect these semantic associations. This is in contrast to evaluative learning mechanisms that have been proposed in EC research (see Hofmann et al., 2010) that suggest that it is the evaluation that is learned. In Study 2, these ideas were explored using emojis (instead of trait words) as the US in the Milner et al. (2017) procedure. Emojis are graphic, schematic depictions of facial expressions or cues of emotions; emojis are used in many social media, such as text messaging, Facebook, etc. Compared to positive words, emojis should be relatively devoid of semantic content, but should elicit positive affect when encountered.

To pursue this idea, Study 2 featured two conditions: a positive EC condition and a pseudo-EC condition. In the positive EC condition positive emojis were used as non-semantic positive US. In the pseudo-EC condition neutral-faced emojis were used as non-semantic neutral US. If the EC effects reported by Milner et al. (2017) do not require semantic content in the US
(e.g., the words positive, sweet, and friendly), but instead reflect evaluative learning, then hypothesized effects in Study 1 and previous research (Milner et al., 2017) will replicate in Study 2.

Method

Participants

Fifty-two participants from Northern Illinois University undergraduate classes (introductory to upper-level lab classes) or from research laboratories participated. Potential participants who had previously completed Study 1 or Study 2 were ineligible for Study 2. One participant was excluded for being affiliated with a project using EC procedures similar to the one used in this study. One participant was excluded for excessive missing data (i.e., answering only one item for each computed measure in the study). Thus, the sample totaled 50 participants (i.e., the planned sample size).

The sample was 26% male and 74% female. The sample ethnicity was 78% White, 12% Black, 4% Asian, 4% Multi-racial and 6% “Other race not listed.” The sample marital status was 88% single, 4% married, 6% cohabiting, and 2% “not listed/do not wish to respond.” Fourteen percent of the sample identified as having Hispanic origin. The mean age was 22.4 ($SD = 4.3$) years. The mean educational level was 14.6 ($SD = 1.7$) years. Four participants (8%) were parents (three participants had one child and one participant had two children).
Power Analysis and Sample Size Planning Information

Because the design of the present study was different from the designs of previous research (Milner et al., 2017) and Study 1, insufficient information was available to estimate the effect size that should be assumed for power analyses. For simplicity and practicality, based on Study 1, twenty-five participants were included in each of the experimental conditions. Given insufficient information to conduct a fully informed a priori power analysis, a sensitivity power analysis was computed. Given the sample size (i.e., 50) and desired Type I/Type II error rate ratio (i.e., .25; Type I error rate = .05, Type II error rate = .20), Study 2 was sufficiently powered to detect both hypothesized mean changes and differences in mean changes (standardized by pooled standard deviation of change scores; \( d_z \)) that were .36 and larger.

Measures and Stimuli

Measures were the same used in Study 1. Except for 11 positive emojis and 11 neutral emojis (described later) as USs in the positive EC and pseudo-EC conditions, respectively, stimuli were the same used in Study 1 and Study 2. In Study 2 Cronbach’s \( \alpha \)s for each dependent measure (pre-EC) were computed and are as follows: positive child evaluations \( \alpha = .78 \), negative child evaluations \( \alpha = .86 \), liking \( \alpha = .79 \), future discipline expectations \( \alpha = .80 \), attributions of hostile intent \( \alpha = .94 \), anger \( \alpha = .90 \), harsh verbal discipline \( \alpha = .90 \), harsh physical discipline \( \alpha = .79 \), doing nothing \( \alpha = .96 \).
Positive Emojis and Neutral Emojis

See Appendix G for emojis used in Study 2. Emojis were approved as part of Unicode 6.0 and copied from emojipedia.org. Many potential positive emojis and neutral emojis could be used; for Study 2 one type of positive emoji (i.e., “happy face emoji;” 😊) and one type of neutral emoji (i.e. “neutral face;” 😞) were selected. Eleven versions of each emoji have been produced by the following (trademarked) corporate entities: Apple, Google, Microsoft, Samsung, LG, HTC, Facebook, Twitter, Mozilla, Emoji One, and Emojidex. For each participant, the 11 versions of the emoji to which participants were randomly assigned (i.e., positive or neutral) were presented as US during the positive EC or pseudo-EC procedure.

Procedure

Except for (1) the differences in the EC stimuli between Study 1 and Study 2, (2) the absence of the backward conditioning condition, and (3) the difference in inter-trial interval (500 ms in Study 2) and number of trials (180 in Study 2), the procedure for Study 2 was the same as in Study 1.

Analytic Strategy

Randomization to experimental conditions ensured that the parameter in the model for any relationship between condition and demographic variables (and unknown/unmeasured variables) is fixed justifiably to zero. With the exception that there is only one a priori contrast of interest per dependent variable (i.e., comparing change scores between the positive EC condition and pseudo-EC condition), the analytic strategy for Study 2 was the same analytic
strategy in Study 1. Additionally, in contrast to Study 1, the MLMs in the present study had only one dummy coded variable comparing the positive EC condition to the pseudo-EC condition.

Results

See Appendix S for Study 2 descriptive statistics for raw data as a function of experimental cell.

Hypothesis Testing

See Appendix T for descriptive statistics for change score data by experimental cell. All hypothesis tests involving pre-EC correlations and a priori comparisons of interest were directional tests (one-tailed $p$-values). For pre-EC correlational hypotheses, the $r$, $p$, and one-sided 95% confidence interval were reported. For each hypothesis test comparing the positive EC condition to the pseudo-EC condition, the $t$-test, $p$-value, mean difference of mean change scores between compared EC conditions (abbreviated $M_{diff}$), and one-sided 95% confidence interval (which corresponds to the information in the directional hypothesis test) were reported. Within EC conditions, the change score mean and standard deviation, $t$-test (one-sample test against zero), one-sided 95% CI for the positive EC condition and 95% CI for the pseudo-EC condition are reported.

Pre-EC Correlational Hypotheses

As hypothesized, pre-EC, positive child evaluations were significantly negatively correlated with attributions of hostile intent, $r = -.27$, $p = .025$, 95% CI[-1.0, -.04]; anger, $r =
Contrary to hypotheses, pre-EC, positive child evaluations were not significantly negatively with harsh physical discipline, $r = .02, p = .45$, 95% CI[-1.0, .25]; harm, $b = -.155, p = .356, RR = .85$, 95% CI[.37, 1.92]; and doing nothing, $r = .09, p = .255$, 95% CI[-1.0, .32].

As hypothesized, pre-EC, liking was significantly positively correlated with positive child evaluations, $r = .63, p < .001$, 95% CI[.46, 1.0]. As hypothesized, pre-EC, liking was significantly negatively correlated with negative child evaluations, $r = -.31, p = .001$, 95% CI [-1.0, -.08]; anger, $r = -.28, p = .025$, 95% CI[-1.0, -.04]; and harsh verbal discipline, $r = -.30, p = .02$, 95% CI[-1.0, -.06]. Contrary to hypotheses, pre-EC liking was not significantly negatively correlated with attributions of hostile intent, $r = -.08, p = .29$, 95% CI[-1.0, .16]; harsh physical discipline, $r = .05, p = .351$, 95% CI[-1.0, .28]; doing nothing, $r = .03, p = .425$, 95% CI[-1.0, .26]; and harm, $b = .02, p = .452$, $RR = 1.02$, 95% CI[.60, 1.74].

As hypothesized, pre-EC, negative child evaluations were significantly positively correlated with attributions of hostile intent, $r = .30, p = .017$, 95% CI[.07, 1.0]; anger, $r = .42, p = .001$, 95% CI[.21, 1.0]; harsh verbal discipline, $r = .40, p = .002$, 95% CI[.18, 1.0]; harsh physical discipline, $r = .26, p = .035$, 95% CI[.02, 1.0]. Contrary to hypotheses, pre-EC, negative child evaluations were not significantly positively correlated with doing nothing, $r = -.06, p = .35$, 95% CI[-.29, 1.0]; and harm, $b = .15, p = .315$, $RR = 1.64$, 95% CI[.62, 2.23].

As hypothesized, pre-EC, attributions of hostile intent were significantly positively correlated with anger, $r = .72, p < .001$, 95% CI[.58, 1.0]; harsh verbal discipline, $r = .54, p < .001$, 95% CI[.35, 1.0]; harsh physical discipline, $r = .33, p = .01$, 95% CI[.11, 1.0]. Contrary to
hypotheses, pre-EC, attributions of hostile intent were not significantly positively predictive of harm, $b = .29$, $RR = 1.33$, $p = .155$, 95% CI[.71, 2.31].

As hypothesized, pre-EC, anger was significantly positively correlated with harsh verbal discipline, $r = .83$, $p < .001$, 95% CI[.74, 1.0]; and harsh physical discipline, $r = .30$, $p = .015$, 95% CI[.07, 1.0]. Contrary to hypotheses, pre-EC, anger was not positively significantly predictive of harm, $b = .37$, $p = .195$, $RR = 1.45$, 95% CI[.58, 3.23].

As hypothesized, pre-EC, harsh verbal discipline was significantly positively correlated with harsh physical discipline, $r = .34$, $p = .008$, 95% CI[.11, 1.0]. Contrary to hypotheses, pre-EC, harsh verbal discipline was not significantly positively predictive of harm, $b = .40$, $p = .19$, $RR = 1.50$, 95% CI[.54, 3.48].

As hypothesized, pre-EC, harsh physical discipline was significantly positively predictive of harm, $b = 5.68$, $p < .001$, $RR = 294.1$, 95% CI[9.4, 7692.7].

Positive Child Evaluations

**Full sample.** As hypothesized, positive EC participants, relative to pseudo-EC participants, evinced significantly greater pre-to-post-EC increases in positive child evaluations, $t(48) = 2.29$, $p = .013$, $M_{diff} = .67$, 95% CI[.18, $\infty$], post hoc power = 74%. Moreover, as hypothesized, in the positive EC condition the mean increase in positive child evaluations was significant, $t(26) = 2.20$, $p = .018$; $M = .51$, $SD = 1.2$, 95% CI[.11, $\infty$]; in the pseudo-EC condition, the mean change in positive child evaluations was not significant, $t(22) = -0.98$, $p = .335$; $M = -.16$, $SD = .80$, 95% CI[-.51, .18].
Reduced sample. Zero participants selected the maximum response value (9) for positive child evaluations pre-EC. Thus, the full sample was not reduced.

Negative Child Evaluations

Full sample. As hypothesized, positive EC participants, relative to pseudo-EC participants, evinced significantly greater pre-to-post-EC reductions in negative child evaluations, \( t(48) = 2.32, p = .012, M_{\text{diff}} = .77, 95\% \text{ CI}[.21, \infty] \), post hoc power = 74%.

Moreover, as hypothesized, in the positive EC condition, the mean reduction in negative child evaluations was significant, \( t(26) = 3.45, p = .001; M = .72, SD = 1.20, 95\% \text{ CI}[.36, \infty] \); in the pseudo-EC condition, the mean change in negative child evaluations was not significant, \( t(22) = -0.22, p = .830; M = -.06, SD = 1.28, 95\% \text{ CI}[-.61, .50] \).

Reduced sample. Zero participants had the minimum response value (0) for negative child evaluations pre-EC. Thus, the full sample was not reduced.

Liking

Full sample. As hypothesized, positive EC participants, relative to pseudo-EC participants, evinced significantly greater pre-to-post-EC increases in liking, \( t(48) = 2.89, p = .003, M_{\text{diff}} = .79, 95\% \text{ CI}[.33, \infty] \), post hoc power = 89%. Contrary to hypotheses, in the positive EC condition, the mean increase in liking was not significant, \( t(26) = 0.96, p = .172; M = .18, SD = 1.00, 95\% \text{ CI}[-.14, \infty] \); in the pseudo-EC condition, the mean change in liking was significant, \( t(22) = -3.12, p = .005; M = -.61, SD = 0.94, 95\% \text{ CI}[-1.01, -.20] \). Thus, the hypothesized EC
condition differences in mean pre-to-post-EC increases in liking were driven mainly by an unexpected significant reduction in liking in the pseudo-EC condition.

**Reduced sample.** Four participants (1 from the positive EC condition and 3 in the pseudo-EC condition) had the maximum response value (9) for liking pre-EC. The reduced sample contained 46 participants. As hypothesized, positive EC, relative to pseudo-EC, had significantly greater pre-to-post-EC increases in liking, \( t(44) = 3.01, p = .002, M_{\text{diff}} = .89, 95\% \text{ CI}[.39, \infty] \), post hoc power = 93%. Contrary to hypotheses, in the positive EC condition, the mean increase in liking was not significant, \( t(25) = 0.96, p = .192; M = .19, SD = 1.02, 95\% \text{ CI}[-.15, \infty] \); in the pseudo-EC condition, the mean change in liking was significant, \( t(22) = -3.22, p = .004; M = -.70, SD = 0.97, 95\% \text{ CI}[-1.16, -.24] \). Thus, the hypothesized EC condition differences in mean pre-to-post-EC increases in liking were driven mainly by an unexpected significant reduction in liking in the pseudo-EC condition.

**Future Discipline Expectations**

**Full sample.** Contrary to hypotheses, positive EC participants, relative to pseudo-EC participants, did not have significantly greater pre-to-post-EC reductions in future discipline expectations, \( t(48) = 0.45, p = .329, M_{\text{diff}} = .17, 95\% \text{ CI}[-.69, \infty] \), post hoc power = 12%. However, as hypothesized in the positive EC condition, the mean reduction in future discipline expectations was significant, \( t(26) = 3.00, p = .003; M = .69, SD = 1.20, 95\% \text{ CI}[.30, \infty] \); in the pseudo-EC condition, the mean change in future discipline expectations was not significant, \( t(22) = 1.68, p = .107; M = .52, SD = 1.49, 95\% \text{ CI}[-.12, 1.17] \).
Reduced sample. Zero participants had the minimum response value (0) for future discipline expectations pre-EC. Thus, the full sample was not reduced.

Attributions of Hostile Intent

Full sample. As hypothesized, positive EC participants, relative to pseudo-EC participants, evinced significantly greater pre-to-post-EC reductions in attributions of hostile intent, \( t(48) = 1.80, p = .039, M_{diff} = .43, 95\% CI [.03, \infty] \), post hoc power = 56\%. Contrary to hypotheses, in the positive EC condition the mean reduction in attributions of hostile intent was not significant, \( t(26) = 1.50, p = .073; M = .28, SD = .97, 95\% CI [-.04, \infty] \); in the pseudo-EC condition, the mean change in attributions of hostile intent was not significant, \( t(22) = -1.08, p = .290; M = -.15, SD = .68, 95\% CI [-.45, .14] \).

The non-significant changes within EC conditions were descriptively different from zero. The distance of these changes (i.e., \( M_{diff} = .43 \)), however was significant. Therefore, the hypothesis was not supported.

Reduced sample. Three participants (one from the positive EC condition and two from the pseudo-EC condition) had the minimum response value (0) for attributions of hostile intent pre-EC. The reduced sample contained 47 participants. As hypothesized, positive EC, relative to pseudo-EC, had significantly greater pre-to-post-EC reductions in attributions of hostile intent, \( t(45) = 1.76, p = .043, M_{diff} = .45, 95\% CI [.02, \infty] \), post hoc power = 56\%. Contrary to hypotheses, in the positive EC condition, the mean reduction in attributions of hostile intent was not significant, \( t(25) = 1.50, p = .073; M = .29, SD = .99, 95\% CI [-.04, \infty] \); in the pseudo-EC
condition, the mean change in attributions of hostile intent was not significant, $t(22) = -1.03, p = .314; M = -.16, SD = 0.72, 95\% \text{ CI}[-.49, .16]$.

**Anger**

**Full sample.** Contrary to hypotheses, positive EC participants, relative to pseudo-EC participants, did not evince significantly greater pre-to-post-EC reductions in anger, $t(48) = 1.56, p = .063, M_{\text{diff}} = .25, 95\% \text{ CI}[-.02, \infty]$; post hoc power = 47%. However, as hypothesized, in the positive EC condition the mean reduction in anger was significant, $t(26) = 2.34, p = .014; M = .27, SD = .60, 95\% \text{ CI}[,07, \infty]$; in the pseudo-EC condition, the mean change in anger was not significant, $t(22) = 0.18, p = .860; M = .02, SD = .52, 95\% \text{ CI}[-.21, .25]$.

**Reduced sample.** Three participants (two from the positive EC condition and two from the pseudo-EC condition) had the minimum response value (0) for anger pre-EC. The reduced sample contained 47 participants. Contrary to hypotheses, positive EC participants, relative to pseudo-EC participants, did not evince significantly greater pre-to-post-EC reductions in anger, $t(45) = 1.59, p = .059, M_{\text{diff}} = .27, 95\% \text{ CI}[-.01, \infty]$; post hoc power = 48%. However, as hypothesized, in the positive EC condition, the mean reduction in anger was significant, $t(24) = 2.35, p = .014; M = .29, SD = .62, 95\% \text{ CI}[,08, \infty]$; in the pseudo-EC condition, the mean change in anger was not significant, $t(21) = 0.18, p = .862; M = .02, SD = .52, 95\% \text{ CI}[-.22, .26]$.

**Harsh Verbal Discipline**

**Full sample.** Contrary to hypotheses, positive EC participants, relative to pseudo-EC participants, did not evince significantly greater pre-to-post-EC reductions in harsh verbal
discipline, $t(48) = 1.38, p = .086, M_{diff} = .13, 95\% \text{ CI}[-.03, \infty], \text{ post hoc power} = 40\%$. However, as hypothesized, in the positive EC condition the mean reduction in harsh verbal discipline was significant, $t(26) = 2.57, p = .008; M = .17, SD = .34, 95\% \text{ CI} [.06, \infty]$; in the pseudo-EC condition, the mean change in harsh verbal discipline was not significant, $t(22) = 0.59, p = .558; M = .04, SD = .31, 95\% \text{ CI} [-.10, .17].$

**Reduced sample.** Twelve participants had the minimum response value (0) for harsh verbal discipline pre-EC. The reduced sample contained 38 participants. Contrary to hypotheses, positive EC participants, relative to pseudo-EC participants, did not evince significantly greater pre-to-post-EC reductions in harsh verbal discipline, $t(36) = 1.08, p = .142, M_{diff} = .12, 95\% \text{ CI} [-.07, \infty], \text{ post hoc power} = 29\%$. As hypothesized, in the positive EC condition, the mean reduction in harsh verbal discipline was significant, $t(20) = 2.65, p = .008; M = .21, SD = .37, 95\% \text{ CI} [.07, \infty]$; in the pseudo-EC condition, the mean change in harsh verbal discipline was not significant, $t(16) = 1.20, p = .248; M = .09, SD = .31, 95\% \text{ CI} [-.07, .25].$

**Harsh Physical Discipline**

**Full sample.** Contrary to hypotheses, positive EC participants, relative to pseudo-EC participants, did not evince significantly greater pre-to-post-EC reductions in harsh physical discipline, $t(48) = 0.53, p = .300, M_{diff} = .02, 95\% \text{ CI}[-.05, \infty], \text{ post hoc power} = 13\%$. Moreover, contrary to hypotheses, in the positive EC condition the mean reduction in harsh physical discipline was not significant, $t(26) = -0.40, p = .652; M = -.01, SD = .16, 95\% \text{ CI} [.06, \infty]$; in the
pseudo-EC condition, the mean change in harsh physical discipline was not significant, \( t(22) = -1.39, p = .179; M = -.03, SD = .12, 95\% \text{ CI}[-.08, .02]. \)

**Reduced sample.** Forty-six participants (23 from the positive EC condition and 23 from the pseudo-EC condition) had the minimum response value (0) for harsh physical discipline pre-EC. The reduced sample contained four participants (all from the positive EC condition). Contrary to hypotheses, in the positive EC condition, the mean reduction in harsh physical discipline was not significant, \( t(3) = -0.36, p = .627; M = -.08, SD = .47, 95\% \text{ CI}[-.63, \infty]. \)

**Doing Nothing**

**Full sample.** As hypothesized, positive EC participants, relative to pseudo-EC participants, evinced significantly greater pre-to-post-EC increases in doing nothing, \( t(48) = 2.33, p = .012, M_{\text{diff}} = .55, 95\% \text{ CI}[.15, \infty], \) post hoc power = 76%. Moreover, as hypothesized, in the positive EC condition the mean increase in doing nothing was significant, \( t(26) = 3.40, p = .001; M = .46, SD = .70, 95\% \text{ CI}[,23, \infty]; \) in the pseudo-EC condition, the mean change in doing nothing was not significant, \( t(22) = -0.45, p = .658; M = -.09, SD = .96, 95\% \text{ CI}[-.50, .32]. \)

**Reduced sample.** Zero participants had the maximum response value (0) for doing nothing pre-EC. Thus, the full sample was not reduced.
Harm

**Full sample.** Contrary to hypotheses, positive EC participants, relative to pseudo-EC participants, did not evince significantly greater pre- to post-EC reductions in harm, $t(48) = 0.47, p = .322, M_{\text{diff}} = .13, 95\% \text{ CI}[-.34, \infty]$, post hoc power = 12%. Also contrary to hypotheses, in the positive EC condition the mean reduction in harm was not significant, $t(26) = 0.00, p = .500$; $M = .00, SD = .28, 95\% \text{ CI}[-.09, \infty]$; in the pseudo-EC condition, the mean change in harm was not significant, $t(22) = -0.44, p = .665$; $M = -.13, SD = 1.42, 95\% \text{ CI}[-.75, .48]$.

**Reduced sample.** Forty-four participants (23 from the positive EC condition and 21 from the pseudo-EC condition) had the minimum response value (0) for harm pre-EC. The reduced sample contained six participants. Contrary to hypotheses, positive EC participants, relative to pseudo-EC participants, did not evince significantly greater pre- to post-EC reductions in harm, $t(4) = 0.53, p = .312, M_{\text{diff}} = 1.5, 95\% \text{ CI}[-4.52, \infty]$, post hoc power = 6%. Contrary to hypotheses, in the positive EC condition, the mean reduction in harm was not significant, $t(3) = 0.00, p = .500$; $M = .00, SD = .82, 95\% \text{ CI}[-.96, \infty]$; in the pseudo-EC condition, the mean change in harm was not significant, $t(1) = -0.33, p = .795$; $M = -1.50, SD = 6.36, 95\% \text{ CI}[-58.68, 55.68]$.

**Planned Exploratory Analyses**

**Analysis of Experimental Design Using Repeated-Measures ANOVA**

While the pre-planned comparisons were focused on the exact hypotheses proposed, additional information about Study 2 can be obtained by examining results for separate EC
Condition (positive EC, pseudo-EC) × Time (pre-EC, post-EC) ANOVAs with repeated measures on the second variable performed on each dependent measure.

**Positive child evaluations.** The EC Condition × Time interaction was significant, $F(1, 48) = 5.22, p = .027, \eta_p^2 = .10, 95\% CI[0.01, 0.24]$. The interaction was probed by examining all simple effects. In the positive EC condition, when pre-EC positive child evaluations ($M = 4.79, SD = 0.81$) and post-EC positive child evaluations ($M = 5.30, SD = 1.37$) were compared, the difference was significant, $t(26) = 2.20, p = .018, 95\% CI[.11, \infty]$. In the pseudo-EC condition, when pre-EC positive child evaluations ($M = 4.73, SD = 1.30$) and post-EC positive child evaluations ($M = 4.57, SD = 1.67$) were compared, the difference was not significant, $t(22) = -0.98, p = .335, 95\% CI [-.51, .18]$. Pre-EC positive child evaluations did not significantly vary between EC condition, $F(1, 48) = 0.03, p = .854, \eta_p^2 = .00, 95\% CI[.00, .04]$; positive EC $M = 4.79, SD = 0.81$; pseudo-EC $M = 4.73, SD = 1.30$.

Post-EC positive child evaluations did not significantly vary between EC conditions, $F(1, 48) = 2.86, p = .097, \eta_p^2 = .06, 95\% CI[.00, .18]$; positive EC $M = 5.30, SD = 1.37$; pseudo-EC $M = 4.57, SD = 1.67$.

The main effect of EC condition was not significant, $F(1, 48) = 1.31, p = .257, \eta_p^2 = .03, 95\% CI[0.00, 0.14]$; positive EC $M = 5.04, SD = 0.95$; pseudo-EC $M = 4.65, SD = 1.44$.

The main effect of time was not significant, $F(1, 48) = 1.36, p = .249, \eta_p^2 = .03, 95\% CI[0.00, 0.14]$; pre-EC $M = 4.76, SD = 1.05$, post-EC $M = 4.96, SD = 1.54$.

**Negative child evaluations.** The EC Condition × Time interaction was significant, $F(1, 48) = 5.39, p = .024, \eta_p^2 = .10, 95\% CI[.01, .24]$. The interaction was probed by examining all
simple effects. In the positive EC condition, when pre-EC negative child evaluations ($M = 2.93$, $SD = 1.34$) and post-EC negative child evaluations ($M = 2.22$, $SD = 1.44$) were compared, the difference was significant, $t(26) = 3.45$, $p = .001$, 95% CI [.36, $\infty$]. In the pseudo-EC condition, when pre-EC negative child evaluations ($M = 2.74$, $SD = 1.56$) and post-EC negative child evaluations ($M = 2.80$, $SD = 2.00$) were compared, the difference was not significant, $t(22) = 0.22$, $p = .830$, 95% CI [−.50, .62].

Pre-EC negative child evaluations did not significantly vary between EC condition, $F(1, 48) = 0.23$, $p = .636$, $\eta^2_p = .00$, 95% CI [.00, .08]; positive EC $M = 2.93$, $SD = 1.34$; pseudo-EC $M = 2.74$, $SD = 1.56$.

Post-EC negative child evaluations did not significantly vary between EC conditions, $F(1, 48) = 1.41$, $p = .241$, $\eta^2_p = .03$, 95% CI [.00, .14]; positive EC $M = 2.22$, $SD = 1.44$; pseudo-EC $M = 2.80$, $SD = 2.00$.

The main effect of EC condition was not significant, $F(1, 48) = 1.31$, $p = .257$, $\eta^2_p = .03$, 95% CI [0.00, 0.14]; positive EC $M = 2.58$, $SD = 1.28$; pseudo-EC $M = 2.77$, $SD = 1.68$.

The main effect of time was not significant, $F(1, 48) = 3.90$, $p = .054$, $\eta^2_p = .10$, 95% CI [0.01, 0.24]; pre-EC $M = 2.84$, $SD = 1.43$, post-EC $M = 2.48$, $SD = 1.73$.

**Liking.** The EC Condition × Time interaction was significant, $F(1, 48) = 8.33$, $p = .006$, $\eta^2_p = .15$, 95% CI [.03, .30]. The interaction was probed by examining all simple effects. In the positive EC condition, when pre-EC liking ($M = 5.99$, $SD = 1.38$) and post-EC liking ($M = 6.17$, $SD = 1.46$) were compared, the difference was significant, $t(26) = 0.97$, $p = .172$, 95% CI [−.14, $\infty$]. In the pseudo-EC condition, when pre-EC liking ($M = 6.13$, $SD = 1.98$) and post-EC liking
(M = 5.52, SD = 2.19) were compared, the difference was significant, \( t(22) = 3.12, p = .005, 95\% \text{ CI [-1.01, -0.20]} \).

Pre-EC liking did not significantly vary between EC condition, \( F(1, 48) = 0.09, p = .766, \eta^2_p = .00, 95\% \text{ CI [.00, .06]} \); positive EC \( M = 5.99, SD = 1.38 \); pseudo-EC \( M = 6.13, SD = 1.98 \).

Post-EC liking did not significantly vary between EC conditions, \( F(1, 48) = 1.57, p = .216, \eta^2_p = .03, 95\% \text{ CI [.00, .14]} \); positive EC \( M = 6.17, SD = 1.44 \); pseudo-EC \( M = 5.52, SD = 2.00 \).

The main effect of EC condition was not significant, \( F(1, 48) = 0.28, p = .599, \eta^2_p = .01, 95\% \text{ CI [0.00, 0.08]} \); positive EC \( M = 6.08, SD = 1.33 \); pseudo-EC \( M = 5.83, SD = 2.04 \).

The main effect of time was not significant, \( F(1, 48) = 2.37, p = .130, \eta^2_p = .05, 95\% \text{ CI [.00, 0.17]} \); pre-EC \( M = 6.05, SD = 1.67 \); post-EC \( M = 5.87, SD = 1.84 \).

Future discipline expectations. The EC Condition × Time interaction was not significant, \( F(1, 48) = 0.20, p = .657, \eta^2_p = .00, 95\% \text{ CI [0.00, 0.08]} \).

The main effect of EC condition was not significant, \( F(1, 48) = 1.08, p = .303, \eta^2_p = .02, 95\% \text{ CI [0.00, 0.32]} \); positive EC \( M = 3.95, SD = 1.74 \); pseudo-EC \( M = 4.54, SD = 2.23 \).

The main effect of time was significant, \( F(1, 48) = 10.18, p = .002, \eta^2_p = .18, 95\% \text{ CI [.04, 0.17]} \); pre-EC \( M = 4.53, SD = 1.90 \); post-EC \( M = 3.91, SD = 2.27 \).

Attributions of hostile intent. The EC Condition × Time interaction was not significant, \( F(1, 48) = 3.24, p = .078, \eta^2_p = .06, 95\% \text{ CI [.00, .19]} \).

The main effect of EC condition was not significant, \( F(1, 48) = 0.03, p = .857, \eta^2_p = .00, 95\% \text{ CI [0.00, 0.04]} \); positive EC \( M = 1.58, SD = 1.44 \); pseudo-EC \( M = 1.51, SD = 1.08 \).
The main effect of time was not significant, $F(1, 48) = 0.27, p = .606, \eta^2 = .00, 95\% CI[.00, 0.08]$; pre-EC $M = 1.59, SD = 1.34$; post-EC $M = 1.51, SD = 1.36$.

**Anger.** The EC Condition $\times$ Time interaction was not significant, $F(1, 48) = 2.42, p = .126, \eta^2 = .05, 95\% CI[.00, .17]$.

The main effect of EC condition was not significant, $F(1, 48) = 0.30, p = .585, \eta^2 = .01, 95\% CI[.00, .08]$; positive EC $M = 1.11, SD = 0.97$; pseudo-EC $M = 0.96, SD = 0.93$.

The main effect of time was not significant, $F(1, 48) = 3.23, p = .078, \eta^2 = .06, 95\% CI[.00, 0.19]$; pre-EC $M = 1.11, SD = 0.93$; post-EC $M = 0.96, SD = 1.04$.

**Harsh verbal discipline.** The EC Condition $\times$ Time interaction was not significant, $F(1, 48) = 1.92, p = .172, \eta^2 = .04, 95\% CI[.00, .16]$.

The main effect of EC condition was not significant, $F(1, 48) = 0.77, p = .384, \eta^2 = .02, 95\% CI[.00, .11]$; positive EC $M = 0.72, SD = 0.80$; pseudo-EC $M = 0.52, SD = 0.81$.

The main effect of time was significant, $F(1, 48) = 4.93, p = .031, \eta^2 = .09, 95\% CI[.00, 0.23]$; pre-EC $M = 0.68, SD = 0.82$; post-EC $M = 0.57, SD = 0.82$.

**Harsh physical discipline.** The EC Condition $\times$ Time interaction was not significant, $F(1, 48) = 0.28, p = .599, \eta^2 = .01, 95\% CI[.00, .08]$.

The main effect of EC condition was not significant, $F(1, 48) = 1.22, p = .274, \eta^2 = .02, 95\% CI[.00, .13]$; positive EC $M = 0.06, SD = 0.20$; pseudo-EC $M = 0.02, SD = 0.06$.

The main effect of time was not significant, $F(1, 48) = 1.29, p = .261, \eta^2 = .03, 95\% CI[.00, 0.13]$; pre-EC $M = 0.03, SD = 0.12$; post-EC $M = 0.05, SD = 0.21$.

**Doing nothing.** The EC Condition $\times$ Time interaction was significant, $F(1, 48) = 5.40, p = .024, \eta^2 = .10, 95\% CI[0.01, 0.24]$. The interaction was probed by examining all simple
effects. In the positive EC condition, when pre-EC doing nothing ($M = 5.78$, $SD = 2.53$) and post-EC doing nothing ($M = 6.26$, $SD = 2.74$) were compared, the difference was significant, $t(26) = 3.54$, $p = .001$, 95% CI [.25, $\infty$]. In the pseudo-EC condition, when pre-EC doing nothing ($M = 6.44$, $SD = 1.73$) and post-EC doing nothing ($M = 6.38$, $SD = 2.01$) were compared, the difference was not significant, $t(22) = -0.34$, $p = .736$, 95% CI [-.48, .34].

Pre-EC doing nothing did not significantly vary between EC condition, $F(1, 48) = 1.14$, $p = .292$, $\eta^2_p = .02$, 95% CI [.00, .13]; positive EC $M = 5.78$, $SD = 2.53$; pseudo-EC $M = 6.44$, $SD = 1.73$.

Post-EC doing nothing did not significantly vary between EC conditions, $F(1, 48) = 0.03$, $p = .861$, $\eta^2_p = .00$, 95% CI [.00, .04]; positive EC $M = 6.26$, $SD = 2.74$; pseudo-EC $M = 6.38$, $SD = 2.01$.

The main effect of EC condition was not significant, $F(1, 48) = 0.37$, $p = .546$, $\eta^2_p = .01$, 95% CI [.00, .09]; positive EC $M = 6.02$, $SD = 2.61$; pseudo-EC $M = 6.41$, $SD = 1.81$.

The main effect of time was not significant, $F(1, 48) = 3.04$, $p = .087$, $\eta^2_p = .06$, 95% CI [.00, .19]; pre-EC $M = 6.09$, $SD = 2.20$, post-EC $M = 6.31$, $SD = 2.41$.

Harm. The EC Condition × Time interaction was not significant, $F(1, 48) = 0.22$, $p = .643$, $\eta^2_p = .01$, 95% CI [.00, .08].

The main effect of EC condition was not significant, $F(1, 48) = 0.00$, $p = .983$, $\eta^2_p = .00$, 95% CI [.00, 1.00]; positive EC $M = 0.70$, $SD = 2.12$; pseudo-EC $M = 0.72$, $SD = 2.44$.

The main effect of time was not significant, $F(1, 48) = 0.22$, $p = .643$, $\eta^2_p = .01$, 95% CI [.00, .08]; pre-EC $M = 0.68$, $SD = 2.11$, post-EC $M = 0.74$, $SD = 2.47$. 
Although the time main effect and the interaction between time and EC condition might appear to be reported erroneously, they are not reported erroneously. This is because one mean difference (i.e., the mean difference in the pseudo-EC condition) pre-to-post-EC is equal to exactly zero, so the overall time effect is equal to the change in the time effect as a function of EC condition.

Post-EC Differences Between EC Conditions for Each Dependent Variable, Estimated at the Mean of Pre-EC Scores (ANCOVA)

It can be desirable to confirm the results of analyses that use difference scores by using alternative statistical methods to explore the data. This was done by using an ANCOVA method. As noted below, the results from the ANCOVA generally duplicated results from the difference score analysis.

Post-EC positive child evaluations. After adjusting for pre-EC positive child evaluations ($M = 4.76$), the effect of EC condition on post-EC positive child evaluations was significant, $F(1, 47) = 5.09, p = .029$; positive EC $LSM = 5.27, SE = 0.20$; pseudo-EC $LSM = 4.60, SE = 0.22$.

Post-EC negative child evaluations. After adjusting for pre-EC negative child evaluations ($M = 2.84$), the effect of EC condition on post-EC negative child evaluations was significant, $F(1, 47) = 5.05, p = .029$; positive EC $LSM = 2.14, SE = 0.23$; pseudo-EC $LSM = 2.89, SE = 0.24$. 

Post-EC liking. After adjusting for pre-EC liking ($M = 6.05$), the effect of EC condition on post-EC liking was significant, $F(1, 47) = 8.06, p = .007$; positive EC $LSM = 6.23, SE = 0.19$; pseudo-EC $LSM = 5.45, SE = 0.20$.

Post-EC future discipline expectations. After adjusting for pre-EC future discipline expectations ($M = 6.05$), the effect of EC condition on post-EC future discipline expectations was not significant, $F(1, 47) = 0.24, p = .629$; positive EC $LSM = 3.83, SE = 0.26$; pseudo-EC $LSM = 4.02, SE = 0.28$.

Post-EC attributions of hostile intent. After adjusting for pre-EC attributions of hostile intent ($M = 1.59$), the effect of EC condition on post-EC attributions of hostile intent was not significant, $F(1, 47) = 2.65, p = .110$; positive EC $LSM = 1.33, SE = 0.16$; pseudo-EC $LSM = 1.71, SE = 0.17$.

Post-EC anger. After adjusting for pre-EC anger ($M = 1.12$), the effect of EC condition on post-EC anger was not significant, $F(1, 47) = 2.08, p = .156$; positive EC $LSM = 0.86, SE = 0.11$; pseudo-EC $LSM = 1.09, SE = 0.12$.

Post-EC harsh verbal discipline. After adjusting for pre-EC harsh verbal discipline ($M = 0.68$), the effect of EC condition on post-EC harsh verbal discipline was not significant, $F(1, 47) = 1.35, p = .250$; positive EC $LSM = 0.52, SE = 0.06$; pseudo-EC $LSM = 0.63, SE = 0.07$.

Post-EC harsh physical discipline. After adjusting for pre-EC harsh physical discipline ($M = 0.03$), the effect of EC condition on post-EC harsh physical discipline was not significant, $F(1, 47) = 1.70, p = .199$; positive EC $LSM = 0.03, SE = 0.03$; pseudo-EC $LSM = 0.08, SE = 0.03$. 
Post-EC doing nothing. After adjusting for pre-EC doing nothing ($M = 6.10$), the effect of EC condition on post-EC doing nothing was significant, $F(1, 47) = 5.68, p = .021$; positive EC $LSM = 6.57, SE = 0.16$; pseudo-EC $LSM = 6.00, SE = 0.17$.

Post-EC harm. After adjusting for pre-EC harm ($M = 0.68$), the effect of EC condition on post-EC harm was not significant, $F(1, 47) = 0.23, p = .632$; positive EC $LSM = 0.68, SE = 0.19$; pseudo-EC $LSM = 0.81, SE = 0.20$.

**Linear Relationships Between Pre-EC Scores and EC Change Scores Within the Positive EC Condition**

Appendix V presents Study 2 correlations between dependent variable pre-EC scores and corresponding change scores within the positive EC condition.

**Linear Relationships Between Continuous Demographic Characteristics and EC Change Scores Within the Positive EC Condition**

All participants had zero children in the positive EC condition. After applying Bonferroni-Holm correction for multiple testing (22 correlations, corrected $\alpha = .002$), age and education were not significantly correlated with EC change scores for any dependent variable.

**EC Change Scores as a Function of Demographic Characteristics Within the Forward Positive EC Condition and the Backward Positive EC Condition**
After applying Bonferroni-Holm correction for multiple testing (10 tests, corrected $\alpha = 0.005$), no significant differences in EC change scores emerged for all nominal demographic variables (i.e., gender, race, and marital status).

Binomial Test for Harm

Among the 44 participants who gave no harm EC, zero participants gave harm post-EC (binomial test: $p = 1.0$, null hypothesis: true probability = 0).

Multi-Level Model Analyses

Results of MLMs for Study 2 are presented in Appendices W-EE. These results are presented for informational purposes only; no inferences were made with respect to generalizing estimates of positive EC condition predictors of pre-to-post EC change estimates (reference group = pseudo-EC condition).

Discussion

Pre-EC correlations between study variables generally supported hypotheses. Specifically, positive child evaluations were significantly negatively correlated with attributions of hostile intent, anger, and harsh verbal discipline. However, positive child evaluations were positively correlated with doing nothing and were not negatively correlated with harsh physical discipline and harm. It is unknown why all pre-EC correlations that involved positive child evaluations were not significant, given that many of the same correlations were significant in the
preliminary study. However, based on mostly supported hypotheses, it would be unwarranted to conclude that the measure of pre-EC positive child evaluations did not reflect the positive child evaluation construct. Moreover, that positive child evaluations were significantly correlated with liking provides confidence that positive child evaluations and liking are heavily evaluatively laden. Yet the fact that the correlation was .63 rather than .90 or higher is consistent with the view that positive child evaluations and liking are not synonymous. It is thought that positive child evaluations contain a semantic component that is not an attribute of the measure of liking. Additional support for this idea comes from the finding that liking was not related to attributions of hostile intent, which is a variable that may be more driven by meaning, such as that the child is positive, friendly, and cooperative, than may be the case for positive child evaluations.

As hypothesized, pre-EC negative child evaluations were significantly positively correlated with attributions of hostile intent, anger, harsh verbal discipline, and harsh physical discipline. Although negative child evaluations did not correlate with doing nothing or harm, that most of the hypothesized relationships involving negative child evaluations were found suggests construct validity for the negative child evaluation measure.

As hypothesized, pre-EC attributions of hostile intent, anger, harsh verbal discipline, and harsh physical discipline were all significantly positively correlated. However, only harsh physical discipline predicted harm in Study 1 and Study 2. Based on this pattern of results, it is concluded that the harsh physical discipline measure shares many conceptual commonalities with the measure of harm (compared to other dependent measures), especially with respect to cognitive and affective processes that may contribute to responses. Indeed, support for this idea comes from research (Fréchette, Zoratti, & Romano, 2015) showing a positive relationship
between self-reported receipt of spanking in childhood and receipt of child physical abuse in childhood; people who reported receiving spanking in childhood were 60 times more likely than those who reported receiving no spanking in childhood to have received child physical abuse.

As hypothesized, positive EC, relative to pseudo-EC, produced significantly greater increases in positive child evaluations from pre- to post-EC. Within the positive EC condition, significant increases in positive child evaluations pre- to post-EC were observed, whereas no such increases were observed in the forward pseudo-EC condition. As in Study 1, this finding is consistent with EC research (e.g., see Hoffman et al., 2010, for meta-analyses of EC effects), showing that, broadly, CS-US pairings involved in EC procedures cause CS valence to “take on” the valence of the US. In addition, the positive EC effect on positive child evaluations replicate prior research (Milner et al., 2017, Experiment 6).

Two mechanisms were proposed to explain the EC effects observed by Milner et al. (2017): a semantic learning mechanism and an evaluative learning mechanism. The semantic learning mechanism differs from evaluative mechanisms proposed in EC research (Hofmann et al., 2010). Study 2 was conducted to learn the extent to which the Milner et al. EC effects can be explained by semantic learning. If semantic learning mediated EC effects observed by Milner et al., then the US used in the Milner et al. studies (the trait words “sweet,” “friendly,” and “cooperative”) taught that the CS (ambiguous child images) had the trait-like properties of being sweet, friendly, and cooperative.

In contrast, if evaluative learning mediated the effects observed by Milner et al. (2017), the EC procedures increased liking for the CS (ambiguous child images) directly. That is, evaluative mental representations of, e.g., “liking,” “good,” “positive,” “pleasant,” were
activated and influence judgments of the CS. Further, different social cognitive processes (discussed previously, e.g., associative learning, implicit misattribution, propositional learning) may underlie the influence of evaluative mental representations on evaluative judgments of the CS. Although Study 2 could not disambiguate these social cognitive processes, the effects of positive EC (with positive emojis as US) on positive child evaluations suggests that semantic learning alone cannot well explain observed effects in Study 2 and in prior research (Milner et al., 2017).

Based on the contentions described above, the positive EC effect on positive child evaluations is unlikely to be explained by the propositional account because the propositional account implies that semantic knowledge must be formulated or retrieved and subsequently evaluated. Relative to using positive trait words (as in Study 1), the positive emojis in Study 2 are likely not to elicit such propositional knowledge directly. Instead the positive emojis were devoid of semantic content but were designed to express pleasantness. As stated earlier evaluative representations (i.e., the mental representation of “good,” “pleasant”) in memory are assumed to be activated when viewing the positive emojis. Thus positive EC effects on positive child evaluations suggest that the results observed by Milner et al. and Study 1 cannot be fully explained by mechanisms posited in the propositional account of EC.

Finally, the positive EC effect on positive child evaluations are consistent with the referential account of EC because the ambiguous child images (CS) preceded the (US). However, the referential account cannot be the only mechanism driving EC effects in the general EC literature. This conclusion is based on evidence provided in Study 1 and prior research that used backward conditioning procedures (Hofmann et al., 2010).
Despite this contribution to understanding social cognitive mechanisms that underlie effects observed in Milner et al. (2017) and in Study 1, Study 2 could not disambiguate the extent to which the formation of referential or predictive associations drive EC effects in general EC research. Indeed, future research is needed to develop or use existing procedures that are better able to measure the extent to which specific processes mediate observed so-called “EC effects.” Although methods in cognitive psychology and social cognition (e.g., subliminal presentation, backward stimuli presentation, task instructions, Hofmann et al., 2010) can provide suggestive evidence that particular process(es) contribute to observed effects (see Spencer, Zanna, & Fong, 2005) for a discussion on moderation-as-process designs), no methods have been used in general EC literature to provide direct measures of such underlying processes.

As hypothesized, positive EC, relative to pseudo-EC, produced significantly greater reductions in negative child evaluations. Milner et al. (2017) sought to change negative child evaluations using positive EC. Negative child evaluations were of central interest because, according to the SIP model of CPA, negative child evaluations are a component of “pre-existing” schemata that abusive and high-risk parents bring to parent-child social interactions. The negative schemata in abusive or high-risk caregivers further influence social information processing across three social cognitive stages and a behavioral response implementation stage. When interpreting and evaluating ambiguous child behaviors during a social interaction, high-risk or abusive parents’ negative child evaluations increase the likelihood that the high-risk or abusive parents infer or attribute antagonistic inferences (i.e., attributions of hostile intent). Thus, negative child evaluations were a target of manipulation in Milner et al., and the Milner et al. research showed that positive EC was effective at reducing negative child evaluations.
Given that positive trait words were used as US in Milner et al. (2017), it was plausible that the effects observed in Milner et al. were caused by semantic learning rather than the evaluative learning posited in general EC research. In other words, the positive EC might have not reduced negative child evaluations directly but instead taught participants that the children depicted in ambiguous child images were sweet, friendly, and cooperative.

The positive EC effects (when using positive emojis, relatively devoid of semantic content) on both positive child evaluations and negative child evaluations suggests that semantic learning cannot fully explain the effects observed in Milner et al. (2017). Thus, Study 2 results contribute to our understanding of plausible social cognitive mechanisms underlying EC effects observed in Milner et al and Study 1.

As hypothesized, positive EC, relative to neutral EC, produced significantly greater reductions in attributions of hostile intent. This finding replicates prior research using a similar positive EC procedure to reduce attributions of hostile intent (Milner et al., 2017, Experiments 3-6). When compared to previously reviewed interventions targeting attributions of hostile intent, administration time for positive EC (approximately 4 min) is briefer than (a) the administration time for Bugental et al.’s (2002) enhanced home visitation condition (cognitive appraisal component was an unknown time duration every home visit \(M = 17\) visits) and (b) the administration time for Rutledge’s (2016) IPT intervention (approximately 25 min). Positive EC’s brief administration time is a desirable property for augmenting existing interventions. Future research might compare the effectiveness (short and long term) of an existing intervention to the effectiveness of the same existing intervention augmented with positive EC.
Contrary to hypotheses, positive EC, relative to pseudo-EC, did not significantly reduce anger, harsh verbal discipline, harsh physical discipline, or harm. However, in the positive EC condition there were significant pre-to-post reductions for anger and harsh verbal discipline. Previous research has shown that positive EC reduces anger, harsh verbal discipline, harsh physical discipline, and harm (Milner et al., 2017, Experiment 6). Second, positive EC-induced reductions in anger, harsh verbal discipline, and harsh physical discipline were in the expected directions.

As in Study 1, pre-EC means for anger, harsh verbal discipline, and harsh physical discipline were low, in part because many participants endorsed the minimum value for each of these variables. Similarly, only six participants gave any harm pre-EC. These “floor” values substantially reduce the ability to test hypotheses regarding positive EC-induced reductions in harsh verbal discipline, harsh physical discipline, and harm because by definition minimum values cannot be reduced (in fact, the minimum values obviously can only increase).

Previous research (Milner et al., 2017) found that positive EC (180 trials) significantly reduced anger, harsh verbal discipline, harsh physical discipline, and harm. In Milner et al., Experiment 6, the pre-EC means for these variables were on a 1 to 10 response scale: anger $M = 1.97$, harsh verbal discipline $M = 1.67$, harsh physical discipline $M = 1.28$, harm $M = 1.90$. It is unknown what produced the especially high frequencies of minimum values on these dependent variables in the present study and in Milner et al.

Future research might avoid these problems by recruiting participants who are likely to have non-minimum pre-EC scores for anger, harsh verbal discipline, harsh physical discipline, and harm. One fruitful option would be to recruit individuals at high-risk for CPA. Unpublished
correlational data show some evidence for positive relationships between CPA risk status (Abuse scale low risk cutoff = 91, Abuse scale high risk cutoff = 215; invalid CAPs excluded; included N = 178) and anger ($r = .26, p = .001$), harsh verbal discipline ($r = .19, p = .010$), harsh physical discipline ($r = .12, p = .104$), and harm (negative binomial regression $b = 1.48, p = .015, RR = 4.38$).
CHAPTER 5
GENERAL LIMITATIONS AND SUMMARY

There were several general limitations across Study 1 and Study 2. Although Study 1 ($N = 77$) had 80% power to detect effects (standardized pre-to-post changes, $d_z$) that were greater than or equal to .29, and Study 2 had 80% power to detect effects that were greater than or equal to $d_z = .36$, tests for within EC condition effects may have been underpowered. For example in Study 1, the forward positive EC condition ($n = 26$) had 80% power to detect effects that were greater than or equal to $d_z = .50$. Thus, for some dependent variables, standardized pre-to-post differences, although in the expected directions, were not significant: liking $d_z = .28$, attributions of hostile intent $d_z = .19$, harsh verbal discipline $= .25$, doing nothing $d_z = .29$.

Another general limitation is that both Study 1 and Study 2 differ from the Milner et al. (2017) studies in ways that make it difficult or impossible to compare the magnitude of EC effects across studies. Except for a dosage study (Milner et al., 2017, Experiment 6), in Milner et al. positive EC contained 180 trials with a 500 ms inter-trial interval and used positive trait words as US. In contrast to Milner et al., Study 1 used 18 trials with a 12 s inter-trial interval in order to preclude backward conditioning from occurring in the forward positive EC condition. Although Milner et al., Experiment 6 explored whether EC dose (1 EC = 180 trials vs. 2 EC = 360 trials) accentuates positive EC-induced changes and found no EC dose effect, it is possible
that low doses (e.g., 18 trials in Study 1) are associated with comparatively weak EC-induced changes.

In contrast to Milner et al. (2017), Study 2 used as US positive emojis instead of positive trait words. However, comparison of the present results to those of Miller et al. is difficult because no valence data have been collected (a) for the US in Milner et al.’s positive EC procedures (i.e., positive trait words) and (b) for the US in Study 1 (positive trait words) and in Study 2 (positive emojis). Future research might present participants with positive trait words and positive emojis used as US in the EC research. For each item, participants might evaluate the US using an item such as, “To what extent does this word/emoji make you feel positive?” Then the valence of the US used in Study 1, Study 2, and previous research (Milner et al., 2017) could be compared across EC paradigms. If the positive trait words and positive emojis have similar positive valence, then EC effects can be compared across these conditions (positive trait words as US vs. positive emojis as US). If the positive trait words and positive emojis do not have similar positive valence, then future research ought to find ways to account for such valence differences. For example, future studies can account for US valence differences between EC conditions by (1) in analyses, adjusting for US valence, or (2) finding US that are matched on level of positive valence between EC conditions.

A second general limitation of Study 1 and Study 2 is that these studies used convenience samples of undergraduate students instead of community parents or parents screened to be high risk for CPA. Although previous research found that positive EC significantly increases student participants’ positive child evaluations and reduces negative child evaluations and future discipline expectations (Milner et al., 2017, Experiment 1 and Experiment 4), no research has
been conducted examining positive EC effects in students on anger, harsh verbal discipline, harsh physical discipline, and harm. It is unknown why positive EC effects would be moderated by whether or not an individual is a parent. In parents, early parenting experiences may cause associations to be formed in memory between children (own child or general children) and evaluative concepts (e.g., positive child evaluations, negative child evaluations, liking), trait hostility, affect, and behavioral tendencies (i.e., likelihood to use harsh verbal discipline, harsh physical discipline, or harm). Although positive EC would be expected to cause such associations to form in individuals having no prior appreciable association between these concepts, it is unknown whether positive EC has differential effects on attitude formation vs. attitude alteration. Moreover, future research is needed to explore how students’ schemas about children may differ from the schemas of parents’ varying in early parenting experiences.

Summary

A social information processing model of child physical abuse predicts that reducing negative evaluations about children should also reduce situational-based attributions of hostile intent, anger, and harsh and harmful parenting behaviors. Previous research provided evidence that use of a positive evaluative conditioning (EC) procedure is effective at reducing these situationally based variables. The studies reported in this dissertation replicated this outcome. Moreover, the results of the studies provided information about the possible mental mechanisms that prompt the effects that positive EC has on these variables.
BIBLIOGRAPHY


APPENDIX A

SAMPLE LAYOUT OF STUDY MEASURES
Please look at the child shown in this picture carefully and then answer the following question:

Do you think this child is:

- Sweet
- Friendly
- Negative
- Hostile
- Cooperative
- Difficult

What is the likelihood that this child will require discipline in the future?
Please look at the child shown in this picture carefully and then answer the following question:

Do you think this child is:

<table>
<thead>
<tr>
<th></th>
<th>Not at all</th>
<th>Somewhat</th>
<th>Extremely</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweet</td>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friendly</td>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hostile</td>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cooperative</td>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficult</td>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

What is the likelihood that this child will require discipline in the future?

1 2 3 4 5 6 7 8 9 10
Please look at the child shown in this picture carefully and then answer the following question:

Do you think this child is:

<table>
<thead>
<tr>
<th></th>
<th>Not at all</th>
<th>Somewhat</th>
<th>Extremely</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweet</td>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friendly</td>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hostile</td>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cooperative</td>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficult</td>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

What is the likelihood that this child will require discipline in the future?

1 2 3 4 5 6 7 8 9 10
Instructions: Please read each of the following descriptions of this child’s behavior and circle the number that represents the likelihood that this child’s behavior was done intentionally to annoy or to irritate. Also respond to the other items.

<table>
<thead>
<tr>
<th>A. She looked away when you spoke to her.</th>
</tr>
</thead>
<tbody>
<tr>
<td>How likely was this behavior done intentionally to irritate you?</td>
</tr>
<tr>
<td><strong>Not likely</strong></td>
</tr>
<tr>
<td>1</td>
</tr>
</tbody>
</table>

To what extent would this behavior make you feel angry?

| **Not at all angry** | **Somewhat angry** | **Very angry** |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

How likely is it that you would yell/shout/scream at the child because of this behavior?

| **Not likely** | **Somewhat likely** | **Very likely** |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

How likely is it that you would slap/hit/spank the child because of this behavior?

| **Not likely** | **Somewhat likely** | **Very likely** |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

How likely is it that you would do nothing because of this behavior?

| **Not likely** | **Somewhat likely** | **Very likely** |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
Instructions: Please view the child picture. Then read the child behavior in bold face and respond to the prompt below the child behavior.

A. She looked away when you spoke to her.

After the child performs this behavior, you might want to harm the child. Below is a picture of an outline of a child. Nearby are 23 stickers that represent pins that you can use to harm the child if you choose to do so.
APPENDIX B

AMBIGUOUS CHILD PHOTOGRAPHS
APPENDIX C

AMBIGUOUS CHILD BEHAVIORS
A. She/he looked away when you spoke to her/him.

B. She/he grimaced as you wiped her/his face.

C. She/he pushed away the cookie when you offered it.

D. She/he raised her brow as you combed her/his hair.

E. She/he lowered her/his head when you showed a toy.

F. She/he whined as you were leaving.
APPENDIX D

SAMPLE VOODOO DOLL TASK
Instructions: Please view the child picture. Then read the child behavior in bold face and respond to the prompt below the child behavior.

A. She looked away when you spoke to her.

After the child performs this behavior, you might want to harm the child. Below is a picture of an outline of a child. Nearby are 23 stickers that represent pins that you can use to harm the child if you choose to do so.
APPENDIX E

INFORMED CONSENT DOCUMENT USED IN PRELIMINARY STUDY
ATTITUDES ABOUT CHILDREN

CONSENT FORM – PARENTS - MTurk (Adults 18 years of age or older)

You are being asked to participate in a study designed to explore people's attitudes about children. Information gathered from this study may help researchers understand the extent to which people's attitudes (i.e., evaluations) about children can be changed.

If you agree to take part in this study, you will either (1) rate pictures of child faces on different traits (2) take part in a computer task in which you make judgments about neutral stimuli, or (3) do both. Subsequently, you will be shown image-word pairs. Following this, you will complete the task(s) you completed before you saw the image-word pairs. Finally, you will be asked to complete questionnaires regarding your thoughts, feelings, beliefs, and behavior about parenting. You will also be asked to provide certain demographic information (i.e., gender, age, educational level, number of children, and marital status) and your opinions about your research experience. The study will take approximately 30 minutes, and you will receive $4 for your participation.

Only the research staff will have access to your data. Your name and telephone number will be kept separate from your data. Your data will be combined with the data from the other parents participating in this study and only group data will be reported.

Your participation in this study is voluntary. While completing the questionnaires, you might feel bored, and you could feel uncomfortable answering some of the items on the questionnaires. You have the right not to answer any specific question as well as withdraw from the study at any time. A decision to withdraw will involve no penalty or loss of benefits to which you are otherwise entitled.

If you have any questions concerning this study you may contact the Principal Investigator, Michael Wagner, Department of Psychology, at 1-815-753-0747. If you have any questions or concerns regarding your rights as a study participant, please contact the Office for Research Compliance at Northern Illinois University at 1-815-753-8588.

I have read the above statements. I understand the purpose and the potential risks and benefits of the study. I have been given the chance to ask questions. I understand that I can withdraw from the study at any time for any reason by exiting out of the web browser. I understand that Northern Illinois University does not provide compensation for treatment of injuries that may occur as a result of participation in this study. I give my informed consent to be a participant in this study.

By clicking NEXT, I consent to participate in this study.
APPENDIX F

DEBRIEFING DOCUMENT USED IN PRELIMINARY STUDY
ATTITUDES ABOUT CHILDREN

DEBRIEFING - Parents

Thank you for participating in the “Attitudes about Children” study!

As you might recall, this study was designed to explore attitudes about children. For this study, you rated images of child faces on several descriptors, or you rated neutral stimuli on pleasantness. Subsequently, you viewed image-word pairs. We predicted that your ratings of the pictures of child faces or of the neutral stimuli would change as a result of the image-word pairs if the word in the pairs was positive.

You can learn more about this research by reading:

APPENDIX G

SAMPLE EMOJIS USED IN STUDY 2
Positive Emojis
Neutral Emojis
APPENDIX H

STUDY 1 DEPENDENT VARIABLE MEANS (SDS) BY EXPERIMENTAL CELL
### Study 1 Dependent Variable Means (SDs) by Experimental Cell

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Forward positive (n = 26)</th>
<th>Backward positive (n = 25)</th>
<th>Forward pseudo (n = 26)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive child evaluations</td>
<td>Pre-EC 5.00 (1.40)</td>
<td>Post-EC 5.79 (1.94)</td>
<td>Pre-EC 5.31 (1.33)</td>
</tr>
<tr>
<td>Negative child evaluations</td>
<td>Pre-EC 2.98 (1.28)</td>
<td>Post-EC 2.16 (1.64)</td>
<td>Pre-EC 3.02 (1.38)</td>
</tr>
<tr>
<td>Liking</td>
<td>Pre-EC 5.77 (1.48)</td>
<td>Post-EC 6.14 (1.81)</td>
<td>Pre-EC 6.15 (1.62)</td>
</tr>
<tr>
<td>Future discipline expectations</td>
<td>Pre-EC 4.56 (1.47)</td>
<td>Post-EC 3.87 (1.96)</td>
<td>Pre-EC 4.49 (1.91)</td>
</tr>
<tr>
<td>Attributions of hostile intent</td>
<td>Pre-EC 2.35 (1.58)</td>
<td>Post-EC 2.16 (1.86)</td>
<td>Pre-EC 2.04 (1.22)</td>
</tr>
<tr>
<td>Anger</td>
<td>Pre-EC 1.32 (1.19)</td>
<td>Post-EC 1.02 (0.94)</td>
<td>Pre-EC 1.35 (1.16)</td>
</tr>
<tr>
<td>Harsh verbal discipline</td>
<td>Pre-EC 0.86 (1.17)</td>
<td>Post-EC 0.73 (1.28)</td>
<td>Pre-EC 0.63 (0.82)</td>
</tr>
<tr>
<td>Harsh physical discipline</td>
<td>Pre-EC 0.20 (0.45)</td>
<td>Post-EC 0.18 (0.50)</td>
<td>Pre-EC 0.08 (0.27)</td>
</tr>
<tr>
<td>Doing nothing</td>
<td>Pre-EC 5.56 (2.43)</td>
<td>Post-EC 6.07 (2.51)</td>
<td>Pre-EC 3.88 (2.23)</td>
</tr>
<tr>
<td>Harm</td>
<td>Pre-EC 1.58 (5.83)</td>
<td>Post-EC 1.19 (4.93)</td>
<td>Pre-EC 4.36 (10.69)</td>
</tr>
</tbody>
</table>

*Note. N = 77.*
APPENDIX I

STUDY 1 DEPENDENT VARIABLE CHANGE SCORE MEANS ($SDS$) BY EVALUATIVE CONDITIONING (EC) CONDITION
### Study 1 Dependent Variable Change Score Means (SDs) by Evaluative Conditioning (EC) Condition

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>EC condition</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Forward positive ($n=26$)</td>
<td>Backward positive ($n=25$)</td>
<td>Forward pseudo ($n=26$)</td>
<td></td>
</tr>
<tr>
<td>Positive child evaluations</td>
<td>1.28 (1.92)</td>
<td>1.13 (1.53)</td>
<td>-0.05 (1.20)</td>
<td></td>
</tr>
<tr>
<td>Negative child evaluations</td>
<td>0.82 (1.21)</td>
<td>1.26 (1.05)</td>
<td>0.23 (0.99)</td>
<td></td>
</tr>
<tr>
<td>Liking</td>
<td>0.37 (1.34)</td>
<td>0.53 (1.02)</td>
<td>-0.03 (1.14)</td>
<td></td>
</tr>
<tr>
<td>Future discipline expectations</td>
<td>0.69 (1.32)</td>
<td>0.67 (1.38)</td>
<td>0.28 (1.06)</td>
<td></td>
</tr>
<tr>
<td>Attributions of hostile intent</td>
<td>0.19 (1.02)</td>
<td>0.51 (0.67)</td>
<td>0.23 (1.03)</td>
<td></td>
</tr>
<tr>
<td>Anger</td>
<td>0.30 (0.75)</td>
<td>0.42 (0.58)</td>
<td>0.03 (0.81)</td>
<td></td>
</tr>
<tr>
<td>Harsh verbal discipline</td>
<td>0.13 (0.51)</td>
<td>0.24 (0.42)</td>
<td>0.01 (0.60)</td>
<td></td>
</tr>
<tr>
<td>Harsh physical discipline</td>
<td>0.02 (0.29)</td>
<td>0.01 (0.04)</td>
<td>-0.06 (0.19)</td>
<td></td>
</tr>
<tr>
<td>Doing nothing</td>
<td>0.51 (1.73)</td>
<td>0.45 (1.68)</td>
<td>0.20 (0.80)</td>
<td></td>
</tr>
<tr>
<td>Harm</td>
<td>0.38 (1.13)</td>
<td>2.68 (7.53)</td>
<td>0.08 (0.89)</td>
<td></td>
</tr>
</tbody>
</table>

*Note. $N=77$.*

*For each dependent variable change scores were computed such that positive values indicate change in the hypothesized direction within the positive EC conditions. Thus, change scores for positive child evaluations, liking, and doing nothing were computed by subtracting pre-EC scores from post-EC scores, and change scores for negative child evaluations, future discipline expectations, attributions of hostile intent, anger, harsh verbal discipline, harsh physical discipline, and harm were computed by subtracting post-EC scores from pre-EC scores.*
APPENDIX J

LINEAR RELATIONSHIPS BETWEEN PRE-EC SCORES AND EC CHANGE SCORES WITHIN THE FORWARD POSITIVE EC CONDITION AND THE BACKWARD POSITIVE EC CONDITION
### Study 1 Correlations between Pre-EC scores for Each Dependent Variable and the Corresponding Change Score\(^a\)
Within the Forward Positive EC Condition and the Backward Positive EC Condition (Combined)

<table>
<thead>
<tr>
<th>Pre-EC dependent variable</th>
<th>Correlation between pre-EC score and corresponding change score</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( r )</td>
<td>( p )</td>
</tr>
<tr>
<td>Positive child evaluations</td>
<td>-0.39</td>
<td>.004</td>
</tr>
<tr>
<td>Negative child evaluations</td>
<td>0.18</td>
<td>.217</td>
</tr>
<tr>
<td>Liking</td>
<td>-0.27</td>
<td>.059</td>
</tr>
<tr>
<td>Future discipline expectations</td>
<td>0.27</td>
<td>.056</td>
</tr>
<tr>
<td>Attributions of hostile intent</td>
<td>0.05</td>
<td>.709</td>
</tr>
<tr>
<td>Anger</td>
<td>0.46</td>
<td>.001</td>
</tr>
<tr>
<td>Harsh verbal discipline</td>
<td>0.17</td>
<td>.226</td>
</tr>
<tr>
<td>Harsh physical discipline</td>
<td>0.16</td>
<td>.270</td>
</tr>
<tr>
<td>Doing nothing</td>
<td>-0.12</td>
<td>.390</td>
</tr>
<tr>
<td>Harm</td>
<td>0.78</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

*Note. \( N = 52 \).*

\(^a\)Change scores always computed in the hypothesized direction. For positive child evaluations, liking, and doing nothing, change scores were computed by subtracting post-EC scores from pre-EC scores; for negative child evaluations, future discipline expectations, attributions of hostile intent, anger, harsh verbal discipline, harsh physical discipline, and harm, change scores were computed by subtracting pre-EC scores from post-EC scores.
APPENDIX K

STUDY 1 RESULTS OF MULTI-LEVEL MODEL ANALYSES FOR POSITIVE CHILD EVALUATIONS
### Study 1 Results of Multi-level Model Analyses for Positive Child Evaluations

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1 ($df = 3$)</th>
<th>Model 2 ($df = 7$)</th>
<th>Model 3 ($df = 11$)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept (Pre-EC, Forward Pseudo-EC)</td>
<td>5.36 (.15)</td>
<td>5.35 (.16)</td>
<td>5.07 (.27)</td>
</tr>
<tr>
<td>Level 1 (trial-level, $n = 1386$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.01 (.10)</td>
<td>0.02 (.17)</td>
<td></td>
</tr>
<tr>
<td>Level 2 (person-level, $n = 77$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forward Positive EC</td>
<td>-0.04 (.25)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Backward Positive EC</td>
<td>0.03 (.25)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Random effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1.22</td>
<td>1.21</td>
<td>1.18</td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.02</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>Level 1 residual variance</td>
<td>1.90</td>
<td>1.90</td>
<td>1.90</td>
</tr>
<tr>
<td>-2*(log likelihood) for model</td>
<td>5872</td>
<td>5874</td>
<td>5871</td>
</tr>
</tbody>
</table>

*Note.* $N = 77$. Standard errors are in parentheses. Model parameters were (restricted) Maximum Likelihood estimates. EC = Evaluative Conditioning; Model 1 was the intercept-only model. Model 2 added the level 1 time predictor (coded, 0 = pre-EC, 1 = post-EC) and random effect for time. Model 3 added the level 2 EC condition predictors (coded, 0 = forward pseudo-EC, 1 = forward or backward positive EC) of post EC coefficient (i.e., cross-level interaction term). All random effects are standard deviations.
APPENDIX L

STUDY 1 RESULTS OF MULTI-LEVEL MODEL ANALYSES FOR NEGATIVE CHILD EVALUATIONS
**Study 1 Results of Multi-level Model Analyses for Negative Child Evaluations**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1 ((df = 3))</th>
<th>Model 2 ((df = 7))</th>
<th>Model 3 ((df = 11))</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept (Pre-EC, forward pseudoEC)</td>
<td>2.63 (.15)</td>
<td>2.56 (.15)</td>
<td>2.76 (.27)</td>
</tr>
<tr>
<td>Level 1 (trial-level, (n = 1386))</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.13 (.11)</td>
<td>0.30 (.18)</td>
<td></td>
</tr>
<tr>
<td>Level 2 (person-level, (n = 77))</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forward Positive EC</td>
<td></td>
<td>-0.14 (.26)</td>
<td></td>
</tr>
<tr>
<td>Backward Positive EC</td>
<td></td>
<td>-0.39 (.26)</td>
<td></td>
</tr>
<tr>
<td><strong>Random effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1.20</td>
<td>1.18</td>
<td>1.19</td>
</tr>
<tr>
<td>Post-EC</td>
<td></td>
<td>0.03</td>
<td>0.01</td>
</tr>
<tr>
<td>Level 1 residual variance</td>
<td>1.98</td>
<td>1.98</td>
<td>1.98</td>
</tr>
<tr>
<td>-2*(log likelihood) for model</td>
<td>5984</td>
<td>5986</td>
<td>5984</td>
</tr>
</tbody>
</table>

**Note.** \(N = 77\). Standard errors are in parentheses. Model parameters were (restricted) Maximum Likelihood estimates. EC = Evaluative Conditioning; Model 1 was the intercept-only model. Model 2 added the level 1 time predictor (coded, 0 = pre-EC, 1 = post-EC) and random effect for time. Model 3 added the level 2 EC condition predictors (coded, 0 = forward pseudo-EC, 1 = forward or backward positive EC) of post EC coefficient (i.e., cross-level interaction term). All random effects are standard deviations.
APPENDIX M

STUDY 1 RESULTS OF MULTI-LEVEL MODEL ANALYSES FOR LIKING
### Study 1 Results of Multi-level Model Analyses for Liking

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1 (df = 3)</th>
<th>Model 2 (df = 4)*</th>
<th>Model 3 (df = 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept (Pre-EC, forward pseudo-EC)</td>
<td>6.07 (.17)</td>
<td>6.03 (.18)</td>
<td>5.82 (.31)</td>
</tr>
<tr>
<td>Level 1 (trial-level, n = 462)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.07 (.13)</td>
<td>0.05 (.22)</td>
<td></td>
</tr>
<tr>
<td>Level 2 (person-level, n = 77)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forward Positive EC</td>
<td></td>
<td></td>
<td>0.17 (.32)</td>
</tr>
<tr>
<td>Backward Positive EC</td>
<td></td>
<td></td>
<td>-0.10 (.32)</td>
</tr>
<tr>
<td><strong>Random effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1.34</td>
<td>1.34</td>
<td>1.38</td>
</tr>
<tr>
<td>Post-EC</td>
<td>--</td>
<td></td>
<td>0.08</td>
</tr>
<tr>
<td>Level 1 residual variance</td>
<td>1.40</td>
<td>1.40</td>
<td>1.40</td>
</tr>
<tr>
<td>-2*(log likelihood) for model</td>
<td>1767</td>
<td>1769</td>
<td>1767</td>
</tr>
</tbody>
</table>

*Note. N = 77. Standard errors are in parentheses. Model parameters were (restricted) Maximum Likelihood estimates. EC = Evaluative Conditioning; Model 1 was the intercept-only model. Model 2 added the level 1 time predictor (coded, 0 = pre-EC, 1 = post-EC) and random effect for time. Model 3 added the level 2 EC condition predictors (coded, 0 = forward pseudo-EC, 1 = forward or backward positive EC) of post EC coefficient (i.e., cross-level interaction term). All random effects are standard deviations.

*Random effect for Post-EC random effect virtually zero; omitted to allow model convergence.
APPENDIX N

STUDY 1 RESULTS OF MULTI-LEVEL MODEL ANALYSES FOR FUTURE DISCIPLINE EXPECTATIONS
### Study 1 Results of Multi-level Model Analyses for Future Discipline Expectations

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1 (df = 3)</th>
<th>Model 2 (df = 7)</th>
<th>Model 3 (df = 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept (Pre-EC, forward pseudo-EC)</td>
<td>4.19 (.18)</td>
<td>4.22 (.20)</td>
<td>4.23 (.34)</td>
</tr>
<tr>
<td>Level 1 (trial-level, n = 462)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-EC</td>
<td>-0.04 (.15)</td>
<td>-0.05 (.26)</td>
<td></td>
</tr>
<tr>
<td>Level 2 (person-level, n = 77)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forward Positive EC</td>
<td></td>
<td>0.27 (.36)</td>
<td></td>
</tr>
<tr>
<td>Backward Positive EC</td>
<td></td>
<td>0.27 (.37)</td>
<td></td>
</tr>
<tr>
<td><strong>Random effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1.40</td>
<td>1.44</td>
<td>1.47</td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.08</td>
<td>0.26</td>
<td></td>
</tr>
<tr>
<td>Level 1 residual variance</td>
<td>1.60</td>
<td>1.60</td>
<td>1.60</td>
</tr>
<tr>
<td>( -2*(\text{log likelihood}) ) for model</td>
<td>1876</td>
<td>1878</td>
<td>1879</td>
</tr>
</tbody>
</table>

**Note.** \( N = 77 \). Standard errors are in parentheses. Model parameters were (restricted) Maximum Likelihood estimates. EC = Evaluative Conditioning; Model 1 was the intercept-only model. Model 2 added the level 1 time predictor (coded, 0 = pre-EC, 1 = post-EC) and random effect for time. Model 3 added the level 2 EC condition predictors (coded, 0 = forward pseudo-EC, 1 = forward or backward positive EC) of post EC coefficient (i.e., cross-level interaction term). All random effects are standard deviations.
APPENDIX O

STUDY 1 RESULTS OF MULTI-LEVEL MODEL ANALYSES FOR ATTRIBUTIONS OF HOSTILE INTENT
Study 1 Results of Multi-level Model Analyses for Attributions of Hostile Intent

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1 (df = 3)</th>
<th>Model 2 (df = 7)</th>
<th>Model 3 (df = 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept (Pre-EC, forward pseudo-EC)</td>
<td>2.08 (.16)</td>
<td>2.08 (.16)</td>
<td>2.22 (.28)</td>
</tr>
<tr>
<td>Level 1 (trial-level, n = 2772)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.01 (.07)</td>
<td>-0.08 (.11)</td>
<td></td>
</tr>
<tr>
<td>Level 2 (person-level, n = 77)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forward Positive EC</td>
<td></td>
<td></td>
<td>0.20 (.16)</td>
</tr>
<tr>
<td>Backward Positive EC</td>
<td></td>
<td></td>
<td>0.04 (.16)</td>
</tr>
<tr>
<td><strong>Random effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1.40</td>
<td>1.37</td>
<td>1.74</td>
</tr>
<tr>
<td>Post-EC</td>
<td></td>
<td>0.07</td>
<td>0.06</td>
</tr>
<tr>
<td>Level 1 residual variance</td>
<td>1.74</td>
<td>1.74</td>
<td>1.74</td>
</tr>
<tr>
<td>-2*(log likelihood) for model</td>
<td>11178</td>
<td>11181</td>
<td>11182</td>
</tr>
</tbody>
</table>

*Note. N = 77. Standard errors are in parentheses. Model parameters were (restricted) Maximum Likelihood estimates. EC = Evaluative Conditioning; Model 1 was the intercept-only model. Model 2 added the level 1 time predictor (coded, 0 = pre-EC, 1 = post-EC) and random effect for time. Model 3 added the level 2 EC condition predictors (coded, 0 = forward pseudo-EC, 1 = forward or backward positive EC) of post EC coefficient (i.e., cross-level interaction term). All random effects are standard deviations.*
APPENDIX P

STUDY 1 RESULTS OF MULTI-LEVEL MODEL ANALYSES FOR ANGER
### Study 1 Results of Multi-level Model Analyses for Anger

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1 (df = 3)</th>
<th>Model 2 (df = 7)</th>
<th>Model 3 (df = 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept (Pre-EC, forward pseudo-EC)</td>
<td>1.15 (.12)</td>
<td>1.13 (.12)</td>
<td>1.12 (.21)</td>
</tr>
<tr>
<td>Level 1 (trial-level, n = 2772)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.05 (.05)</td>
<td>0.07 (.09)</td>
<td></td>
</tr>
<tr>
<td>Level 2 (person-level, n = 77)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forward Positive EC</td>
<td></td>
<td>0.02 (.12)</td>
<td></td>
</tr>
<tr>
<td>Backward Positive EC</td>
<td></td>
<td></td>
<td>-0.06 (.12)</td>
</tr>
<tr>
<td><strong>Random effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1.05</td>
<td>1.01</td>
<td>1.03</td>
</tr>
<tr>
<td>Post-EC</td>
<td></td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>Level 1 residual variance</td>
<td>1.30</td>
<td>1.30</td>
<td>1.30</td>
</tr>
<tr>
<td>-2*(log likelihood) for model</td>
<td>9566</td>
<td>9567</td>
<td>9573</td>
</tr>
</tbody>
</table>

**Note.** N = 77. Standard errors are in parentheses. Model parameters were (restricted) Maximum Likelihood estimates. EC = Evaluative Conditioning; Model 1 was the intercept-only model. Model 2 added the level 1 time predictor (coded, 0 = pre-EC, 1 = post-EC) and random effect for time. Model 3 added the level 2 EC condition predictors (coded, 0 = forward pseudo-EC, 1 = forward or backward positive EC) of post EC coefficient (i.e., cross-level interaction term). All random effects are standard deviations.
APPENDIX Q

STUDY 1 RESULTS OF MULTI-LEVEL MODEL ANALYSES FOR HARSHP VERBAL DISCIPLINE
Study 1 Results of Multi-level Model Analyses for Harsh Verbal Discipline

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1 (df = 3)</th>
<th>Model 2 (df = 7)</th>
<th>Model 3 (df = 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept (Pre-EC, forward pseudo-EC)</td>
<td>0.74 (.12)</td>
<td>0.72 (.11)</td>
<td>0.87 (.20)</td>
</tr>
<tr>
<td>Level 1 (trial-level, n = 2772)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.02 (.04)</td>
<td>0.06 (.06)</td>
<td></td>
</tr>
<tr>
<td>Level 2 (person-level, n = 77)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forward Positive EC</td>
<td>-0.04 (.09)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Backward Positive EC</td>
<td>-0.06 (.09)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Random effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1.00</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.05</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Level 1 residual variance</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td>-2*(log likelihood) for model</td>
<td>7929</td>
<td>7938</td>
<td>7938</td>
</tr>
</tbody>
</table>

Note. N = 77. Standard errors are in parentheses. Model parameters were (restricted) Maximum Likelihood estimates. EC = Evaluative Conditioning; Model 1 was the intercept-only model. Model 2 added the level 1 time predictor (coded, 0 = pre-EC, 1 = post-EC) and random effect for time. Model 3 added the level 2 EC condition predictors (coded, 0 = forward pseudo-EC, 1 = forward or backward positive EC) of post EC coefficient (i.e., cross-level interaction term). All random effects are standard deviations.
APPENDIX R

STUDY 1 RESULTS OF MULTI-LEVEL MODEL ANALYSES FOR HARSH PHYSICAL DISCIPLINE
### Study 1 Results of Multi-level Model Analyses for Harsh Physical Discipline

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1 (df = 3)</th>
<th>Model 2 (df = 7)</th>
<th>Model 3 (df = 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept (Pre-EC, forward pseudo-EC)</td>
<td>0.21 (.07)</td>
<td>0.20 (.07)</td>
<td>0.33 (.12)</td>
</tr>
<tr>
<td>Level 1 (trial-level, n = 2772)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.00 (.02)</td>
<td>0.02 (.02)</td>
<td></td>
</tr>
<tr>
<td>Level 2 (person-level, n = 77)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forward Positive EC</td>
<td></td>
<td>-0.04 (.04)</td>
<td></td>
</tr>
<tr>
<td>Backward Positive EC</td>
<td></td>
<td></td>
<td>-0.04 (.04)</td>
</tr>
<tr>
<td><strong>Random effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.64</td>
<td>0.62</td>
<td>0.62</td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.04</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>Level 1 residual variance</td>
<td>0.38</td>
<td>0.38</td>
<td>0.38</td>
</tr>
<tr>
<td>-2*(log likelihood) for model</td>
<td>2890</td>
<td>2890</td>
<td>2900</td>
</tr>
</tbody>
</table>

**Note.** N = 77. Standard errors are in parentheses. Model parameters were (restricted) Maximum Likelihood estimates. EC = Evaluative Conditioning; Model 1 was the intercept-only model. Model 2 added the level 1 time predictor (coded, 0 = pre-EC, 1 = post-EC) and random effect for time. Model 3 added the level 2 EC condition predictors (coded, 0 = forward pseudo-EC, 1 = forward or backward positive EC) of post EC coefficient (i.e., cross-level interaction term). All random effects are standard deviations.
APPENDIX S

STUDY 1 RESULTS OF MULTI-LEVEL MODEL ANALYSES FOR DOING NOTHING
### Study 1 Results of Multi-level Model Analyses for Doing Nothing

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1 (df = 3)</th>
<th>Model 2 (df = 7)</th>
<th>Model 3 (df = 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept (Pre-EC, forward pseudo-EC)</td>
<td>5.50 (.28)</td>
<td>5.49 (.28)</td>
<td>6.56 (.44)</td>
</tr>
<tr>
<td>Level 1 (trial-level, n = 2772)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.03 (.08)</td>
<td>-0.02 (.14)</td>
<td></td>
</tr>
<tr>
<td><strong>Level 2 (person-level, n = 77)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forward Positive EC</td>
<td>0.12 (.19)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Backward Positive EC</td>
<td>0.02 (.20)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Random effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>2.40</td>
<td>2.37</td>
<td>2.17</td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.05</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>Level 1 residual variance</td>
<td>2.10</td>
<td>2.10</td>
<td>2.10</td>
</tr>
<tr>
<td>-2*(log likelihood) for model</td>
<td>12270</td>
<td>12273</td>
<td>12259</td>
</tr>
</tbody>
</table>

*Note. N = 77. Standard errors are in parentheses. Model parameters were (restricted) Maximum Likelihood estimates. EC = Evaluative Conditioning; Model 1 was the intercept-only model. Model 2 added the level 1 time predictor (coded, 0 = pre-EC, 1 = post-EC) and random effect for time. Model 3 added the level 2 EC condition predictors (coded, 0 = forward pseudo-EC, 1 = forward or backward positive EC) of post EC coefficient (i.e., cross-level interaction term). All random effects are standard deviations.*
APPENDIX T

STUDY 2 DEPENDENT VARIABLE MEANS (SDS) BY EXPERIMENTAL CELL
Study 2 Dependent Variable Means (SDs) by Experimental Cell

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Positive ( n = 27 )</th>
<th>Pseudo ( n = 23 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-EC</td>
<td>Post-EC</td>
</tr>
<tr>
<td>Positive child evaluations</td>
<td>4.79 (0.81)</td>
<td>5.30 (1.37)</td>
</tr>
<tr>
<td>Negative child evaluations</td>
<td>2.93 (1.34)</td>
<td>2.22 (1.43)</td>
</tr>
<tr>
<td>Liking</td>
<td>5.99 (1.38)</td>
<td>6.17 (1.46)</td>
</tr>
<tr>
<td>Future discipline expectations</td>
<td>4.30 (1.67)</td>
<td>3.60 (2.00)</td>
</tr>
<tr>
<td>Attributions of hostile intent</td>
<td>1.74 (1.62)</td>
<td>1.44 (1.46)</td>
</tr>
<tr>
<td>Anger</td>
<td>1.24 (0.98)</td>
<td>0.98 (1.06)</td>
</tr>
<tr>
<td>Harsh verbal discipline</td>
<td>0.80 (0.87)</td>
<td>0.63 (0.75)</td>
</tr>
<tr>
<td>Harsh physical discipline</td>
<td>0.06 (0.15)</td>
<td>0.07 (0.27)</td>
</tr>
<tr>
<td>Doing nothing</td>
<td>5.80 (2.54)</td>
<td>6.26 (2.74)</td>
</tr>
<tr>
<td>Harm</td>
<td>0.70 (2.11)</td>
<td>0.70 (2.15)</td>
</tr>
</tbody>
</table>

*Note. N = 50.*
APPENDIX U

STUDY 2 DEPENDENT VARIABLE CHANGE SCORE MEANS ($SDS$) BY EVALUATIVE CONDITIONING (EC) CONDITION
Study 2 Dependent Variable Change Score\textsuperscript{a} Means (SDs) by Experimental Cell

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Positive ((n = 27))</th>
<th>Pseudo ((n = 23))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive child evaluations</td>
<td>0.51 (1.20)</td>
<td>-0.16 (0.80)</td>
</tr>
<tr>
<td>Negative child evaluations</td>
<td>0.72 (1.08)</td>
<td>-0.06 (1.28)</td>
</tr>
<tr>
<td>Liking</td>
<td>0.19 (1.00)</td>
<td>-0.61 (0.94)</td>
</tr>
<tr>
<td>Future discipline expectations</td>
<td>0.69 (1.20)</td>
<td>0.52 (1.49)</td>
</tr>
<tr>
<td>Attributions of hostile intent</td>
<td>0.30 (1.00)</td>
<td>-0.11 (0.72)</td>
</tr>
<tr>
<td>Anger</td>
<td>0.27 (0.60)</td>
<td>0.02 (0.52)</td>
</tr>
<tr>
<td>Harsh verbal discipline</td>
<td>0.17 (0.34)</td>
<td>0.04 (0.31)</td>
</tr>
<tr>
<td>Harsh physical discipline</td>
<td>-0.0 (0.16)</td>
<td>-0.03 (0.12)</td>
</tr>
<tr>
<td>Doing nothing</td>
<td>0.46 (0.70)</td>
<td>-0.09 (0.96)</td>
</tr>
<tr>
<td>Harm</td>
<td>0.00 (0.28)</td>
<td>-0.13 (1.42)</td>
</tr>
</tbody>
</table>

\textit{Note.} \(N = 50\).

\textsuperscript{a}For each dependent variable change scores were computed such that positive values indicate change in the hypothesized direction within the positive EC condition. Thus, change scores for positive child evaluations, liking, and doing nothing were computed by subtracting pre-EC scores from post-EC scores, and change scores for negative child evaluations, future discipline expectations, attributions of hostile intent, anger, harsh verbal discipline, harsh physical discipline, and harm were computed by subtracting post-EC scores from pre-EC scores.
APPENDIX V

LINEAR RELATIONSHIPS BETWEEN PRE-EC SCORES AND EC CHANGE SCORES WITHIN THE POSITIVE EC CONDITION
### Study 2 Correlations between Pre-EC scores for Each Dependent Variable and the Corresponding Change Score Within the Positive EC Condition

<table>
<thead>
<tr>
<th>Pre-EC dependent variable</th>
<th>Correlation between pre-EC score and corresponding change score</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r</td>
<td>p</td>
</tr>
<tr>
<td>Positive child evaluations</td>
<td>-0.10</td>
<td>.285</td>
</tr>
<tr>
<td>Negative child evaluations</td>
<td>0.31</td>
<td>.118</td>
</tr>
<tr>
<td>Liking</td>
<td>-0.28</td>
<td>.156</td>
</tr>
<tr>
<td>Future discipline expectations</td>
<td>0.06</td>
<td>.758</td>
</tr>
<tr>
<td>Attributions of hostile intent</td>
<td>0.46</td>
<td>.015</td>
</tr>
<tr>
<td>Anger</td>
<td>0.17</td>
<td>.397</td>
</tr>
<tr>
<td>Harsh verbal discipline</td>
<td>0.54</td>
<td>.004</td>
</tr>
<tr>
<td>Harsh physical discipline</td>
<td>-0.47</td>
<td>.014</td>
</tr>
<tr>
<td>Doing nothing</td>
<td>0.17</td>
<td>.386</td>
</tr>
<tr>
<td>Harm</td>
<td>-0.06</td>
<td>.744</td>
</tr>
</tbody>
</table>

*Note. N = 27.*

*a Change scores always computed in the hypothesized direction. For positive child evaluations, liking, and doing nothing, change scores were computed by subtracting post-EC scores from pre-EC scores; for negative child evaluations, future discipline expectations, attributions of hostile intent, anger, harsh verbal discipline, harsh physical discipline, and harm, change scores were computed by subtracting pre-EC scores from post-EC scores.*
APPENDIX W

STUDY 2 RESULTS OF MULTI-LEVEL MODEL ANALYSES FOR POSITIVE CHILD EVALUATIONS
### Study 2 Results of Multi-level Model Analyses for Positive Child Evaluations

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1 (df = 3)</th>
<th>Model 2 (df = 7)</th>
<th>Model 3 (df = 6)*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept (Pre-EC, Forward Pseudo-EC)</td>
<td>4.86 (.17)</td>
<td>4.79 (.17)</td>
<td>4.64 (.26)</td>
</tr>
<tr>
<td>Level 1 (trial-level, n = 900)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.14 (.11)</td>
<td>0.02 (.16)</td>
<td></td>
</tr>
<tr>
<td>Level 2 (person-level, n = 50)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive EC</td>
<td></td>
<td></td>
<td>0.21 (.22)</td>
</tr>
<tr>
<td><strong>Random effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1.15</td>
<td>0.80</td>
<td>1.14</td>
</tr>
<tr>
<td>Post-EC</td>
<td></td>
<td>0.26</td>
<td>--</td>
</tr>
<tr>
<td>Level 1 residual variance</td>
<td>1.61</td>
<td>1.60</td>
<td>1.61</td>
</tr>
<tr>
<td>-2*(log likelihood) for model</td>
<td>3528</td>
<td>3527</td>
<td>3528</td>
</tr>
</tbody>
</table>

Note. *N = 50. Standard errors are in parentheses. Model parameters were (restricted) Maximum Likelihood estimates. EC = Evaluative Conditioning; Model 1 was the intercept-only model. Model 2 added the level 1 time predictor (coded, 0 = pre-EC, 1 = post-EC) and random effect for time. Model 3 added the level 2 EC condition predictors (coded, 0 = pseudo-EC, 1 = positive EC) of post EC coefficient (i.e., cross-level interaction term). All random effects are standard deviations.

*In Model 3 time random effect prevented model convergence.
APPENDIX X

STUDY 2 RESULTS OF MULTI-LEVEL MODEL ANALYSES FOR NEGATIVE CHILD EVALUATIONS
### Study 2 Results of Multi-level Model Analyses for Negative Child Evaluations

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1 (df = 3)</th>
<th>Model 2 (df = 7)</th>
<th>Model 3 (df = 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept (Pre-EC, Forward Pseudo-EC)</td>
<td>2.66 (.21)</td>
<td>2.62 (.21)</td>
<td>2.65 (.31)</td>
</tr>
<tr>
<td>Level 1 (trial-level, n = 900)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.08 (.11)</td>
<td>0.25 (.17)</td>
<td></td>
</tr>
<tr>
<td>Level 2 (person-level, n = 50)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive EC</td>
<td></td>
<td></td>
<td>0.30 (.23)</td>
</tr>
<tr>
<td><strong>Random effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1.41</td>
<td>1.38</td>
<td>1.40</td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.06</td>
<td>0.06</td>
<td>0.05</td>
</tr>
<tr>
<td>Level 1 residual variance</td>
<td>1.71</td>
<td>1.70</td>
<td>1.70</td>
</tr>
<tr>
<td>-2*(log likelihood) for model</td>
<td>3645</td>
<td>3646</td>
<td>3646</td>
</tr>
</tbody>
</table>

*Note. N = 50. Standard errors are in parentheses. Model parameters were (restricted) Maximum Likelihood estimates. EC = Evaluative Conditioning; Model 1 was the intercept-only model. Model 2 added the level 1 time predictor (coded, 0 = pre-EC, 1 = post-EC) and random effect for time. Model 3 added the level 2 EC condition predictors (coded, 0 = pseudo-EC, 1 = positive EC) of post EC coefficient (i.e., cross-level interaction term). All random effects are standard deviations.*
APPENDIX Y

STUDY 2 RESULTS OF MULTI-LEVEL MODEL ANALYSES FOR LIKING
### Study 2 Results of Multi-level Model Analyses for Liking

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1 ((df = 3))</th>
<th>Model 2 ((df = 4))(^{a})</th>
<th>Model 3(^{b})</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept (Pre-EC, Forward Pseudo-EC)</td>
<td>6.03 (.18)</td>
<td>5.90 (.25)</td>
<td>--</td>
</tr>
<tr>
<td>Level 1 (trial-level, (n = 900))</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.13 (.16)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Level 2 (person-level, (n = 50))</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive EC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Random effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1.59</td>
<td>1.58</td>
<td>--</td>
</tr>
<tr>
<td>Post-EC</td>
<td></td>
<td></td>
<td>--</td>
</tr>
<tr>
<td>Level 1 residual variance</td>
<td>1.35</td>
<td>1.35</td>
<td>--</td>
</tr>
<tr>
<td>(-2*(log likelihood) for model)</td>
<td>1144</td>
<td>1145</td>
<td>--</td>
</tr>
</tbody>
</table>

**Note.** \(N = 50\). Standard errors are in parentheses. Model parameters were (restricted) Maximum Likelihood estimates. EC = Evaluative Conditioning; Model 1 was the intercept-only model. Model 2 added the level 1 time predictor (coded, 0 = pre-EC, 1 = post-EC) and random effect for time. Model 3 added the level 2 EC condition predictors (coded, 0 = pseudo-EC, 1 = positive EC) of post EC coefficient (i.e., cross-level interaction term). All random effects are standard deviations.

\(^{a}\)In Model 2 time random effect prevented model convergence. \(^{b}\)Model 3 did not converge with time random effect removed.
APPENDIX Z

STUDY 2 RESULTS OF MULTI-LEVEL MODEL ANALYSES FOR FUTURE DISCIPLINE EXPECTATIONS
### Study 2 Results of Multi-level Model Analyses for Future Discipline Expectations

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1 ($df = 3$)</th>
<th>Model 2 ($df = 4$)</th>
<th>Model 3 ($df = 6$)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed effects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept (Pre-EC, Forward Pseudo-EC)</td>
<td>$4.22 (.28)$</td>
<td>$4.24 (.30)$</td>
<td>$4.56 (.43)$</td>
</tr>
<tr>
<td>Level 1 (trial-level, $n = 900$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-EC</td>
<td>$-0.04 (.18)$</td>
<td>$-0.04 (.27)$</td>
<td></td>
</tr>
<tr>
<td>Level 2 (person-level, $n = 50$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive EC</td>
<td>$0.00 (.37)$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random effects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>$1.89$</td>
<td>$1.89$</td>
<td>$1.89$</td>
</tr>
<tr>
<td>Post-EC</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Level 1 residual variance</td>
<td>$1.50$</td>
<td>$1.50$</td>
<td>$1.50$</td>
</tr>
<tr>
<td>$-2^\ast$ (log likelihood) for model</td>
<td>$1212$</td>
<td>$1213$</td>
<td>$1212$</td>
</tr>
</tbody>
</table>

*Note. $N = 50$. Standard errors are in parentheses. Model parameters were (restricted) Maximum Likelihood estimates. EC = Evaluative Conditioning; Model 1 was the intercept-only model. Model 2 added the level 1 time predictor (coded, 0 = pre-EC, 1 = post-EC) and random effect for time. Model 3 added the level 2 EC condition predictors (coded, 0 = pseudo-EC, 1 = positive EC) of post EC coefficient (i.e., cross-level interaction term). All random effects are standard deviations.

*In Model 2 and Model 3 time random effect prevented model convergence.
APPENDIX AA

STUDY 2 RESULTS OF MULTI-LEVEL MODEL ANALYSES FOR ATTRIBUTIONS OF HOSTILE INTENT
### Study 2 Results of Multi-level Model Analyses for Attributions of Hostile Intent

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1 (df = 3)</th>
<th>Model 2 (df = 7)</th>
<th>Model 3 (df = 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fixed effects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept (Pre-EC, Forward Pseudo-EC)</td>
<td>1.56 (.18)</td>
<td>1.56 (.18)</td>
<td>1.51 (.26)</td>
</tr>
<tr>
<td>Level 1 (trial-level, n = 1800)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.01 (.07)</td>
<td>0.06 (.11)</td>
<td></td>
</tr>
<tr>
<td>Level 2 (person-level, n = 50)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive EC</td>
<td>-0.08 (.15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Random effects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1.26</td>
<td>1.20</td>
<td>1.21</td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
</tr>
<tr>
<td>Level 1 residual variance</td>
<td>1.55</td>
<td>1.54</td>
<td>1.54</td>
</tr>
<tr>
<td>-2*(log likelihood) for model</td>
<td>6836</td>
<td>6837</td>
<td>6839</td>
</tr>
</tbody>
</table>

**Note.** N = 50. Standard errors are in parentheses. Model parameters were (restricted) Maximum Likelihood estimates. EC = Evaluative Conditioning; Model 1 was the intercept-only model. Model 2 added the level 1 time predictor (coded, 0 = pre-EC, 1 = post-EC) and random effect for time. Model 3 added the level 2 EC condition predictors (coded, 0 = pseudo-EC, 1 = positive EC) of post EC coefficient (i.e., cross-level interaction term). All random effects are standard deviations.
APPENDIX BB

STUDY 2 RESULTS OF MULTI-LEVEL MODEL ANALYSES FOR ANGER
### Study 2 Results of Multi-level Model Analyses for Anger

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1 (df = 3)</th>
<th>Model 2 (df = 7)</th>
<th>Model 3 (df = 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept (Pre-EC, Forward Pseudo-EC)</td>
<td>1.04 (.13)</td>
<td>1.04 (.13)</td>
<td>0.95 (.18)</td>
</tr>
<tr>
<td>Level 1 (trial-level, n = 1800)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.01 (.07)</td>
<td>0.02 (.10)</td>
<td></td>
</tr>
<tr>
<td>Level 2 (person-level, n = 50)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive EC</td>
<td>-0.02 (.14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Random effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.92</td>
<td>0.84</td>
<td>0.84</td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.16</td>
<td>0.16</td>
<td>0.16</td>
</tr>
<tr>
<td>Level 1 residual variance</td>
<td>1.38</td>
<td>1.38</td>
<td>1.38</td>
</tr>
<tr>
<td>-2*(log likelihood) for model</td>
<td>6405</td>
<td>3403</td>
<td>3406</td>
</tr>
</tbody>
</table>

*Note. N = 50. Standard errors are in parentheses. Model parameters were (restricted) Maximum Likelihood estimates. EC = Evaluative Conditioning; Model 1 was the intercept-only model. Model 2 added the level 1 time predictor (coded, 0 = pre-EC, 1 = post-EC) and random effect for time. Model 3 added the level 2 EC condition predictors (coded, 0 = pseudo-EC, 1 = positive EC) of post EC coefficient (i.e., cross-level interaction term). All random effects are standard deviations.*
APPENDIX CC

STUDY 2 RESULTS OF MULTI-LEVEL MODEL ANALYSES FOR HARSH VERBAL DISCIPLINE
### Study 2 Results of Multi-level Model Analyses for Harsh Verbal Discipline

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1 (df = 3)</th>
<th>Model 2 (df = 7)</th>
<th>Model 3 (df = 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept (Pre-EC, Forward Pseudo-EC)</td>
<td>0.62 (.11)</td>
<td>0.62 (.11)</td>
<td>0.50 (.16)</td>
</tr>
<tr>
<td>Level 1 (trial-level, n = 1800)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.00 (.06)</td>
<td>0.03 (.08)</td>
<td></td>
</tr>
<tr>
<td>Level 2 (person-level, n = 50)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive EC</td>
<td>-0.05 (.11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Random effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.78</td>
<td>0.72</td>
<td>0.72</td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.12</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>Level 1 residual variance</td>
<td>1.11</td>
<td>1.11</td>
<td>1.11</td>
</tr>
<tr>
<td>-2*(log likelihood) for model</td>
<td>5633</td>
<td>5632</td>
<td>5634</td>
</tr>
</tbody>
</table>

*Note. N = 50. Standard errors are in parentheses. Model parameters were (restricted) Maximum Likelihood estimates. EC = Evaluative Conditioning; Model 1 was the intercept-only model. Model 2 added the level 1 time predictor (coded, 0 = pre-EC, 1 = post-EC) and random effect for time. Model 3 added the level 2 EC condition predictors (coded, 0 = pseudo-EC, 1 = positive EC) of post EC coefficient (i.e., cross-level interaction term). All random effects are standard deviations.*
APPENDIX DD

STUDY 2 RESULTS OF MULTI-LEVEL MODEL ANALYSES FOR HARSH PHYSICAL DISCIPLINE
### Study 2 Results of Multi-level Model Analyses for Harsh Physical Discipline

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1 ($df = 3$)</th>
<th>Model 2 ($df = 7$)</th>
<th>Model 3 ($df = 6$)*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept (Pre-EC, Forward Pseudo-EC)</td>
<td>0.04 (.02)</td>
<td>0.03 (.02)</td>
<td>0.01 (.04)</td>
</tr>
<tr>
<td>Level 1 (trial-level, $n = 1800$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.02 (.03)</td>
<td>0.01 (.03)</td>
<td></td>
</tr>
<tr>
<td>Level 2 (person-level, $n = 50$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive EC</td>
<td>0.03 (.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Random effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.14</td>
<td>0.00</td>
<td>0.14</td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.16</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Level 1 residual variance</td>
<td>0.39</td>
<td>0.38</td>
<td>0.39</td>
</tr>
<tr>
<td>-2*(log likelihood) for model</td>
<td>1794</td>
<td>1767</td>
<td>1805</td>
</tr>
</tbody>
</table>

Note. $N =$ 50. Standard errors are in parentheses. Model parameters were (restricted) Maximum Likelihood estimates. EC = Evaluative Conditioning; Model 1 was the intercept-only model. Model 2 added the level 1 time predictor (coded, 0 = pre-EC, 1 = post-EC) and random effect for time. Model 3 added the level 2 EC condition predictors (coded, 0 = pseudo-EC, 1 = positive EC) of post EC coefficient (i.e., cross-level interaction term). All random effects are standard deviations.

*In Model 3 time random effect prevented model convergence.
APPENDIX EE

STUDY 2 RESULTS OF MULTI-LEVEL MODEL ANALYSES FOR DOING NOTHING
Study 2 Results of Multi-level Model Analyses for Doing Nothing

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1 (df = 3)</th>
<th>Model 2 (df = 7)</th>
<th>Model 3 (df = 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept (Pre-EC, Forward Pseudo-EC)</td>
<td>6.20 (.32)</td>
<td>6.16 (.34)</td>
<td>6.28 (.50)</td>
</tr>
<tr>
<td>Level 1 (trial-level, n = 1800)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-EC</td>
<td>.07 (.11)</td>
<td>0.27 (.15)</td>
<td></td>
</tr>
<tr>
<td>Level 2 (person-level, n = 50)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive EC</td>
<td></td>
<td></td>
<td>-0.38 (.21)</td>
</tr>
<tr>
<td><strong>Random effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>2.24</td>
<td>1.46</td>
<td>1.62</td>
</tr>
<tr>
<td>Post-EC</td>
<td>0.42</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>Level 1 residual variance</td>
<td>1.90</td>
<td>1.88</td>
<td>1.88</td>
</tr>
<tr>
<td>-2*(log likelihood) for model</td>
<td>7609</td>
<td>7604</td>
<td>7600</td>
</tr>
</tbody>
</table>

Note. N = 50. Standard errors are in parentheses. Model parameters were (restricted) Maximum Likelihood estimates. EC = Evaluative Conditioning; Model 1 was the intercept-only model. Model 2 added the level 1 time predictor (coded, 0 = pre-EC, 1 = post-EC) and random effect for time. Model 3 added the level 2 EC condition predictors (coded, 0 = pseudo-EC, 1 = positive EC) of post EC coefficient (i.e., cross-level interaction term). All random effects are standard deviations.