Implicit attitudes toward children may be unrelated to child abuse risk

Heather J. Risser a,*, John J. Skowronski b, Julie L. Crouch b

a University of Illinois at Chicago, USA
b Northern Illinois University, USA

Abstract

Objective: To explore whether adults possess implicit attitudes toward children and whether those attitudes are especially negative among respondents who are high in child physical abuse (CPA) risk.

Methods: The present study used an implicit evaluative priming procedure. In this procedure, participants were instructed to make decisions about the evaluative implications of target words. These words were preceded by photographs of child faces or adult faces displaying positive, neutral, or negative expressions. Reaction times for the evaluative decisions were used as an index of the extent to which photos invoked negative or positive evaluative reactions.

Results: Results from 2 studies, the first conducted on a student sample (N=90) and the second on a parent sample (N=95), demonstrated that evaluative congruence between the facial expressions displayed in photographs and the target words facilitated responses. Furthermore, the results suggested that regardless of CPA risk, child faces, relative to adult faces, facilitated responses to negative target words, suggesting an out-group bias. This implicit out-group bias was not moderated by respondents’ CPA risk status.

Conclusions: Faces of children, relative to faces of adults, appear to activate negative information structures that facilitate evaluative decisions of negative stimuli, suggesting an out-group bias. Given that out-group biases typically lead to less favorable treatment of out-group members, additional research is needed to examine the pervasiveness of negative evaluative biases toward children and the potential implications of such biases on children’s lives. Further, research examining whether high CPA risk parents and low CPA risk parents differ in how they manage initial negative evaluative reactions is needed.

© 2011 Elsevier Ltd. All rights reserved.

The Social Information Processing (SIP) model of physically abusive parenting (Crouch & Milner, 2005; Milner, 1993, 2000, 2003) describes cognitive processes that direct parenting behavior. In this model, repeated experiences with parenting and with child stimuli contribute to the development of parenting knowledge structures called schemata. Schemata include parent–child interaction scripts as well as knowledge about child attributes or characteristics (e.g., Risser, Lovejoy, & Magilano, 2005). Importantly, some schemata are highly accessible, and hence, are more easily activated than others (Bargh & Williams, 2006). These easily activated schemata are especially likely to guide subsequent social information processing, influencing everything from the stimuli to which a parent attends to the behaviors selected in parent–child interactions.

* This research was partially supported by an NIMH training grant (MH 019952-08) awarded to Joel S. Milner and was originally conducted as the first author’s doctoral dissertation under the direction of John J. Skowronski and Joel S. Milner.

* Corresponding author address: Interdisciplinary Center for Research on Violence, University of Illinois-Chicago MC 141, 1007 W. Harrison St., Chicago, IL 60607-7137, USA.

0145-2134/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.chiabu.2011.02.008
Schemata can contain both semantic knowledge and evaluative knowledge. Semantic knowledge includes beliefs about developmental milestones (e.g., that by 3 years of age children should be toilet trained) and general expectations about child behavior and characteristics (e.g., children are noisy; see Holden & Miller, 1999; Stolz, 1967). Evaluative knowledge includes attitudes about children, a particular child, or a child behavior (e.g., a 3-year-old who is not toilet trained is bad; quiet children are good). Importantly, each knowledge type may be linked to a specific mechanism underlying abusive parenting behavior; these may guide subsequent CPA risk reduction interventions. For example, deficits in semantic knowledge may indicate a need for parent training leading to the acquisition of appropriate developmental expectations. In comparison, the presence of parent negative evaluative biases may point to the usefulness of mindfulness training, which might reduce the influence of negative biases on parent thoughts and behaviors.

Research has just begun to explore the cognitive structures and social information processing tendencies that might characterize parents who are high in child physical abuse (CPA) risk. For example, recent results suggest that parents high in CPA risk possess highly accessible knowledge structures that link children with negative traits. One set of studies (Farc, Crouch, Milner, & Skowronski, 2008) found that high CPA risk parents were especially likely to rate children displaying ambiguous facial expressions as hostile and difficult. However, one ambiguity in the Farc et al. (2008) results reflects the classic distinction that attitude theorists have made between beliefs about an attitude object and evaluative reactions to the attitude object (e.g., Breckler & Wiggins, 1989). From the standpoint of this distinction, it is not clear whether Farc et al.'s results were due to beliefs about children (the semantic component of attitudes) or to evaluative reactions to children. One possibility is that high CPA risk parents are especially likely to have developed schemata containing a default belief that children are hostile and difficult. If this were the case, then this belief could be invoked to assign a negative interpretation to ambiguous child behavior (e.g., a neutral facial expression). A second possibility is that high CPA risk parents are especially likely to negatively evaluate children whose faces convey neutral expressions. Note that this response does not necessarily involve a judgment about, or interpretation of, the child's behavior. Instead, the response simply reflects an evaluative reaction to a stimulus (e.g., a negative reaction to the neutral-expression child face).

The Farc et al. (2008) results could have been caused by CPA risk group differences in either of these components of child-relevant schemata. The primary goal of the research described in the present article is to gain information about which of these components may have been responsible for the Farc et al. (2008) results. The research does so by probing for evidence of automatic negative evaluative responses to neutral-expression child faces.

Such automatic activation of evaluative components of knowledge structures is often cited as a property of attitudes. For example, Fazio, Sanbonmatsu, Powell, and Kardes (1986) suggested that when a strong association exists between an attitude object and an evaluation, that association can be automatically activated by simply exposing the individual to the attitude object. Subsequent research has shown that this activation may occur even when the association is relatively weak (Bargh, Chalken, Raymond, & Hymes, 1996).

The evaluative priming technique relies on this effect (see Fazio & Olson, 2003). The technique attempts to assess, in a subtle and non-reactive manner, the association between an attitude object and an evaluation. The technique is predicated on the idea that responses to a question about the valence of a target stimulus can be made especially quickly when there is a match in valence between a priming stimulus and the target (Arkes & Tetlock, 2004). For example, in one set of studies (Fazio, Jackson, Dunton, & Williams, 1995), on each of a series of trials, a priming photo (either a White face or a Black face) was followed by either a positive adjective (e.g., pleasant) or a negative adjective (e.g., awful). On each trial, participants made a decision about adjective valence. Fazio et al. (1995) found in-group biases. White students' responses to positive adjectives were facilitated when preceded by White faces, and responses to negative adjectives were facilitated when preceded by Black faces. In contrast, Black students' responses to positive adjectives showed facilitation when preceded by Black faces, and responses to negative adjectives showed facilitation when preceded by White faces.

Importantly, the evaluative priming method is thought to capture evaluative effects (reflecting a match in attitude valence; e.g., good, bad), but not semantic effects (reflecting knowledge of, or beliefs about, an attitude object). Accordingly, the authors reasoned that this technique could help to determine whether Farc et al.'s (2008) data reflected negative evaluative responses to children whose facial expressions were neutral, or whether the data reflected semantic beliefs that prompted negative interpretations of the neutral facial expressions exhibited by the children.

To better understand this reasoning, consider a thought experiment. Imagine that a photograph of a child displaying a neutral expression is followed by a negative word. If this photo prompts a negative evaluative response, then responses to the question of whether the word is positive or negative should be speeded (a priming effect). The emergence of such a priming effect in people high in CPA risk, but not in people low in CPA risk, would suggest that in the Farc et al. (2008) study parents high in CPA risk rated the neutral-faced child negatively because that neutral child face prompted a negative evaluative response. Now consider an alternative result: for all parents, regardless of CPA risk, no priming effect emerges when a neutral-faced child photo is followed by a negative word. The absence of a priming effect on these trials in all parents, regardless of their risk status, would not support an evaluative explanation of Farc et al.'s (2008) results. Instead, such a result would suggest that high CPA risk parents were especially likely to rate the neutral-faced child as hostile and difficult because the lack of expression on the face was interpreted as the child acting in a hostile manner.

The present article reports results from two studies designed to explore these ideas. The evaluative priming task used in these studies: (1) presented photos of children whose facial expressions were neutral, (2) followed those faces with negative words, and (3) observed whether the faces speeded responses to the question of whether the words were negative or positive. Our primary interests were: (1) whether response facilitation (revealed as shorter response latencies) emerged...
on these trials, and (2) whether the amount of facilitation differed for people who differed in CPA risk. Except for the samples used, the two experiments described in the present article were virtually identical. The first experiment used a college student sample. The second experiment used a parents sample. These two experiments allow us to assess the cognitive processes of potential parents and to compare them to the cognitive processes of actual parents. Such comparisons can yield insight into the extent to which child rearing alters child-related cognition.

One possible result of these studies would be that there would not be a CPA risk group difference in the extent to which neutral-valenced child photos speeded responses to negative words. We also realized that should such a null effect occur, we needed evidence that this null effect was not a consequence of flawed procedures. We reasoned that if the evaluative priming task was working correctly, we should find the usual evaluative priming effects when the expression depicted in the photo explicitly matched the valence of the subsequently presented word. Accordingly, we included trials of child photos that also depicted positively valenced and negatively valenced facial expressions. We also included positive words. Should priming effects occur in response to such stimuli, it would be hard to attribute to poor methods any null priming result obtained for the neutral photos.

However, responses on these valenced-face trials also have theoretical implications. For example, it may be that a CPA risk group difference emerges in response to the words following the neutral photos, but that the valenced facial expressions eliminate CPA risk group differences in evaluative priming. If this were the case, then both groups would show equal facilitation when the valence of the facial expression matched the valence of the target word. Such results would point to the idea that at-risk parents do not evaluate children negatively—only neutral-expressioned children. However, there is another possibility. Although intuitively unlikely, high CPA risk parents may be especially likely to harbor negative attitudes toward children, or low CPA risk people may be especially likely to harbor positive attitudes towards children, regardless of facial expression. Should this be the case, then it is possible that a CPA risk group difference might emerge, even on those trials on which the child photos depict valenced facial expressions.

Finally, we wondered whether the CPA risk group effects that might emerge were specific to schemata containing information about children, or whether they reflected schemata containing information about people in general. To explore this idea, we also included photos of adults in our evaluative priming task. On the one hand, CPA risk group effects could be limited to photos of child faces. For example, it could be that child faces displaying neutral expressions prime responses to negative trait words, but not adult faces displaying such expressions. Such a result would suggest that high CPA risk and low CPA risk parents might differ in the schemata that drive their responses toward children, but not in the schemata that drive their responses to adults. On the other hand, it is possible that the CPA risk assessment measure that we used (the Child Abuse Potential [CAP] Inventory; Milner, 1986) reflects general assumptions and beliefs about others. If this were the case, then one might expect similar priming effects to emerge in response to both the child photos and the adult photos.

**Experiment 1**

**Method**

**Participants.** Participants were 132 undergraduates enrolled in an introductory psychology course at a large midwestern university. All participants received credit toward completion of a course research option. Because they exhibited evidence of response distortion on the CAP inventory (described below), the CPA risk status of some participants (n = 42) could not reliably be determined. Exclusion of these participants resulted in a final sample of 90 respondents. The mean age of participants in this the final sample was 18.9 years (SD = 1.3). Most were female (60%). The sample reflected diversity in ethnic heritage (Caucasian, 61.2%; African-American, 16.7%; Latino, 12.2%; Asian-American, 3.3%, and Other, 6.6%).

Following the guidelines outlined in the CAP scoring manual (Milner, 1986), participants who scored below 91 on the CAP were assigned to the low-risk group (n = 30, M = 47.0, SD = 22.4). Participants who scored between 91 and 165 on the CAP were assigned to the moderate-risk group (n = 28, M = 119.3, SD = 22.7). Participants who scored 166 and above on the CAP were assigned to the high-risk group (n = 32, M = 259.5, SD = 63.3). [All analyses exploring the parental risk group variable were conducted using this CAP categorization scheme, which allows the data to be analyzed in an ANOVA format. However, we recognize that some authors (e.g., Cohen & Cohen, 1983) recommend against this procedure. They would suggest that the original score on the CAP scale be used as a predictor in regression analyses. We did the analyses both ways. The inferential conclusions derived from the analyses did not differ. We present the results from the ANOVA because they are simply described.]

**Materials**

**Child Abuse Potential (CAP) Inventory.** The CAP Inventory (Milner, 1986) is a 160-item, self-report questionnaire with a forced-choice response format that is designed to measure potential for child physical abuse. The CAP Inventory contains a 77-item child physical abuse scale and 3 validity scales (lie, random response, and inconsistency). The validity scales are used in various combinations to form three response distortion indices (faking-good, faking-bad, and random response) that can be used to discard participants whose responses are suspect.

Research supports the construct validity of the CAP’s physical child abuse scale, as well as its concurrent validity (for a review see Milner, 1986, 1993). For example, in cross-validation studies using physical child abusers and demographically
matched comparison parents, the CAP provided overall correct classification rates above 80% (e.g., Milner, 1989; Milner, Gold, & Wimberley, 1986). Temporal stability estimates range from .90 for 1-week to .75 for 3-months. Internal consistencies for responses from parents range from .92 to .98 (Milner, 1986).

**Adjectives used in the priming task.** Fifteen evaluative adjectives with positive connotations (e.g., attractive, charming) and 15 evaluative adjectives with negative connotations (e.g., annoying, repulsive) were used in the priming task. These exact words were used in several previous evaluative priming studies (e.g., Fazio et al., 1995; R. Fazio, personal communication, October 23, 2005). Using sentences that started with “He is…” and using each adjective in the study, the average Flesch-Kincaid grade level for all adjectives was 5, indicating that the average reading level of the adjectives was equivalent to that of a fifth grade reading level. An additional 10 adjectives (5 positive and 5 negative) were used on practice trials.

**Stimulus photographs used in the priming task.** Digital color photographs of 30 child faces and 30 adult faces available for download from the internet (www.punchstock.com) were used as primes. In each group of 30 photos, 10 photos depicted positive facial expressions, 10 depicted neutral/ambiguous facial expressions, and 10 depicted negative facial expressions.

The 60 photographs were selected from a pool of 332 photographs. Selections were based on ratings provided by undergraduates who participated in a pretest study (N = 103). Each pretest participant viewed each of the 332 photographs on a computer monitor accompanied by the question “How would you rate this individual?” and a 7-point ordered category scale with response options ranging from –3 (very negative) to 3 (very positive). Stickers with the values –3 through 3 were affixed to keyboard keys Z through M, and these were used by participants to make their responses. Ratings for each photograph were averaged across all participants.

Mean ratings and standard deviations for each photograph were calculated and used (in addition to facial expression, race, and gender) to select the 60 photographs used in the main study. The positive photographs evinced means ranging from 1.83 to 2.54, with standard deviations ranging from .70 to 1.09. The negative photographs evinced means ranging from –1.51 to –2.37, with standard deviations ranging from .85 to 1.37. The neutral photographs evinced means ranging from –.96 to .87 with standard deviations ranging from 1.01 to 1.54. The heightened standard deviation for the neutral faces is consistent with the idea that there was uncertainty about whether those photographs were positive or negative.

To select photos for use in the experiments child photographs and adult photographs were matched individually within valence category (e.g., positive, neutral, and negative) across five criteria: mean rating, standard deviation of rating, ethnicnicity, gender, and type of emotional expression (e.g., a negatively rated picture with a sad expression was paired with another negatively rated sad expression, not a negatively rated angry expression and vice versa). With the exception of seven pairs (six pairs did not match on race, one pair did not match on race or gender), all other pairs matched on all five criteria.

**Practice trial photographs.** Photographs used on practice trials were 15 child photos and 15 adult photos taken from the original pool of 332 that were not used in the priming task. There were 10 positive photographs, 10 neutral/ambiguous photographs, and 10 negative photographs (as determined by means from the pretest study described above). There were 14 photographs of females, 15 photographs of males, and 1 infant photograph (of undetermined gender).

**Filler task photographs.** The evaluative priming task protocol also included several filler tasks. Several presented photos as stimuli. In all cases, the photos used on these tasks were culled from the original pool of 332 that were not used in the priming task or on practice trials.

**Procedure.** On arrival at the lab, participants read and signed a consent form that described the purpose, risks, benefits, and voluntary nature of the study. After signing, each participant was seated in front of a desktop computer. Participants were asked to read the instructions on the computer screen and follow the prompts.

The study’s procedure followed Fazio’s 7-phase evaluative priming protocol (e.g., Fazio et al., 1995). Key phases were Phase 1, in which participant baseline data was collected, and Phase 4, in which the evaluative priming task was administered. Phases 2, 3, and 5 either: (1) provided practice with aspects of the task, (2) presented filler tasks, or (3) maintained the cover story. Participants completed the CAP and provided demographic data in Phases 6 and 7.

The first phase began with a set of instructions directing participants to attend to words to be presented, 1 at a time, in the middle of the computer screen. On each trial, the presentation of each word was preceded by a row of asterisks. Participants were asked to press the Z key if the word was “good” and the M key if the word was “bad.” They were instructed to respond as quickly and accurately as possible. Participants completed 10 practice trials. After completing these, as a reminder of the task requirements the instructions again appeared on the computer screen. Participants then completed 2 blocks of trials. One of the target adjectives (15 positive and 15 negative) appeared on each of the 30 trials in each block. On each trial, participants judged whether the word on the screen was “good” or “bad.” As in the practice task, each word was preceded by a row of asterisks. The mean response latency for the two presentations of each adjective served as a participant’s baseline response latency to each. [In the original Fazio et al. (1995) protocol, prior to data analysis these baseline measures were subtracted from the response latencies obtained on the priming task (phase 4). However, results of our main analyses were highly similar regardless of whether the difference scores were used or whether the raw latencies were used. Given this similarity and the oft-cited problems with analysis of difference scores (e.g., Charter & Feldt, 2009), in this article only results derived from the raw latencies are reported.]

Phases 2 and 3 were filler tasks. In Phase 2 a series of 12 faces (6 children and 6 adults) were presented, 1 at a time. In Phase 3 the participants saw a series of faces, some of which appeared in Phase 2, and were asked to decide whether they had seen each face before. Participants were instructed to press the A key if the photograph had been presented previously and the L key if the photograph was new. Each of 24 photographs (the 12 photos seen in Phase 2, 6 new child photos, and
Table 1
Response latency means and standard deviations for adjective valence by face valence (positive, neutral, negative).

<table>
<thead>
<tr>
<th>Sample</th>
<th>Student M</th>
<th>Student SD</th>
<th>Parent M</th>
<th>Parent SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive adjectives</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive valence</td>
<td>622.72a,b</td>
<td>80.75</td>
<td>752.77d</td>
<td>168.39</td>
</tr>
<tr>
<td>Neutral valence</td>
<td>637.18a</td>
<td>82.35</td>
<td>760.82e</td>
<td>182.67</td>
</tr>
<tr>
<td>Negative valence</td>
<td>639.01b</td>
<td>84.43</td>
<td>810.20f</td>
<td>273.02</td>
</tr>
<tr>
<td>Negative adjectives</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive valence</td>
<td>650.37c</td>
<td>87.91</td>
<td>807.98f</td>
<td>195.77</td>
</tr>
<tr>
<td>Neutral valence</td>
<td>642.75</td>
<td>89.33</td>
<td>807.91g</td>
<td>236.63</td>
</tr>
<tr>
<td>Negative valence</td>
<td>635.67c</td>
<td>88.46</td>
<td>782.66f</td>
<td>231.33</td>
</tr>
</tbody>
</table>

Notes: Student sample, N = 90. Parent sample, N = 95. Means that share a superscript were significantly different in post hoc tests using the Neuman–Keuls procedure. Tests were conducted on adjective pairs within each Sample × Face Valence cell.

6 new adult photos) was presented individually and remained on the computer screen until the participant responded by pressing 1 of the 2 response keys.

Phase 4 presented the experimental priming task. This phase began with a set of instructions that informed participants that the previous tasks would be combined. Participants were told that the judgment of word meaning is thought to be an automatic skill so individuals should be able to make judgments about the words even if they are asked to do another task. In this case, the other task was attending to and remembering faces. Participants were told that it was important that they attend to the faces that preceded the words because, as on prior tasks (in Phases 2 and 3) they would be asked to judge whether or not they had seen the face previously. Participants also were instructed to judge whether the word presented was “good” or “bad.”

On each trial, a row of asterisks appeared on the screen, cuing participants to the imminent appearance of a photo. The photographs served as primes, potentially facilitating responses to the target words, 1 of which appeared after each photo. As in Fazio et al. (1995) photographs were displayed for 300 ms. A word appeared on the screen about 300 ms after the face disappeared from the screen. Fazio (2001) reported that this short stimulus onset asynchrony (SOA; the interval between prime presentation and target presentation) is important, allowing priming effects to be interpreted in terms of automatic activation of affect. Priming effects using longer SOAs (>300 ms) could be caused by controlled processes instead of automatic processes.

To complete the task, participants first completed a block of 30 practice trials. These trials used photographs and words that were not presented in the priming task. The priming task was presented immediately after completion of the practice trials. This priming task contained 4 blocks. Each block consisted of 60 trials in which each of the 60 photographs was presented once and each of the 30 adjectives was presented twice. Within each block each photograph and adjective were presented in random order without replacement. Over the course of the 4 blocks, each of the 60 photographs was presented 4 times and paired with 2 of the 15 positive adjectives and 2 of the 15 negative adjectives. On each trial, the speed with which participants judged each word to be “good” or “bad” was recorded. This response latency was the dependent measure in the experiment. Consistent with established practice, a 100 ms floor and a 5000 ms ceiling were applied to these latencies. These cutoffs limit the effect that outliers have on latency means and standard deviations (see Ratcliff, 1993).

Phase 5 presented a recognition task. Participants decided whether pictures had been presented previously. Twenty-four photographs were presented: 12 child photos and 12 adult photos. The photo set contained both new (6 child and 6 adult), and previously seen (six child and six adult) faces. This task was used solely to maintain the integrity of the cover story.

The experiment’s last two phases involved completion of demographic questions (Phase 6) and the CAP Inventory (Phase 7). After participants completed the CAP they were thanked for their participation and were given their debriefing.

Results and discussion

In an initial set of analyses, response latencies from Phase 4 of the priming task were entered into a four-variable [CPA Risk (high, moderate, low) × Adjective Valence (positive, negative) × Face Type (child, adult) × Face Valence (positive, neutral, negative)] mixed ANOVA. CPA risk was the sole between-subjects variable in the design.

The significant Adjective Valence × Face Valence interaction, $F(2, 172) = 10.95, p < .001$, partial $\eta^2 = .113$ replicates the standard evaluative priming effect obtained by those who have used the Fazio paradigm (e.g., Fazio et al., 1986; see Table 1 for means and standard deviations). This conclusion is supported by the results of pairwise mean comparisons evaluated using the Newman–Keuls procedure. Results of these comparisons indicated that there were significantly shorter latencies for positive adjectives when preceded by positively valenced faces than when preceded by either neutrally valenced faces or by negatively valenced faces. In comparison, response latencies for negative adjectives were significantly shorter when preceded by negatively valenced faces than when preceded by positively valenced faces, but neither of these latencies differed from the latency for negative words preceded by neutrally valenced faces. This evaluative priming effect was not
Table 2
Means and standard deviations for response latency for adjective valence by face type (child vs. adult).

<table>
<thead>
<tr>
<th></th>
<th>Sample</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Student</td>
<td>Parent</td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Positive adjective</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child face</td>
<td>634.79</td>
<td>83.25</td>
<td>783.17</td>
</tr>
<tr>
<td>Adult face</td>
<td>631.09</td>
<td>77.50</td>
<td>760.57</td>
</tr>
<tr>
<td>Negative adjective</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child face</td>
<td>635.74^a</td>
<td>88.82</td>
<td>786.09^a</td>
</tr>
<tr>
<td>Adult face</td>
<td>650.18^b</td>
<td>85.65</td>
<td>815.26^b</td>
</tr>
</tbody>
</table>

Notes: Student Sample, N = 90. Parent Sample, N = 95. Means that share a superscript were significantly different in post hoc tests using the Neuman–Keuls procedure. Tests were conducted on adjective pairs within each Sample × Face type cell.

The similarity in responding for high CPA risk and low CPA risk participants also extended to their responses to different face types. For example, the analysis yielded a significant Adjective Valence × Face Type interaction, $F(1, 86) = 7.32, p < .01$, partial $\eta^2 = .078$ (see Table 2 for means and standard deviations). The Newman–Keuls paired mean comparison procedure used to explore this interaction revealed no significant response latency difference between face types that preceded positive adjectives. For negative adjectives, however, response latencies were significantly shorter when preceded by a child face ($M = 635.74$), than when preceded by an adult face ($M = 650.18$). This result again replicates the results provided by Fazio et al. (1986), in that it likely reflects a preference of our college-aged respondents for their in-group (other adults) over the out-group (children). This interaction was not moderated by CPA risk [CPA Risk × Adjective Valence × Face Type interaction, $F(2, 86) = 1.80, p = .17$, partial $\eta^2 = .040$; CPA Risk × Face Type interaction, $F(2, 86) = 2.22, p = .10$, partial $\eta^2 = .021$]. Hence, the in-group evaluative preference for adults reflected in these responses was equally shared by high risk respondents and low-risk respondents.

These results suggest that CPA risk is unrelated to evaluative reactions to faces. However, the analyses yielding those results may cast too large a net. Recall that one interest of this paper was to better understand the results provided by Farc et al. (2008), which found that high CPA risk participants were especially likely to rate neutral child faces as reflecting the trait of hostility. Accordingly, in a second set of focused analyses we explored the possibility that high risk participants and low risk participants had different evaluative reactions to the neutral child faces that they encountered in this experiment. To compensate for the reduction in the number of observations entered into this second set of analyses and to maximize analytic power, in this second set of analyses we compared the latencies of responses to words that were preceded by neutral child faces (obtained in Phase 4) to the latencies of responses to words that were not preceded by faces (obtained in Phase 1). We entered the latencies obtained into a CPA Risk (high, moderate, low) × Adjective Valence (positive, negative) × Trial Type (unprimed, primed) mixed ANOVA. CPA risk was the sole between-subjects variable in the design.

Converging with the results of the main overall analysis, the risk group variable did not predict response latencies in any main effect or interaction [largest $F(2, 82) = .98, p = .38$, partial $\eta^2 = .023$]. These data are not consistent with the idea that risk group is related to the tendency to experience a negative evaluative reaction to neutral child faces.

Only 3 significant results emerged from the analysis, and only one is of theoretical interest. Responses to negative words were more greatly facilitated by exposure to the neutral child face primes (baseline $M = 910.09$; priming $M = 629.61$) than were responses to positive words (baseline $M = 770.15$; priming $M = 633.84$). $F(1, 82) = 3.96, p = .050$, partial $\eta^2 = .051$. This result dovetails with the result observed in the main analysis, and is consistent with negative evaluation of out-groups: That is, for the young adult sample, out-group members (children) appear to be evaluated negatively. Interpretation of both the significant priming main effect, $F(1, 82) = 26.91, p < .001$, partial $\eta^2 = .247$, and the word valence main effect that approached significance, $F(1, 82) = 3.56, p = .063$, partial $\eta^2 = .042$, are obviously both qualified by the significant interaction. Hence, no interpretation of these effects is offered.

Experiment 2

Results from Experiment 1 indicated that CPA risk, as indexed by scores on the CAP, was not related to the automatic evaluation of children displaying neutral facial expressions. This result suggests that the results provided by Farc et al. (2008), in which CAP-indexed risk status was related to the tendency to rate neutral child faces as displaying the trait of hostility, were caused by the activation of semantic interpretive frameworks which were used to interpret the faces, and not by automatic evaluative responses to the faces. However, the absence of such effects in Experiment 1 could have been a function of the non-parent student sample that was used. That is, it is possible that parenting experience might cause high-, moderate-, and low-risk parents to have stronger attitudes about children than high-, moderate-, and low-risk nonparents. To explore this possibility, and to again
explore whether it is possible that the Farc et al. (2008) results were caused by parents’ automatic evaluative responses to child faces, an evaluative priming study was conducted with parents as participants.

Method

Participants. Parents 18 years of age and older were recruited from social service agencies and daycare centers in the Midwest. Recruitment efforts involved distribution of fliers, contacts made at community programs and daycare centers, solicitations made from birth announcements, and word of mouth.

A total sample pool of 141 parents participated, but the CAP scoring scheme classified the responses of 46 parents as invalid. Thus, the final study sample consisted of 95 parents. Parents who scored below 91 on the CAP comprised the low-risk group (N = 40), parents who scored between 91 and 165 on the CAP comprised the moderate-risk group (N = 20), and parents who scored 166 and above on the CAP comprised the high-risk group (N = 35). No significant differences (all ps > .05) emerged on any of the demographic variables for parents that were included in the final study sample and parents that were excluded due to invalid CAP responses.

The mean age (in years) of the individuals in the final sample was 31.8 (SD = 9.5). The number of children living with parents ranged from 0 to 13, with a mean of 2.4 children (SD = 2.0). The participants in the sample completed a mean of 13.5 years of education (SD = 2.9), and contained both males (34.8%) and females (65.2%). The sample was ethnically diverse [Caucasian (53.7%), African-American (33.6%), Latino (6.3%), Asian-American (1.1%), Native American (1.1%) and other (4.2%)] and reflected diversity in marital status [unmarried (44.2%), married (41.1%), divorced/separated (10.5%), and other (4.2%)]. The mean income reported by the final study sample was $26,740 (SD = $27,111). Analyses indicated that the low-, moderate-, and high-risk groups did not significantly differ on any assessed demographic variable.

Procedure. Differences between the materials, measures, and procedures used in Experiment 2 and those used in Experiment 1 were minor. Among those differences were: (1) the content of the demographic questionnaire changed slightly to reflect the adult–parent nature of the sample; (2) the building and rooms in which Experiment 2 was conducted differed from those used for Experiment 1; (3) participants received $25 (US) as compensation for participation; and (4) because we wanted to make sure that the parent sample was comfortable with the computer task, the number of practice trials in Phase 4 was increased from 30 to 60.

Results and discussion

In an initial set of analyses, response latencies from Phase 4 of the priming task were entered into a 4-variable [CPA Risk (high, moderate, low) × Adjective Valence (positive, negative) × Face Type (child, adult) × Face Valence (positive, neutral, negative)] mixed ANOVA. CPA risk was the sole between-subjects variable.

The analysis yielded a significant Adjective Valence × Face Valence interaction, $F(2, 180) = 13.85, p < .001$, partial $\eta^2 = .133$, (see Table 1 for means and standard deviations), replicating the standard evaluative priming effect obtained by those who have used the Fazio paradigm (e.g., Fazio et al., 1986). This conclusion is supported by the results of pairwise mean comparisons evaluated using the Newman–Keuls procedure: (1) Response latencies to positive adjectives were significantly shorter when preceded by positive faces than when preceded by negative faces, but did not differ from positive adjectives preceded by neutral faces; and (2) response latencies to negative adjectives were significantly shorter when preceded by negative faces than when preceded by positive faces or neutral faces. This evaluative priming effect was not moderated by CPA risk [Adjective Valence × Face Valence × CPA Risk interaction, $F(4, 180) = .10, p = .98$, partial $\eta^2 = .002$; Adjective Valence × Face Valence × CPA Risk interaction, $F(4, 180) = .10, p = .98$, partial $\eta^2 = .002$]. Hence, it appears that evaluative reactions to the photos did not depend on respondent CPA risk.

Similarity in responding for high CPA risk participants and low CPA risk participants extended to responses to different face types. For example, the analysis yielded a significant Adjective Valence × Face Type interaction, $F(1, 90) = 4.14, p = .045$, partial $\eta^2 = .044$ (see Table 2 for means and standard deviations). Exploration of this interaction using the Newman–Keuls procedure revealed no significant response latency difference between face types that preceded positive adjectives. However, for negative adjectives, response latencies were significantly shorter when preceded by a child face ($M = 786.09$), than when preceded by an adult face ($M = 815.26$). This result replicates the results provided by Fazio et al. (1986), reflecting a preference of parents for their in-group (adults) over an out-group (children). This interaction was not moderated by CPA risk [CPA Risk × Adjective Valence × Face Type interaction, $F(2, 90) = .07, p = .94$, partial $\eta^2 = .001$; CPA Risk × Face Type interaction, $F(2, 90) = 1.26, p = .29$, partial $\eta^2 = .027$]. Hence, the observed in-group evaluative preference for adults was equally shared by high risk respondents and low risk respondents.

As in Experiment 1, a second set of focused analyses explored the possibility that high CPA risk participants and low CPA risk participants had different evaluative reactions to the neutral child faces. To compensate for the reduction in the number of observations entered into this second set of analyses and to maximize analytic power, in this second set of analyses we compared the latencies of responses observed in response to words that were preceded by neutral child faces (obtained in Phase 4) to the latencies of responses to words that were not preceded by those faces (obtained in Phase 1). We entered the latencies obtained into a CPA Risk (high, moderate, low) × Adjective Valence (positive, negative) × Trial Type (unprimed, primed) mixed ANOVA. CPA risk was the sole between-subjects variable in the design.
Converging with the results of the main overall analysis, the risk group variable did not predict response latencies in any main effect or interaction [largest F(2, 91) = 2.01, p = .14, partial η² = .042]. Thus, these data are not consistent with the idea that risk group is related to the tendency to evaluate neutral child faces in a negative fashion.

Three significant results emerged from the analysis; only 1 is of theoretical interest. Responses to negative adjectives were more greatly facilitated by exposure to the neutral child face primes (baseline M = 892.93; priming M = 762.02) than responses to positive adjectives (baseline M = 824.95; priming M = 761.62), F(1, 91) = 14.51, p < .001, partial η² = .137. This result is consistent with the idea that adults negatively evaluate out-group members (children) relative to in-group members (adults).

Interpretation of both the significant priming main effect, F(1, 91) = 20.67, p < .001, partial η² = .185, and the significant Adjective Valence main effect, F(1, 91) = 12.69, p = .001, partial η² = .122, are obviously both qualified by the significant interaction. Hence, we offer no interpretations of these effects.

General discussion

A primary goal of our studies was to better understand mechanisms driving the results obtained by Farc et al. (2008). They found that parents high in CPA risk were especially likely to rate children displaying neutral emotional expressions as hostile and difficult. However, the exact cause of this effect was unclear. That is, these effects could have occurred because high risk parents are especially likely to believe (i.e., possess semantic knowledge) that children are hostile, and thus, were especially likely to interpret the facial expressions as reflecting hostility. Alternatively, these results could reflect the fact that high CPA risk parents were especially likely to have a negative evaluative reaction to children displaying neutral facial expressions.

To examine an evaluation-based explanation for these results, we used an evaluative priming task in which we presented photos of children whose facial expressions were neutral, followed those faces with negative adjectives, and observed whether the faces speeded responses to a question about the adjective’s valence. Our primary interests were: (1) whether response facilitation emerged on these trials, and (2) whether the amount of facilitation differed for people who differed in CPA risk. Other task trials examined whether positive facial expressions and negative facial expressions facilitated responses to valence-congruent adjectives, and whether such facilitation varied by respondent CPA risk.

Results of the two studies reported herein indicate that adult respondents (students and parents alike), regardless of CPA risk, display evidence of automatic negative evaluative reactions to child photos. The data also suggested that this effect is a consequence of adults’ implicit negative evaluation of their age out-group (children) relative to their age in-group (adults), and that this in-group/out-group bias emerges regardless of the facial expression depicted in the photos.

Although both high CPA risk parents and low CPA risk parents exhibited implicit negative evaluative biases toward children, it remains possible that these parent groups may differ in how they manage these initial negative evaluative reactions. For example, low CPA risk parents may find it easier to switch their attention to positive or benign aspects of child-related situations, and in doing so, prevent escalation of negative evaluative and affective reactions. As a second example, consider that while all parents may initially react to a crying child with implicit negativity; low CPA risk parents may be more effective in instantiating alternative thoughts and interpretations that mitigate this initial negative reaction (e.g., “its amazing how one so little can be so loud,” or “sounds like somebody needs a nap”). Indeed, accumulating evidence indicates that high CPA risk parents may have deficits in attentional control (Crouch et al., in press), which could manifest as a reduced capacity to shift attention from dominant negative evaluative reactions to more positive/benign interpretations or attributions. Thus, although high and low CPA risk parents may not differ with respect to their automatic negative evaluative reactions to children, such biases may have an especially large impact on subsequent interpretive, affective, and behavioral responses in high CPA risk parents.

Further, a variety of other cognitive, emotional, and personality factors may moderate how out-group biases toward children are overtly expressed by high and low CPA risk parents (Hewstone, Rubin, & Willis, 2002). For example, heightened perceptions of differences in status between the in-group and out-group members may actually lead members of the high status group to exhibit more generous, forgiving, and magnanimous behavior toward lower status out-group members. Thus, to the extent that low CPA risk parents achieve greater differentiation of self (i.e., view self as separate from the child; Skowron & Platt, 2005) and make clearer distinctions with respect to the roles of parents and children (Bavolek & Keene, 1999), they may behave with more beneficence toward children perceived as lower status out-group members.

However, when interactions with out-group members involve strong negative emotions (e.g., anger, contempt), hostile and aggressive behaviors directed toward out-group members become more likely (Hewstone et al., 2002). Thus, to the extent that CPA risk status is differentially associated with negative emotional states, then one might expect that such emotions may moderate the expression of aggression toward children perceived as out-group members. Indeed, high CPA risk parents are especially prone to experiencing negative affective states (e.g., anger, distress; Milner, 1986) and parental negative affect is predictive of harsh; punishing, and aggressive behaviors directed toward children (Mammen, Kolko, & Pilkonis, 2002; Rodriguez & Richardson, 2007).

Whereas negative emotions increase the likelihood of out-group derogation, other factors, such as the ability to experience empathy for out-group members, may help reduce overt expression of out-group biases (Finlay & Stephan, 2000). Given that low CPA risk parents tend to obtain higher scores on measures of dispositional empathy compared to high CPA risk parents (Rodriguez & Richardson, 2007; Wiehe, 2003), it is possible that CPA risk group differences in empathy may moderate the overt expression of negative implicit biases toward children. It also is noteworthy that empathic responses tend to be weaker...
for out-group members then for in-group members (Society for Neuroscience, 2009), a phenomenon which may erode the already limited capacity for empathy in high CPA risk parents as they interact with children.

Another, perhaps more subtle, expression of out-group bias involves unequal application of rules of social justice to in-group members versus out-group members. Although it is easy to see how justification of aggression toward children due to an out-group bias might contribute to physical abuse, a broader implication with respect to use of corporal punishment of children also might be considered. More specifically, implicit out-group biases toward children may underlie widespread beliefs regarding the justification of aggression toward children and foster the double standard that says it is permissible to strike a child but illegal to do the same to an adult. Several studies involving nationally representative samples from the United States suggest that between 67% and 94% of parents reported using corporal punishment to discipline toddlers and up to 93% of young adults report being spanked as children (e.g., Flynn, 1998; Graziano & Namaste, 1990; Straus & Stewart, 1999; Wissow, 2001). The extent to which pervasive out-group biases toward children underlie unequal protection of children and adults from physical aggression remains an important empirical question.

Of course, one might argue that our conclusions essentially rest on a null effect—the absence of risk group differences in negative evaluative reactions to children. The standard argument is that one ought to be very concerned about drawing strong conclusions from null effects. However, there are exceptions to every rule, and such exceptions are applicable here. The logic for caution when interpreting null effect findings revolves around the idea that a null result can occur either because an effect is not actually there, or the null result can be a consequence of a methodology that is not adequate to detect an effect. In this regard, we note that the methods and measures used in the present experiments have been extensively used and validated. For example, the CAP has been extensively validated as a measure of CPA risk (see Milner, 1986, 1989). Moreover, parental risk status was measured in the present studies using exactly the same instrument (the CAP) employed by Farc et al. (2008). Thus, it is not very plausible that the failure to detect a risk group effect in the studies reported in the present article reflects questionable practices with respect to CPA risk group classification.

Similarly, one cannot point to deficits in our evaluative priming procedure as a cause of the null result. The evaluative priming paradigm used in the studies described in the present article produced evidence of both facial expression-driven evaluative priming effects and of in-group preference effects. Both effects have emerged in earlier studies using this paradigm (e.g., Fazio et al., 1986). Given our replications of these effects, we believe that the best conclusion is that the paradigm used in the present studies detected automatic evaluative priming effects, and it did so in two ways (via facial expressions and in-group preference).

Accordingly, in our view, the failure to obtain moderation of evaluative priming by CPA risk in the present research should best be viewed as an informative companion to the results provided by Farc et al. (2008), serving to narrow the range of cognitive mechanisms that might explain the results obtained in that study. Our data suggest that those results were caused by a tendency of high risk parents to interpret the facial expressions as a manifestation of hostility, and not by an implicit negative evaluative reaction to such photos.

Such a conclusion has prior precedent. First, it is often argued that different experimental tasks can tap into different kinds of knowledge (e.g., semantic knowledge as opposed to evaluative knowledge; see Judd, Blair, & Chapleau, 2004). Moreover, although semantic knowledge (e.g., beliefs about people) and evaluative knowledge (preferences for or judgments of people) are often assumed to be related, evidence suggests that semantic knowledge and evaluative knowledge can sometimes be independently represented in knowledge structures (Eagly & Chaiken, 1993). For example, Judd et al. (2004) reported that tasks examining racial stereotypes and prejudices can evoke separate and distinct effects on social information processing for semantic information (e.g., group-based stereotypic beliefs) and evaluative information (i.e., prejudices derived from group membership information).

However, one needs to be careful about the extent to which one uses the present data to claim that there are no differences in implicit attitudes toward children exhibited by abusers and non-abusers. Although measurement of CPA risk using the CAP (Milner, 1986) has been found to be a valid proxy for abuse risk, it is not equivalent to assessing actual abusive caregiver behavior. Additionally, the stimuli used in the present studies depicted children who were unknown to the participants; whether similar results would be obtained when the priming stimulus is a parent’s own child is an open question. Moreover, though results of a power analysis suggest that a sufficiently large sample was used to detect reasonably sized effects, it could be that a parent risk effect was present, but the sample sizes used in the present studies were too small to detect it. This would be remedied by replication with a larger sample size.

Finally, it should be noted that the implicit evaluative priming procedure is only one method that can be applied to the pursuit of automatic evaluative responses to stimuli. Other instruments that can be used to similar purpose are the Implicit Association Test (Noesek & Hansen, 2008) and the Affect Misattribution Procedure (Payne, Burkley, & Stokes, 2008). Importantly, results obtained using these three implicit attitude assessment procedures are sometimes dissimilar. Hence, the possibility remains that one might be able to use one of these other techniques to garner evidence suggestive of especially negative implicit reactions to neutrally faced children among parents high in abuse risk. Should that be the case, then our conclusion with respect to the mechanism responsible for the Farc et al. (2008) results would need to be revised.

Acknowledgement

The authors would like to acknowledge the contributions of Joel S. Milner.