Journal of Experimental Child Psychology 109 (2011) 187–200

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Measuring implicit attitudes of 4-year-olds: The Preschool Implicit Association Test

Dario Cvencek [a,*], Anthony G. Greenwald [b], Andrew N. Meltzoff [a]

[a] Institute for Learning & Brain Sciences, University of Washington, Seattle, WA 98195, USA
[b] Department of Psychology, University of Washington, Seattle, WA 98195, USA

ARTICLE INFO

Article history:
Received 20 August 2010
Revised 8 November 2010
Available online 6 January 2011

Keywords:
Social cognition
Gender attitudes
In-group bias
Sex typing
Implicit Association Test
Preschool children

ABSTRACT

The Preschool Implicit Association Test (PSIAT) is an adaptation of an established social cognition measure (IAT) for use with preschool children. Two studies with 4-year-olds found that the PSIAT was effective in evaluating (a) attitudes toward commonly liked objects (*flowers* = *good*) and (b) gender attitudes (*girl* = *good* or *boy* = *good*). The gender attitude PSIAT was positively correlated with corresponding explicit attitude measures and also children's actual sex. The new implicit and explicit measures of gender attitudes demonstrated discriminant validity; each predicted variance in children's gendered play activities beyond that predicted by the other. Discussion describes potential uses of the PSIAT to investigate development of societally significant attitudes and stereotypes at younger ages than are achievable with currently available methods.

Introduction

Can the attitudes of preschool children be measured without using self-report? Implicit measures of attitudes based on response latency techniques, such as the Implicit Association Test (IAT) (Greenwald, McGhee, & Schwartz, 1998), are already routinely used in research with adults. They have been adapted for children as young as 6 years of age (Baron & Banaji, 2006; Cvencek, Meltzoff, & Greenwald, in press; Dunham, Baron, & Banaji, 2006; Rutland, Cameron, Milne, & McGeorge, 2005; Steffens, Jelenec, & Noack, 2010). However, cognitive demands of these procedures have so far prevented their use with still younger children (Castelli, Zogmaister, & Tomelleri, 2009; Olson & Dunham, 2010). This article describes adaptations of the IAT that enable its successful use with 4-year-olds.

* Corresponding author.
E-mail address: dario1@u.washington.edu (D. Cvencek).

0022-0965/$ - see front matter
doi:10.1016/j.jecp.2010.11.002

The term *implicit* has been used to refer to measurement methods that (a) do not require introspective access, (b) reduce the amount of available mental control to produce the response, and (c) reduce the role of intentional processes (Nosek, Greenwald, & Banaji, 2007). Implicit *attitudes*, as measured by the IAT, are thought of as links that connect a group category (e.g., *male* or *female*) to a valence category (e.g., *good* or *bad*). The IAT provides an implicit measure of attitudes by comparing response speeds in a double-categorization task that combines a concept classification (e.g., boy vs. girl) and an attribute classification (e.g., good vs. bad), with either or both of those dimensions represented as pictures or words. For example, if a participant is significantly faster when *girl* and *good* share a response button than when *boy* and *good* do, she can be said to have an *implicit* attitude that links "girls" rather than "boys" with a positive valance of "good".

Developing implicit measures of attitudes for young children is potentially useful. First, such measures permit assessment of spontaneous aspects of social cognition that may lie below children's conscious awareness. Second, explicit measures might not always capture attitudes accurately in this age group, in part, because of children's emerging tendency to respond in socially desirable ways (Brody, Rozek, & Muten, 1985). Third, to identify the origins of implicit attitudes and self-concepts that have already been demonstrated in elementary school children (Baron & Banaji, 2006; Cvencek et al., in press; Degner & Wentura, 2010; Rutland et al., 2005), researchers will require new methods that are appropriate for preschool-age children (see Olson & Dunham, 2010).

For the first test of the Preschool IAT (PSIAT), two attitudes that have been thoroughly assessed in older children and adults were chosen. Previous results show that flowers elicit more positive attitudes than insects on implicit measures (Baron & Banaji, 2006; Thomas, Burton Smith, & Ball, 2007). Thus, the effectiveness of the PSIAT to test this pervasive attitude was assessed first (Experiment 1). In Experiment 2, the adapted PSIAT was used to investigate gender attitudes of 4-year-olds. Previous research on implicit measures with older children and adults has repeatedly shown that females associate female gender more strongly with positive valence (*female* = *good*) than they associate males with a positive valence (*male* = *good*), whereas males associate both genders equally strongly with positive valence (Ebert & Steffens, 2008; Rudman & Goodwin, 2004; Skowronski & Lawrence, 2001).

In addition to conducting implicit measures, explicit verbal reports on the same children were also recorded. Past research with adults has shown that correlations between implicit and explicit measures vary depending on the domain tested (Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005; Nosek, Greenwald, & Banaji, 2005). Not much is known about these relations during childhood because relatively few studies have used both implicit and explicit measures with the same children. The two experiments reported here conducted both implicit and explicit assessments on the same children to begin to build the relevant database; however, the main focus of this article is not to theorize about the relation between these two measures but rather to present an adaptation of a method and provide the evidence for its consistency, validity, and proper use.

Three hypotheses were examined: (a) the PSIAT will provide evidence of implicit preference for flowers (Experiment 1); (b) girls will demonstrate significant positive attitudes toward female gender on both implicit and self-report measures, whereas boys will demonstrate more neutral gender attitudes (Experiment 2); and (c) the PSIAT and explicit measures of children's attitudes will not be duplicates of one another but rather will provide different and incremental information for predicting children's gendered activities (Experiment 2).

Experiment 1: Establishing the PSIAT method

Method

Participants

A total of 65 4-year-olds (M = 54.51 months, SD = 1.20, 32 girls and 33 boys) were recruited for this experiment by telephone from the university's participation pool. According to parental reports, all children were typically developing and the racial/ethnic makeup of the sample was 83.1% White, 3.1% Asian, 1.5% Native American, 7.7% multiracial (i.e., more than one race), and 4.6% other/unknown, with 7.7% reported to be of Hispanic ethnicity. Each family received $10 in exchange for participation.

Procedure

Children were tested individually in a quiet room at the university and were told that they first would be "asked some questions" and then would "play a game on a computer." The explicit measure was administered first. Previous meta-analyses of IAT studies in adults found no effect on the implicit and explicit measure means as a function of order in which the implicit and explicit measures were presented (Nosek et al., 2005, $k = 65$), nor did they find that the order of presentation had an effect on implicit–explicit correlations (Hofmann et al., 2005, $k = 261$). Given the robust absence of order effects from these two independent meta-analyses, there is no reason to believe that a fixed order of measures in the current two experiments constitutes a severe limitation.

Explicit measure of flower–insect attitudes. The explicit measure was administered as two bipolar scales using the same response format as in Harter and Pike's (1984) Pictorial Scale. For each scale, children were shown two composite images: one with four flowers and one with four insects (see supplementary material). The experimenter explained each composite image while pointing to it (e.g., "On the left we have some flowers" and "On the right we have some bugs"). For each scale, children were asked to report on a 4-point scale (a) which composite image they liked more (e.g., "Do you like flowers more or do you like bugs more?") and (b) the degree to which they liked the selected composite image (e.g., "How much do you like flowers [bugs]? A little or a lot?"). The latter was done by children's pointing to one of two circles (1.1 and 2.3 cm in diameter corresponding to *a little* and *a lot*) assigned the values of 1 and 2. The responses from two questions were averaged to arrive at the explicit score with lower and upper bounds of −2 and +2; positive values indicated that children reported liking flowers more. The order of the two questions was counterbalanced across participants, as was the left–right assignment of flower and insect composite images. These self-report measures were not administered to 12 of the 65 participants because they had not been developed at testing.

The PSIAT and its use as an implicit measure. In adults, the IAT is a computerized double-categorization task that measures relative association strengths among concepts without self-report (Greenwald et al., 1998). In this experiment, the standard adult IAT was adapted for use with 4-year-olds. Modifications were similar to those in previous child IAT procedures (Baron & Banaji, 2006; Cvencek et al., in press; Dunham et al., 2006; Rutland et al., 2005; Steffens et al., 2010), including an adapted keyboard and presentation of spoken words, but extended well beyond them to make the test suitable for preschoolers. Full details of all the adaptations are provided below.

PSIAT adaptations for 4-year-olds. The development of the PSIAT included a number of adaptations beyond past work on adults and school-age children. The significant adaptations are described below.

Apparatus. Fig. 1 is a pictorial representation of the PSIAT apparatus. Children were seated at a small desk facing a computer (56-cm diagonal screen). The keyboard was furnished with two large response buttons ($14 \times 10.5 \times 0.5$ cm) that were color-coded—one orange and one green (see Fig. 1). An orange stripe and a green stripe (each 10 cm wide) appeared at the left and right edges of the computer screen. These colors indicated to children which of the two buttons to push to provide a correct response (e.g., the correct response for a stimulus "on the orange side" would be the orange button).

Stimuli. Each PSIAT task alternated between picture and word stimuli from one trial to the next. The dimensions of the picture stimuli were 7×10.5 cm. To eliminate the need for reading, all words were digitally recorded by a female speaker as 24-bit .wav files. This use of recordings by a female native English speaker is customary in past IAT research with children (Baron & Banaji, 2006; Dunham et al., 2006). Words also appeared on the screen written in neutral gray, thereby serving as a cue for children to respond. Similarly, the onset of each picture was synchronized with a short beep sound. That way, each trial involved both an audio presentation (i.e., spoken word or beep sound) and a visual presentation (i.e., written word or picture). The intertrial interval, between the response and the next stimulus, was 500 ms.

Visual reminders. For each of the four concepts (*flowers, insects, good,* and *bad*), children had visual reminders that remained in view during each task (Fig. 1). The reminders for *flowers* and *insects*

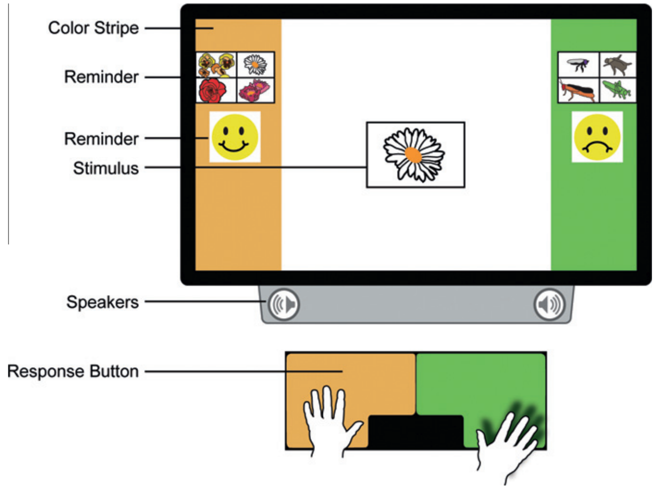

Fig. 1. Child's view of the apparatus used for the flower–insect attitude PSIAT. During the PSIAT, children categorize stimuli from four categories (*flower, insect, good,* and *bad*). The four categories are assigned to two response buttons: one operated with the left hand and the other operated with the right hand. Stimuli are presented one at the time in the center of the screen. The presentation of each stimulus on the screen is synchronized with the sound presentation through the speakers. Colored stripes on either side of the screen (one orange and one green) contain visual reminders for each of the four categories. Reminders for flowers and insects appear as a collage of all flower and insect stimuli, respectively. Reminders for good and bad words appear (below the two collages) as smiley and frowny faces, respectively. The child's task is to sort the stimuli as soon as they appear in the center of the screen by pressing one of the response buttons.

appeared as a collage of all flower and insect stimuli, respectively. The reminders for *good* and *bad* words appeared as "smiley" and "frowny" faces, respectively.

Length. To reduce the need for children to sustain their attention for a long period of time, the standard IAT length of 180 trials (Greenwald, Nosek, & Banaji, 2003) was reduced by 20%. Each task involving the discrimination between two concepts consisted of 16 trials, and each task involving the discrimination among four concepts consisted of 24 trials, for a total of 144 trials.

Instructions and feedback. To ensure that children understood and completed each PSIAT, the experimenter was present during the session and gave instructions orally. Children were told to respond to the stimulus "as soon as you know which one it is" and were encouraged to poise their hands above the buttons so as to respond "as soon as possible." Category reminders placed in the top left and top right of the computer screen (see Fig. 1) signified the correct and error responses. In a task where flowers and good words shared the left response key (and insects and bad words shared the right response key), pressing the left button constituted the *correct response* for pictures of flowers and good words and pressing the right button constituted the correct response for pictures of insects and bad words (see Fig. 1).

In this task, pressing the right button constituted an *error response* for pictures of flowers and good words and pressing the left button constituted an error response for pictures of insects and bad words. Error responses (i.e., hitting the wrong button) were followed by a red question mark appearing below the stimulus. Following an error response, children could not advance to the next trial until they provided the correct response. Latency was recorded to the occurrence of the correct response, thereby creating a built-in error penalty. This definition of what constitutes correct and error responses, as well as the use of a built-in error penalty, is standard in IAT research with adults (Greenwald et al., 2003). Children were given positive feedback after completion of every task (e.g., told "good job," shown thumbs up, or given a sticker).

PSIAT measure of flower–insect attitudes. During the flower–insect attitude PSIAT, children first practiced sorting pictures of *flowers* and *insects*. For example, they responded to pictures of flowers by pressing the left response button and responded to pictures of insects by pressing the right response button. After that, children practiced sorting *good* words (good, happy, fun, and nice) and *bad* words (bad, yucky, mean, and mad) using the same two response buttons.

Following these two single-discrimination tasks (16 trials each), children completed two combined discrimination tasks in which all four concepts were used. Each combined discrimination task consisted of two blocks of 24 trials for a total of 48 trials per combined task. During the combined tasks, two of the four concepts were mapped onto the same response button. In one instructional condition, pictures of *flowers* and *good* words shared a response button and pictures of *insects* and *bad* words shared the other response button. In the other instructional condition, two of the response assignments were reversed; pictures of *insects* and *good* words shared one response button.

An IAT score (*D*) (Greenwald et al., 2003) was calculated by comparing children's response speed in one instructional condition relative to the other instructional condition. As is typical in IAT research with adults, the standard *D-as-is* measure was used as a unit of analysis. This measure makes use of built-in error penalties as described in detail by Greenwald and colleagues, 2003. In line with previous research on implicit measures demonstrating robust associations between flowers and positive valence, positive scores indicated stronger association of *flowers* with *good* and of *insects* with *bad* than vice versa (see Table 1 for a detailed report of average latencies on the PSIAT measure).

The gender of the experimenter was counterbalanced across participants, as was the left–right orientation of *flower–insect* reminders, with the left–right orientation of reminders of *good* and *bad* categories remaining unchanged throughout the task. The order of the two instructional conditions was counterbalanced within each PSIAT.

Consistent with procedures from previous IAT studies with adults (Greenwald et al., 2003) and children (Cvencek et al., in press), PSIAT data were excluded for participants who were identified as outliers on the basis of the three preestablished criteria: (a) $\geqslant$10% of responses faster than 300 ms, (b) error rate of $\geqslant$35%, or (c) average response latency 3 standard deviations above the mean response latency for the whole sample. These criteria excluded 9 participants (13.8%), leaving $N = 56$ children (27 girls and 29 boys) for analysis (a more detailed discussion of data reduction considerations can be found in the supplementary material). The analyses following data reduction provided increased power, but the pattern of significant results remained unchanged.

Results and discussion

Fig. 2 displays the means for girls and boys on the implicit and explicit measures of flower–insect attitudes (both in the same standard deviation units). The implicit measure was entered as the dependent variable in a 2 (Gender: male or female) $\times$ 2 (Order of Combined Tasks: flowers/good first or insects/good first) between-subjects analysis of variance (ANOVA) to examine differences by gender and order of combined tasks. On the implicit measure, there was strong evidence for positive attitudes toward flowers, as indicated by a statistically significant intercept, $F(1, 52) = 14.06$, $p < .001$, $d = 0.50$. The same analysis revealed a main effect of gender, with girls associating *flowers* with *good* more strongly than did boys, $F(1, 52) = 5.03$, $p < .001$, $d = 0.62$. Further analyses using one-sample *t* tests showed that the gender effect was due to the fact that the association of *flowers* with *good* was statistically significant for girls alone, $t(26) = 3.98$, $p < .0001$, $d = 0.77$, but not for boys alone, $t(28) = 1.12$, $p = .27$, $d = 0.21$. The 2 $\times$ 2 ANOVA also revealed a main effect of order, indicating that children demonstrated

Table 1
Mean response times (in milliseconds) of male and female participants in PSIAT testing flower–insect attitudes in Experiment 1 and gender attitudes in Experiment 2.

	Combined task	Girls	Boys
Experiment 1	Flowers/good + insects/bad	1529 (278)	1714 (386)
	Flowers/bad + insects/good	1812 (361)	1803 (432)
Experiment 2	Girls/good + boys/bad	2048 (460)	2210 (511)
	Girls/bad + boys/good	2512 (550)	2131 (416)

Note: Standard deviations are in parentheses.

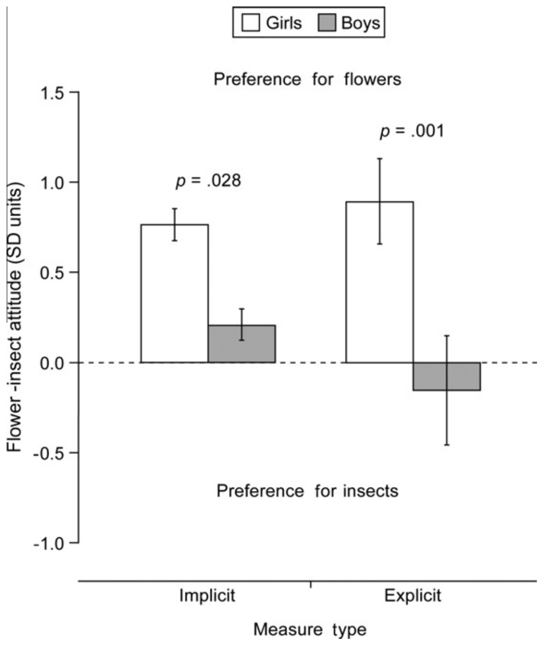

Fig. 2. Sex differences (in standard deviation [*SD*] units) on implicit and explicit measures of flower–insect attitudes of 4-year-olds. $N = 56$ (27 girls and 29 boys) for the implicit measure, and $N = 53$ (28 girls and 25 boys) for the explicit measure. Error bars represent standard errors. The *p* values indicate significance of the difference between girls and boys.

stronger positive attitudes toward flowers when they completed the flowers/good + insects/bad block first, $F(1, 52) = 5.84$, $p < .001$, $d = 0.67$. Further follow-up analyses of this order effect using one-sample *t* tests indicated that the significant positive attitude toward flowers was demonstrated when *flowers/good* was the first combined task, $t(28) = 3.90$, $p = .001$, $d = 0.72$, but not when *flowers/bad* was the first combined task, $t(26) = 0.89$, $p = .38$, $d = 0.17$. No significant gender-by-order interaction was observed ($p > .16$).

The explicit measure was entered as the dependent variable in a one-way ANOVA with gender as a between-subjects factor. On the explicit measure, there was evidence for positive attitudes toward flowers, $F(1, 49) = 5.34$, $p = .03$, $d = 0.31$. The analysis also revealed a main effect of gender, indicating that girls self-reported stronger positive attitudes toward flowers than did boys, $F(1, 49) = 12.34$, $p = .001$, $d = 0.49$. The correlation between implicit and explicit measures of flower–insect attitudes was $r = .22$, $p = .11$ (more detailed analyses examining the relationship between implicit and explicit measures can be found in the supplementary material).

Cronbach's alpha for the implicit measure, calculated from two *D* measures computed from word and picture trials of each PSIAT, was $\alpha = .88$. Cronbach's alpha for the explicit measure, calculated from the two questions of the flower–insect attitude measure, was only $\alpha = .53$, leading to the development of a different approach to the explicit measure in Experiment 2.

Experiment 1 established the PSIAT method by showing that (a) the 4-year-olds were able to complete the PSIAT and that (b) the current adaptation of the IAT was sensitive to evaluative preferences for flowers over insects. These findings are in line with similar results obtained with 6-year-olds (Baron & Banaji, 2006). In the current study, the implicit preference for flowers was stronger in girls than in boys—an effect that was also reported by Baron and Banaji (2006) for 6-year-olds, although it did not reach statistical significance in that study. In addition, the overall effect size (i.e., Cohen's *d*) for the flower–insect attitude measure was $d = 0.50$, which is comparable to the findings with 6-year-olds by Baron and Banaji, who reported an overall effect size of $d = .55$.

Experiment 2: Gender attitudes

Experiment 2 examined (a) the PSIAT's sensitivity to individual differences in gender attitudes, (b) the relationship of children's gender attitudes to children's gender, and (c) the predictive validity of two gender attitude measures (implicit and explicit) with children's engagement in gendered activities. The Preschool Activities Inventory (Golombok & Rust, 1993) measures gender-typical play activities in young children, according to parental reports, and is sensitive to within-gender variation. In addition, the Preschool Activities Inventory exhibits high split-half reliability (ranging from .73 to .80 for children between 30 and 57 months of age) and high test–retest reliability (.66 for girls and .62 for boys) over a 1-year period (Golombok et al., 2008).

Method

Participants

The participants were 75 typically developing 4-year-olds (M = 53.5 months, SD = 3.2, 38 girls and 37 boys). The same method of recruitment and remuneration as in Experiment 1 was used. According to parental reports, the racial/ethnic composition of the sample was 85.4% White, 2.7% African American, 1.3%, Asian, 1.3% Native American, and 9.3% multiracial, with 2.2% of the participants being of Hispanic ethnicity.

Procedure

The design, manipulations, and procedure were the same as in Experiment 1 except for the following differences, which were instituted to investigate gender attitudes.

PSIAT measure. During the gender attitude PSIAT, children sorted stimuli representing four concepts: *boy, girl, good,* and *bad.* The same *good* and *bad* words from Experiment 1 were used in Experiment 2. Four pictures of boys and four pictures of girls were used to represent the *boy* and *girl* concepts, respectively. In one instructional condition, *good* words and *boy* pictures shared a response key and *bad* words and *girl* pictures shared the other response key. The other instructional condition reversed the response assignments for *boy–girl* pictures. Positive scores indicated stronger association of *good* with *girl* and of *boy* with *bad* than vice versa.

Using the same exclusion criteria as in Experiment 1, additional PSIAT data were dropped from the analyses due to excessively fast or slow responding or excessive errors (7 participants [9.3%]), and 4 additional children (5.3%) did not finish their PSIAT, leaving N = 64 children for analysis. The implicit measure of gender attitudes was administered before the explicit measure.

Explicit measure. Given the poor internal consistency of the two-item explicit measure in Experiment 1, a new explicit measure was developed for Experiment 2 based on a format of similar measures used with 6-year-olds (Baron & Banaji, 2006; Dunham et al., 2006). For the explicit measure, children's task was to "tell the computer which picture you like more." This was done by pressing a button of the same color on the adapted keyboard (Fig. 1). For example, if children liked the picture of a girl and that picture was presented "on the green side," children would indicate their liking for that picture by pushing the green button. Unlike in the PSIAT, the explicit task was not a response latency measure and children were not urged to respond as quickly as possible.

On each trial, children were shown two face drawings, presented side by side, and asked to indicate which drawing they liked more. Eight face drawings were used (four boys and four girls) for a total of 16 trials. These 16 trials reflect all possible male–female pairings for a total of 16 unique face pairings that were randomized in their presentation (for a complete list of all face drawing stimuli, see supplementary material). Half of the computer screen had a green background, and the other half had an orange background. On each side of the screen, a drawing was centered against the colored background so that one drawing of a boy and one drawing of a girl appeared on each trial. Children were instructed to take their time and think carefully about their preferences. The measure was scored as a proportion of responses (out of 16) in which children selected the girl character. A value of .50 was subtracted from each child's score. That way, rational zero points were retained for the explicit measure of gender attitudes; positive values indicated liking for a girl character, and negative values indicated liking for a boy character (with 0 indicating equal liking for boy and girl characters).

Gendered activities. The Preschool Activities Inventory (Golombok & Rust, 1993) was used to assess children's gendered activity. The inventory is a 24-item survey measuring three aspects of children's everyday behaviors. The measure was completed by parents, who indicated how frequently during the past month their children played with particular toys (e.g., dolls), engaged in particular activities (e.g., playing house), or showed particular characteristics (e.g., "likes pretty things"). This was done by choosing one of the five options on a bipolar scale where one end corresponded to *never* and the other end corresponded to *very often.* Three quarters (75%) of parents ($N = 48$) chose to complete the survey. The measure was scored by summing the *female* items and then subtracting the *male* items. That way, positive scores indicated more feminine behavior and negative scores indicated more masculine behavior (with zero indicating equally masculine and feminine behavior).

Results and discussion

Gender attitudes: Implicit and explicit measures

Fig. 3 displays individual scores expressed in the same standard deviation units for each of the three measures (raw means and standard deviations for all measures used in Experiments 1 and 2 can be found in the supplementary material). The implicit measure was entered as the dependent variable in a 2 (Gender: male or female) × 2 (Order of Combined Tasks: girls/good first or boys/good first)

Note: This version of Figure 3 is a replacement of the original, correcting data for the implicit gender attitude measure.

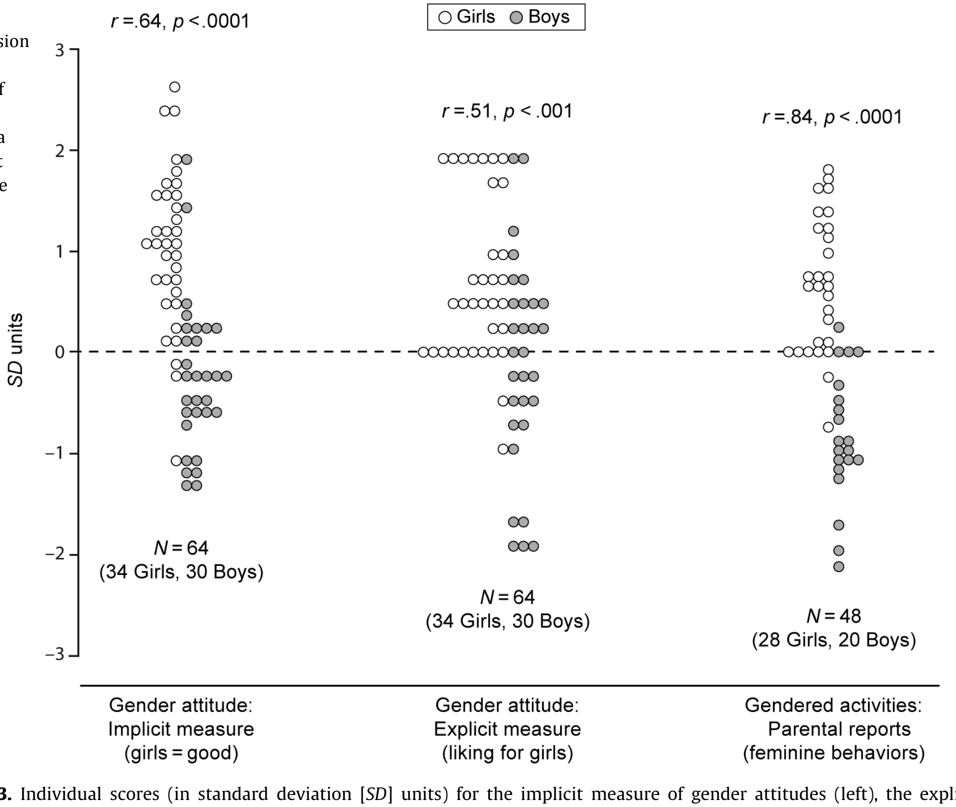

Fig. 3. Individual scores (in standard deviation [*SD*] units) for the implicit measure of gender attitudes (left), the explicit measure of gender attitudes (middle), and a measure of gendered activities (right) of 4-year-olds. Coefficients at the top of score distributions (*r*s) are effect sizes for correlations with children's actual sex. For each of the measures, the labels on the *x* axis indicate (in parentheses) the direction associated with positive values.

ANOVA to examine differences by gender and order of combined tasks. On the implicit measure, there was strong evidence for positive attitudes toward girls, as indicated by a statistically significant intercept, $F(1, 60) = 15.24$, $p < .0001$, $d = 0.49$. This analysis also revealed the expected main effect of gender, with girls demonstrating a stronger association of *girls* with *good* (relative to *boys* with *good*) than did boys, $F(1, 60) = 43.48$, $p < .0001$, $d = 1.70$. There was no main effect of order nor was there a gender-by-order interaction (*p*s > .35).

The explicit measure was entered as the dependent variable in a one-way ANOVA with gender as a between-subjects factor. On the explicit measure, there was a significant effect showing an overall explicit positive attitude toward girls, as indicated by a statistically significant intercept, $F(1, 62) = 13.91$, $p < .0001$, $d = 0.43$. The one-way ANOVA of the explicit measure also revealed the expected main effect of gender; given a choice between two pictures, one of a girl and the other of a boy, girls were more likely to report liking the girl picture than were boys, $F(1, 62) = 21.93$, $p < .001$, $d = 0.66$.

Both implicit and explicit gender attitude measures had rational zero points, indicating the absence of a preference for either gender (i.e., an equally positive attitude toward boys and girls). On the implicit measure, two separate one-way ANOVAs, one for girls and one for boys (each with the order of combined tasks as a between-subjects factor), indicated that preference for own gender was statistically significant for girls, $F(1, 32) = 53.88$, $p < .0001$, $d = 1.26$, but not for boys, $F(1, 28) = 3.80$, $p = .061$, $d = 0.36$. Similarly, on the explicit measure, two separate one-way ANOVAs, one for girls and one for boys, indicated that preference for own gender reached statistical significance for girls, $F(1, 33) = 56.70$, $p < .0001$, $d = 0.94$, but not for boys, $F(1, 19) = 0.31$, $p = .58$, $d = -0.07$. These effects are in line with similar work on attitudes in adults, as discussed below.

Implicit and explicit measures of gender attitudes were positively correlated ($r = .43$, $p = .0003$). The correlations with children's actual sex were strong for both the implicit ($r = .64$, $p < .0001$) and explicit ($r = .51$, $p < .0001$) measures.

Cronbach's alpha for the implicit measure, calculated in the same fashion as for Experiment 1, was $\alpha = .85$. Cronbach's alpha for the explicit measure, calculated from matched 8-trial subsets of the gender attitude measure, was $\alpha = .88$.

Gendered activities

As expected, the Preschool Activities Inventory differentiated between boys and girls in terms of their play activities, with girls engaging in more feminine behaviors ($M = 12.64$, $SD = 9.24$) than boys ($M = -15.35$, $SD = 8.72$), $t(46) = 10.58$, $p < .0001$, $d = 3.16$. Cronbach's alphas for the gender activities measure, calculated separately from the 12 male and 12 female items, were $\alpha = .84$ and $\alpha = .90$, respectively. The data from the Preschool Activities Inventory showed strong correlations with both the implicit ($r = .52$, $p = .0001$) and explicit ($r = .50$, $p = .0003$) gender attitude measures.

Relation between implicit and explicit gender attitudes and gendered activities

To examine how well the implicit and explicit attitude measures predicted children's gendered activities, two hierarchical regressions were conducted. In both analyses, the Preschool Activities Inventory score was entered as a criterion. In one regression model, the PSIAT score was entered at Step 1 and the explicit gender attitude score was entered at Step 2. In the other regression model, the order of the two predictors was reversed. The zero-order correlations were significant at Step 1 for both the implicit predictor ($R^2 = .28$, $\beta = .53$, $p < .0001$) and the explicit predictor ($R^2 = .26$, $\beta = .51$, $p < .0001$). The effect of each predictor also remained significant in a simultaneous regression at Step 2 for both the implicit measure ($\Delta R^2 = .12$), $F(1, 45) = 8.91$, $p = .005$, and the explicit measure ($\Delta R^2 = .10$), $F(1, 45) = 7.40$, $p = .009$, showing that the two measures provided incremental prediction. Taken together, the results indicated that the two measures were not just duplicates of one another; each measure added value at Step 2 beyond the value already demonstrated by the other measure at Step 1.

General discussion

The current experiments report the first investigations of the PSIAT, which is an adaptation of the IAT so that it can be used with 4-year-olds. Evidence of the new procedure's value derives from its showing: (a) expected preference for flowers (relative to insects) on IAT measures previously used with older children and adults, (b) sensitivity to sex differences in children's implicit attitudes, (c) predictive validity in correlations with parental reports of gender-typed behavior, and (d) discriminant validity in correlational analyses that included both PSIAT and self-report measures.

Evidence for validity of PSIAT

Expected implicit attitude patterns

On the flower–insect PSIAT of Experiment 1, adult data reveal implicit preference for flowers relative to insects (Greenwald et al., 1998). The flower–insect IAT has often been used as an initial test of novel forms of the IAT because the ability to detect a preference for flowers (relative to insects) indicates that the measure is working. In Experiment 1, this effect was replicated in preschool children and the overall effect size was $d = 0.50$, which corresponds to a conventional medium effect size for Cohen's *d*.

Sensitivity to sex differences

Both of the experiments reported here found differences between boys' and girls' implicit attitudes. In Experiment 1, the implicit preference for flowers was stronger in girls than in boys. Although this had not been predicted, it was not surprising. Little boys, who are presumed—in the widely known nursery rhyme—to be made of frogs, snails, and puppy dogs' tails, might be expected to be fonder of insects than girls are. Just this sex difference was reported by Baron and Banaji (2006) for 6-year-olds, although they did not find it to be statistically significant. Baron and Banaji did, however, find that girls' greater liking for flowers was statistically significant on their self-report attitude measure. Experiment 1 confirmed Baron and Banaji's finding of girls' significantly greater self-reported preference for flowers relative to insects. Experiment 1's PSIAT also found statistical significance for girls' greater preference for flowers, bolstering indications of the PSIAT's usefulness with young children.

For both its PSIAT and self-report measures of gender attitudes, Experiment 2 found a pattern of sex differences that was expected from previous findings with adults and older children. Using adult IAT measures, Rudman and Goodwin (2004) reported that women were significantly more positive toward their own gender than were men. Ebert and Steffens (2008) found that this effect already occurs in elementary school children, suggesting that it is not confined to adults. Eagly and Mladinic (1989) previously reported this same pattern for self-report measures. Experiment 2's confirmation of this sex difference pattern on both types of measures indicated not only that the PSIAT was capable of revealing the pattern but also that the pattern is evident at a younger age than was previously established.

This finding informs our theories for understanding the origins and development of implicit gender attitudes. It has been suggested that the strong pro-female attitudes among adult females and the lack of pro-male attitudes among adult males may be due to multiple influences (Skowronski & Lawrence, 2001), with one of these influences being that positive attitudes toward one's mother generalize to positive attitudes toward women in general (*female* = *good* even if one's own sex is male). The second influence may be in-group bias; people generally prefer groups that they are a part of over groups to which they do not belong (Heyman & Legare, 2004). For females, both of these influences work in the same direction, resulting in strong pro-female attitudes. For males, the two influences work against each other. This opens a developmental line of research because it suggests that attachment to one's mother during early childhood may moderate the relationship between in-group bias and out-group bias in young boys. Using the PSIAT technique developed here, we plan to explore the developmental origins of in-group biases for gender and how they may relate to attitudes toward one's own mother, other significant female caretakers, and other social developmental issues.

Given the test content and the results of past experiments, girls were expected to have more positive attitudes toward flowers (Experiment 1) and more positive attitudes toward girls (Experiment 2) than were boys of the same age. However, another possibility was that girls would show more positive attitudes on an IAT across the board regardless of test content (perhaps due, e.g., to executive function differences that could be correlated with IAT performance). This will need to be tested in future studies, but we do not favor the latter alternative; past literature has not shown strong gender differences at this early age on executive function tasks such as working memory, inhibition, and cognitive flexibility (e.g., Cragg & Nation, 2009; Davidson, Amso, Anderson, & Diamond, 2006).

Predictive validity

The availability of parental reports of gender-typed activities in Experiment 2 allowed a test of the predictive validity of both the PSIAT and self-report measures of gender attitudes. Each type of measure was an effective predictor of gender-typed activities ($r = .53$ for the PSIAT, $r = .51$ for self-report). In addition, each type of measure predicted variance in gender-typed activities independently of the other type, as shown by the hierarchical regression findings. Although there exist a few prior studies that established predictive validity of IAT measures for children age 10 years and above (Pieters, van der Vorst, Engels, & Wiers, 2010; van Goethem, Scholte, & Wiers, 2010), the current Experiment 2 provides the first demonstration of predictive validity of an IAT measure involving naturally occurring social groups in children as young as 4½ years old (see Dunham, Baron, & Carey, in press, for a similar demonstration with 5½-year-olds involving minimal groups).

Discriminant validity

When examined only in terms of mean dependent measure values for boys and girls, the data for both experiments were quite similar. This might suggest that the PSIAT and self-report were interchangeable measures that were redundant with one another. However, other evidence established that they are distinct measures. One indication was in the correlations between parallel PSIAT and self-report measures. In Experiment 1, the PSIAT and self-report were nonsignificantly correlated ($r = .22$, a low correlation that may be due, in part, to the poor internal consistency of the explicit measure). In Experiment 2, although the two measures were significantly correlated ($r = .43$), this correlation may be explained by both of them tapping a sex difference in gender attitudes. This significant implicit–explicit correlation could also be seen as an interesting finding insofar as impression management processes might influence some explicit responses to a lesser degree at this young age than in adults, thereby leading to a greater congruence between implicit and explicit gender attitude measures in children than is found with gender attitude measures in adults. Additional important evidence for discriminant validity was provided by the mutually incremental predictive validity pattern found in the hierarchical regression analysis of Experiment 2.

Psychometrics

Internal consistencies of the PSIAT measures in both experiments were obtained by correlating IAT measures computed from odd-numbered versus even-numbered trials in each experiment. These internal consistencies revealed satisfactorily high values of Cronbach's alpha, namely, $\alpha = .88$ for Experiment 1 and $\alpha = .85$ for Experiment 2. Another useful psychometric property of IAT measures is their having a rational zero point, which is a consequence of the method's design (Greenwald, Rudman, Nosek, & Zayas, 2006). For the flower–insect PSIAT, the value of zero indicates equal strengths of the associations measured by the two combined tasks; stated differently, it indicates equal liking for flowers and insects. For the gender attitude PSIAT, the value of zero indicates equal liking for girls and boys. This rational zero point property is useful in permitting tests of cognitive consistency within sets of associations, as was done in the study of math gender stereotypes of young children by Cvencek et al. (in press).

Internal validity

Studies of IAT measures with adults often reveal an undesired effect of the order in which the two combined tasks are administered. Expectation of these occasional order effects has prompted a standard recommendation to counterbalance the order of combined tasks when using IAT measures (Nosek et al., 2007). Such counterbalancing was used in the current research, which found an order effect for the flower–insect PSIAT in Experiment 1 but not for the gender attitude PSIAT in Experiment 2.

When an order effect is obtained, a plausible interpretation is that performance of the first combined task can strengthen associations measured in that task. This, in turn, will boost the IAT-measured strength of the associations that are assessed in the first combined task. This order effect may be largest when the associations being measured by the IAT or PSIAT are initially weak, as may have been the case for the valence associations with flowers and insects in the current Experiment 1. By comparison, the lack of an order effect in Experiment 2 suggests that the gender–valence associations assessed in that experiment were more strongly established. In future work with young children, researchers may want to increase the number of practice trials on the switch category following the first combined task. This has reduced order effects considerably in adults (Greenwald et al., 2003) and may do so in children as well.

Comparison of PSIAT method with other child IAT methods

The adult IAT is a useful tool in social psychology studies of attitudes (e.g., Greenwald, Poehlman, Uhlmann, & Banaji, 2009). A child IAT was successfully developed for elementary school children, but methodological problems prevented it from being used with preschoolers despite considerable effort by interested researchers (Baron & Banaji, 2006; Cvencek et al., in press; Dunham et al., 2006; Rutland et al., 2005; Skowronski & Lawrence, 2001). Three procedural refinements in the current work helped to overcome the obstacles encountered by previous researchers: (a) presentation of all stimuli in two modalities (i.e., simultaneous audio and visual presentations), (b) visual reminders that use multiple stimuli to represent a particular concept (see Fig. 1), and (c) a 20% decrease in trial number from the standard IAT presentation of 180 trials (see Method in Experiment 1). In the following paragraphs, we discuss each of these improvements in turn.

Our first adaptation involved multimodal presentation of stimuli. In previous work, when stimulus words and pictures were presented unimodally on alternating trials to children (i.e., audio presentation of words and visual presentation of pictures), children's response speed to picture trials differed significantly from their response speed to word trials. To permit computation of one IAT score from two distributions of response latencies, resulting child IAT measures often required computation and averaging of two effect sizes: one for word stimuli and one for picture stimuli (e.g., Baron & Banaji, 2006; Dunham et al., 2006). The use of multimodal stimulus presentation (i.e., visual and auditory presentations on each trial) avoids this added statistical complexity; high Cronbach's alphas for the PSIAT measures indicated that response latencies for pictures and words were highly consistent with one another.

The second adaptation involved more detailed visual reminders for children than the ones used in previous research. The use of multiple stimuli to represent a concept (e.g., four pictures of flowers as a reminder for the *flowers* concept [see Fig. 1]) likely reduces cognitive task demands associated with retaining multiple stimuli in working memory while sorting four concepts using two response buttons, as required by the IAT. We suspect that such cognitive overload might have contributed to past difficulties.

The third adaptation involved a shorter protocol. To our knowledge, past research with children has always used a standard adult IAT length. The shorter number of trials used here facilitates sustaining children's attention during the task while still yielding effect sizes comparable to those of standard-length child IAT measures (Baron & Banaji, 2006).

A primary focus of this article was to establish an IAT method that could work with prekindergarten children, and this has been established. Further work will be needed to see the lower age boundary at which these tests can be administered and yield valid data, a topic that interests the current authors.

Future usefulness of PSIAT

A substantial challenge in creating an IAT measure for children was to adapt the instructions and response tasks to children's limited abilities to read, to comprehend abstractions, and to use computer keyboards. The multiple indications that the PSIAT provided useful findings in the current Experiments 1 and 2 show that the PSIAT was, indeed, successfully adapted to the limited abilities of children of preschool age. Success of the adaptation was further indicated by the low participant loss rate in the current two experiments—14% overall, which is barely larger than the 5–10% loss rates typical of IAT experiments with adult participants.

There are two primary explanations for the IAT's apparent ability to access social cognitive processes that are not revealed by self-report measures in adults (Greenwald et al., 2002). First, lack of introspective access, perhaps due to the knowledge being acquired from long-forgotten experiences, may make the knowledge unavailable to self-report. Second, self-report measures may be distorted by self-presentational strategies such as trying to make a good impression on others. Such impression management is especially likely in socially sensitive domains such as in studies of attitudes and stereotypes involving race, gender, age, religion, ethnicity, and disability.

Children's self-reports are likewise potentially subject to lack of introspective access and distortion by impression management (Brody et al., 1985; Richman, Berry, Bittle, & Himan, 1988), and these social desirability concerns become stronger during childhood and adolescence (Heyman & Legare, 2005; Rutland et al., 2005). The current experiments' indications of success of the PSIAT suggests that there is now available a method that can be extended to measure otherwise inaccessible social cognitive knowledge of preschoolers. As was true of the IAT with adults, the PSIAT is straightforwardly adaptable to measure a wide variety of valence associations and may eventually be adapted to measure stereotypes and other categories of associative knowledge in preschoolers.

A specific likely future use of the PSIAT is suggested by findings of a recent meta-analysis of the IAT's predictive validity (Greenwald et al., 2009). The meta-analysis produced an unexpected finding that implicit measures of prejudices and stereotypes predict racially discriminatory behavior more strongly than do parallel self-report measures. That pattern of findings focuses importance on the task of identifying the most important developmental experiences that shape valence and other associations with racial categories. The availability of the PSIAT will make it possible to conduct studies with preschoolers that can start to tease apart the relative influences of parents, peers, and mass media, to list just the most likely formative experiences.

Acknowledgment

This research was supported by a grant from the National Science Foundation (OMA-0835854) to the LIFE Science of Learning Center. The University of Washington is the owner of the intellectual property in the PSIAT. The University has applied for a patent and manages all intellectual property rights for the PSIAT. For information on licensing the PSIAT, please contact University of Washington, Center for Commercialization – Technology Licensing (E-mail:license@u.washington.edu).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jecp.2010.11.002.

References

Baron, A. S., & Banaji, M. R. (2006). The development of implicit attitudes: Evidence of race evaluations from ages 6 and 10 and adulthood. *Psychological Science, 17*, 53–58.

Brody, L. R., Rozek, M. K., & Muten, E. O. (1985). Age, sex, and individual differences in children's defensive styles. *Journal of Clinical Child Psychology, 14*, 132–138.

Castelli, L., Zogmaister, C., & Tomelleri, S. (2009). The transmission of racial attitudes within the family. *Developmental Psychology, 45*, 586–591.

Cragg, L., & Nation, K. (2009). Shifting development in mid-childhood: The influence of between-task interference. *Developmental Psychology, 45*, 1465–1479.

Cvencek, D., Meltzoff, A. N., & Greenwald, A. G. (in press). Math-gender stereotypes in elementary school children. *Child Development*.

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia, 44*, 2037–2078.

Degner, J., & Wentura, D. (2010). Automatic prejudice in childhood and early adolescence. *Journal of Personality and Social Psychology, 98*, 356–374.

Dunham, Y., Baron, A. S., & Banaji, M. R. (2006). From American city to Japanese village: A cross-cultural investigation of implicit race attitudes. *Child Development, 77*, 1268–1281.

Dunham, Y., Baron, A. S., & Carey, S. (in press). Consequences of "minimal" group affiliations in children. *Child Development*.

Eagly, A. H., & Mladinic, A. (1989). Gender stereotypes and attitudes toward women and men. *Personality and Social Psychology Bulletin, 15*, 543–558.

Ebert, I. D., & Steffens, M. C. (2008). Generality and correlates of the implicit preference for women. In *Paper presented at the 11th Jena workshop on intergroup processes: Positive behaviour between social groups—processes, problems, and promises*. Germany: Oppurg.

Golombok, S., & Rust, J. (1993). The Pre-School Activities Inventory: A standardized assessment of gender role in children. *Psychological Assessment, 5*, 131–136.

Golombok, S., Rust, J., Zervoulis, K., Croudace, T., Golding, J., & Hines, M. (2008). Developmental trajectories of sex-typed behavior in boys and girls: A longitudinal general population study of children aged 2.5–8 years. *Child Development, 79*, 1583–1593.

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109*, 3–25.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74*, 1464–1480.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85*, 197–216.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology, 97*, 17–41.

Greenwald, A. G., Rudman, L. A., Nosek, B. A., & Zayas, V. (2006). Why so little faith? A reply to Blanton and Jaccard's (2006) skeptical view of testing pure multiplicative theories. *Psychological Review, 113*, 170–180.

Harter, S., & Pike, R. (1984). The pictorial scale of perceived competence and social acceptance for young children. *Child Development, 55*, 1969–1982.

Heyman, G. D., & Legare, C. H. (2004). Children's beliefs about gender differences in the academic and social domains. *Sex Roles, 50*, 227–239.

Heyman, G. D., & Legare, C. H. (2005). Children's evaluation of sources of information about traits. *Developmental Psychology, 41*, 636–647.

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin, 31*, 1369–1385.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2005). Understanding and using the Implicit Association Test: II. *Method variables and construct validity. Personality and Social Psychology Bulletin, 31*, 166–180.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Automatic processes in social thinking and behavior* (pp. 265–292). New York: Psychology Press.

Olson, K. R., & Dunham, Y. (2010). The development of implicit social cognition. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 241–254). New York: Guilford.

Pieters, S., van der Vorst, H., Engels, R. C. M. E., & Wiers, R. W. (2010). Implicit and explicit cognitions related to alcohol use in children. *Addictive Behaviors, 35*, 471–478.

Richman, C. L., Berry, C., Bittle, M., & Himan, K. (1988). Factors related to helping behavior in preschool-age children. *Journal of Applied Developmental Psychology, 9*, 151–165.

Rudman, L. A., & Goodwin, S. A. (2004). Gender differences in automatic in-group bias: Why do women like women more than men like men? *Journal of Personality and Social Psychology, 87*, 494–509.

Rutland, A., Cameron, L., Milne, A., & McGeorge, P. (2005). Social norms and self-presentation: Children's implicit and explicit intergroup attitudes. *Child Development, 76*, 451–466.

Skowronski, J. J., & Lawrence, M. A. (2001). A comparative study of the implicit and explicit gender attitudes of children and college students. *Psychology of Women Quarterly, 25*, 155–165.

Steffens, M. C., Jelenec, P., & Noack, P. (2010). On the leaky math pipeline: Comparing implicit math-gender stereotypes and math withdrawal in female and male children and adolescents. *Journal of Educational Psychology, 102*, 947–963.

Thomas, S., Burton Smith, R., & Ball, P. (2007). Implicit attitudes in very young children: An adaptation of the IAT. *Current Research in Social Psychology, 13*, 75–85.

van Goethem, A. A. J., Scholte, R. H. J., & Wiers, R. W. (2010). Explicit and implicit bullying attitudes in relation to bullying behavior. *Journal of Abnormal Child Psychology, 38*, 829–842.