Available online at www.sciencedirect.com
SCIENCE DIRECT®
Cognitive Development 19 (2004) 223–240
COGNITIVE DEVELOPMENT

Observers' proficiency at identifying pretense acts based on behavioral cues

Rebekah A. Richert, Angeline S. Lillard*

Department of Psychology, University of Virginia, 102 Gilmer Hall, P.O. Box 400400, Charlottesville, VA, 22904-4400, USA

Received 30 June 2002; received in revised form 20 December 2003; accepted 26 January 2004

Abstract

Discriminating what is pretense from what is real is a fundamental problem in development. Research has addressed the proficiency with which adults and children discriminate between play fighting and real fighting, and yet none (to our knowledge) has investigated discrimination of other kinds of pretense and real acts. In addition, little is known about what aspects of pretender behavior (as opposed to pretend content) might cue pretense interpretations. In two experiments, 8–20 s clips showing pretense and real snack behaviors were presented to adult and child participants. All participants distinguished between pretense and real behaviors at better than chance level. Furthermore, certain features (specific looking patterns and mistimed behaviors) were most prominent in the videotapes that were most often correctly identified. This provides empirical support for the suggestion that these cues, as opposed to more commonly cited cues, like smiles, might serve as important indicators of pretense for children and adults.

Keywords: Pretense; Discrimination; Development; Social cognition; Behavioral cues

Imagine watching the following scenario unfold for the first time. An adult woman picks up a banana. Rather than peeling and eating it, she holds one end to her ear and begins laughing and speaking into the other end. She pauses and nods occasionally, as if in conversation. The only other time you have seen a woman perform similar types of behaviors is with a telephone. Seen for the first time, these bizarre actions with a banana could be interpreted a number of different ways: the woman might really think the banana is speaking to her, the woman might actually think the banana is a telephone, or the woman

* Corresponding author. Fax: +1-434-982-4694.
E-mail address: lillard@virginia.edu (A.S. Lillard).

0885-2014/$ – see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.cogdev.2004.01.001

might simply be pretending the banana is a telephone. For adults, the first case could indicate schizophrenia, and the second, severe confusion. Only the third would be an appropriate interpretation. For a very young child, however, to whom bananas and telephones are novel entities, the first two interpretations might make sense. Yet, in taking the adult's action seriously, the child's interpretation of the real world would become jumbled: bananas do not talk, and they are not telephones. Being able to make a pretense interpretation of pretense events is crucial to understanding the real world.

American parents begin pretending with their infants within the first year (Haight & Miller, 1993; Kavanaugh, Whittington, & Cerbone, 1983; Tamis-LeMonda & Bornstein, 1991). By 2 years old, infants have been shown to participate in tightly controlled, experimenter-provided pretense scenarios, indicating that they understand pretense (e.g., Harris & Kavanaugh, 1993). In particular, it has been argued that pretending 2-year-olds understand four features of pretense: pretend stipulations, causal powers, the suspension of objective truth, and an unfolding causal chain (Harris, 2000). This ability seems surprising if considered in the context of the onslaught of new information to which infants are continuously exposed. It has been suggested that even the real actions of people must be to infants much like a foreign language is to an adult: the individual words are hard to distinguish and it is difficult to make sense of the babble (Brand, Baldwin, & Ashburn, 2001). Yet, despite this large quantity of new information, infants appear to learn to make sense of the "real" actions they observe. For example, by the end of their first year, infants demonstrate the ability to "parse" action streams into intentional units (Baldwin & Baird, 1999).

Pretense behaviors, however, serve a fundamentally different purpose than real behaviors. Pretense is not intended to be taken literally, and in fact involves the *distortion* of reality. Thus, the introduction of pretense behaviors at the end of the first year could easily confuse infants' newly developing expectations about real behaviors. When mother pretend-eats off the spoon, she does not put it in her mouth. How is the infant to make sense of that action stream? Do infants even make sense of such acts?

There is surprisingly very little research on how proficiently even adults and older children identify pretense acts. When they are proficient, we know little about the basis on which they make their judgments. One possibility is that pretense judgments are based on content cues. However, content cannot always serve to assist pretense interpretation, especially for young children (Lillard & Witherington, 2004). The focus of the present research is on potential *behavioral* cues to pretense, cues that could be useful even when content is not.

Although there is little research on how proficiently people can identify pretending, there is some literature on the similar behavior of play fighting. Among young mammals, play fighting is common, and the inability to discriminate pretend from real fighting would presumably be disastrous for a species.

Specific behavioral alterations appear to cue other mammals that one is playing. For example, *canids* appear to "play bow" to invite another to engage in social play (Bekoff, 1977; Bekoff & Allen, 1998). This "play bow" might occur either before an attack or after accidentally hurting the play partner. The cues to play fighting are not limited to the "play bow" before or after the attack, but exist in the course of the play fighting as well. A study with Eastern coyotes suggested that in play, bites toward the tail, flank, legs, abdomen, or back lasted a significantly shorter time and were more stereotyped than in real fighting scenarios (Hill & Bekoff, 1977). Such behavioral modifications might communicate to play

partners that a fight is not real. A well-documented characteristic of mother–infant play in primates is the mother's tendency to self-handicap or inhibit her behavior (e.g., Biben & Suomi, 1993; Fagen, 1981). This means mothers will put themselves at a "disadvantage" when play fighting with younger animals by allowing the younger one to dominate at times and by not completing harmful actions on the youngster. It is likely that these mothers are holding back so as not to hurt their offspring. In the process, they might also communicate that the fight is simply play so the youngster does not hurt the mother either. Squirrel monkeys display three specific behaviors that lead to play: leaping toward a play partner, rolling onto one's back, and performing "acrobatic antics" while hanging from a perch (Biben & Suomi, 1993).

In sum, among animals, specific behavior patterns, behavioral initiations, and truncated behaviors might serve to indicate to others that one is play fighting, not fighting for real. Clearly, the fact that animals engage in play fighting suggests there are nonverbal, behavioral means by which the message "this is play" is delivered (Bateson, 1972).

Behavioral cues of play fighting have been documented in human children, as well. For example, rough and tumble play is characterized by positive and neutral facial expressions (Boulton, 1991) as well as by using less force and deliberately not making contact with the play partner (Smith & Boulton, 1990). Furthermore, the likelihood that children and teachers can tell the difference between videotaped real and play fighting sequences is associated with the presence of these modifications in physical actions and facial expressions (Boulton, 1993). Thus, an ability to distinguish between real fighting and play fighting has been documented in children and adults.

Children, however, are not only faced with distinguishing between pretend and real fighting scenarios. Parents pretend a variety of scenarios with their children that may embody behavioral changes as well. Changes in behavioral patterns in pretense may be especially crucial for infants' correct interpretation of actions, given the immaturity of their language. It has been suggested that mothers may scaffold their infants within particular play frames characterized by exaggerated contours, marked changes of tempo, and systematic repetitions (Bruner, 1979; Gergely & Watson, 1999; Kaye, 1982). Sound effects also appear to be very common in pretense (DeLoache & Plaetzer, 1985; Farver, 1992; Nicholich, 1977) and might serve as cues. Many researchers, most notably Piaget (1962), have suggested that "knowing" smiles are the best key to pretense (see also Wellman & Hickling, 1993). Despite theorists' interest in the non-linguistic aspects of pretend play (Bekoff & Byers, 1981; Ruesch & Bateson, 1951), little research has been devoted to the non-verbal aspects of pretense, outside of play fighting, that might serve as cues to the nature of the activity.

To address these issues, recent research in our laboratory has compared parallel situations in which mothers pretend to have a snack and really have a snack with their children (Lillard & Witherington, 2004). This research has revealed an array of notable differences in behaviors when mothers are pretending. First, mothers' looking patterns varied in pretend and real interactions. In pretense, mothers tended to spend proportionally more time looking at their children than at the task.[1] Second, mothers' movements, like grasping real

[1] Mothers' looking to the child or to the task were not necessarily reciprocal, since the mother could also have been looking around the room.

or imaginary food and bringing it to their mouths, were often faster in pretense. In addition, they held their hands at their mouths longer when fake-eating than when really eating. Pretend movements were thus mistimed in relation to real movements. Third, mothers made sound effects (saying "glug glug!" when drinking) frequently and almost exclusively during pretense. Fourth, smiling showed distinct patterns. Mothers smiled more in pretense interactions (although they frequently smiled in real interactions as well). More telling, perhaps, mothers smiled significantly more often in pretense scenarios with apparent reference to their own actions; in real situations, they rarely smiled about their own actions. In addition, smiles lasted longer, on average, in the pretense condition. These and several other significant differences were observed in mothers' behaviors in pretend and real snack interactions with their 18-month-old children. In other research we have found that these maternal behavioral changes are consistent with infants from 15 to 24 months (Lillard et al., 2004). Such behavioral differences might serve as cues that an action is pretense.

To determine whether these behavioral changes do in fact cue pretense, videotapes of mothers pretending were shown to observers to determine under what circumstances they could best distinguish pretense from real events. Although the eventual interest is in how babies make pretense interpretations, it seemed prudent to initially examine what cues are present when older children and adults most easily make pretense interpretations. Importantly, content cues were blocked during presentation of these scenarios. This enabled us to examine how pretender behaviors might serve pretense interpretation, even in the absence of obvious cues from content.

1. Experiment 1

In three phases, Experiment 1 examined pretense identification skills in adults and children of very brief segments of behavior. Previous research has indicated that people are quite accurate at judging other types of intentions besides pretense, particularly deceptive intentions, even when the judgments are based on brief videotaped observations, termed "thin splices" (Albright, Kenny, & Malloy, 1988; Funder & Colvin, 1988). Meta-analytic results reveal that using longer video clips for behavioral observation does not necessarily yield greater accuracy (Ambady & Rosenthal, 1992), and it has been noted that people often are more accurate when they rely on quick intuition rather than more considered introspection to judge another's behaviors (Ambady, Bernieri, & Richeson, 2000; Wilson & Schooler, 1991). Thus, shorter clips of behaviors often elicit more accurate judgments from observers.

Based on these findings, in the first phase, adults were shown short segments of pilot video clips for the Lillard and Witherington (2004) study reviewed above, in which mothers were either pretending to have a snack or really having a snack of Cheerios and water with their 18-month-old children.[2] The purpose of the second phase was to explore a possible developmental progression in children's ability to distinguish between pretense and real actions, therefore, children were shown the half of the video clips that were more easily identified by the adults. In the third phase, the relationship between adult participants'

[2] The actual studies used juice in a metal pitcher, but pilot work had used water in a glass pitcher.

responses and specific behavioral modifications in mothers' interactions were examined to determine whether participants use certain cues in identifying pretense actions.

1.1. Experiment 1a

1.1.1. Method

1.1.1.1. Participants. Participants for this experiment were 40 undergraduates ($M = 19$; range = 18–21 years) recruited from the Psychology Department participant pool at a major university. Participants received class credit for their participation.

1.1.1.2. Materials. Thirty-six video clips from Lillard and Witherington (2004) were chosen to show participants. Clips ranged in length from 8.2 to 12.76 s ($M = 10.41$ s, S.D. = 1.41). In half of these clips, mothers were pretending to have a snack with their 18-month-old children ($M = 10.54$ s; range = 9.11–12.76). In the other half of these clips, mothers were really having a snack with their 18-month-old children ($M = 10.28$ s; range = 8.26–12.75). The mean lengths of the clips in the two groups were not significantly different. Clips were chosen by the following criteria: (a) the word "pretend" was not used, and (b) clips showed a complete action. In addition, effort was made to reduce obvious contextual cues. To do this, a piece of paper was taped to the bottom of the television screen to cover the table, hiding whether Cheerios were actually present. In addition, because a glass pitcher revealed whether water was present, when the pitcher rose above the paper on the screen (e.g., if the mother was pouring from the pitcher), the pitcher was covered in black using video editing techniques.

1.1.1.3. Procedure. All participants were shown the 36 video clips in random order. To negate a possible training effect, half of the participants viewed the clips continuously from Clips 1 through 36. The other half first viewed Clips 19 through 36, and then Clips 1 through 18. Following each clip, participants were asked to rate the mother's actions on a 4-point scale: "real," "probably real," "probably pretend," and "pretend."

1.1.2. Results and discussion

For analysis, the 4-point scale was collapsed into a dichotomous scale, with "probably real" and "real" responses counted as "real" and "probably pretend" and "pretend" responses counted as "pretend." Clips were then coded as 1 for correct and 0 for incorrect. Two-tailed *t* tests compared the percentages of correct responding on each type of video clip to chance responding (50%). Participants responded at levels that were significantly better than chance for both the pretend ($t[39] = 7.39$, $P < 0.001$) and the real clips ($t[38] = 4.40$, $P < 0.001$). In addition, there was no effect of clip order nor any tendency to improve on the later clips viewed.

1.2. Experiment 1b

The results of Experiment 1a suggested that adults could identify whether or not someone is pretending, at a better than chance level and even from very short splices of behavior.

Adults, however, have far more experience than children do at observing and identifying behaviors. Experiment 1b was designed to explore whether *children* can identify when someone is pretending.

1.2.1. Method

1.2.1.1. Participants. Participants included seventeen 4-year-olds ($M = 4.4$; range = 3.3–5.5), seventeen 7-year-olds ($M = 7.0$; range = 6.1–7.7), and 28 college undergraduates ($M = 18$ years). Children were recruited from preschools and elementary schools in a small city. Undergraduates were recruited from the Psychology Department participant pool at a major university and received class credit for their participation.

1.2.1.2. Materials. Materials included two videotapes. One videotape was used as a training tape for the children and is described below. The second, stimulus videotape was used with both children and adults. It showed a subset of 18 of the clips used in Experiment 1a, because watching all 36 clips was taxing for children. The 18 clips (9 pretend and 9 real) that were easiest for the adults in Experiment 1a were chosen to show to children because the point of interest was whether children could also distinguish scenarios that we knew adults could distinguish well.

1.2.1.3. Procedure. A female experimenter visited local preschools and elementary schools and told the classes that she needed their help studying what children know about pretending. Permission slips were handed out to children to take home to their parents. Approximately a week later, the experimenter or a female undergraduate research assistant returned to the school and interviewed children whose parents had returned permission slips. Children were interviewed one at a time in a quiet room of the school. The children's procedure involved two stages: a training phase and an experimental phase.

The training phase was incorporated to establish that children could reliably identify the difference between pretend and real acts on video. Four training clips were chosen for each child from eight clips ($M = 8.80$ s, S.D. = 2.01) in which a woman was either pretending to or was really brushing her teeth, brushing her hair, drawing a house, or cutting a piece of paper. In counterbalanced order, children were shown four of these video clips, two portraying real acts and two portraying pretend acts. After viewing each clip, children were asked, for example, "Is she pretending to brush her teeth or really brushing her teeth?" If children responded incorrectly, the correct answer was explained. For example, "Look. Lori doesn't have a real toothbrush, so she is pretending to brush her teeth." Content was pointed out as a potential cue, so as not to lead children to notice a particular pretense behavior. All children responded correctly on at least three of the four training clips.

In the experimental phase, participants were shown the 18 most correctly identified video clips from Experiment 1a, showing 9 pretense (adult $M = 84\%$ correct) and 9 real ($M = 78\%$ correct) episodes. Since there was no practice effect in the Experiment 1a, these clips were presented in a single random order. For the children, following each clip, the experimenter asked the child, "Is the mom pretending to have a snack or really having a snack?" Question order was counterbalanced within and across children.

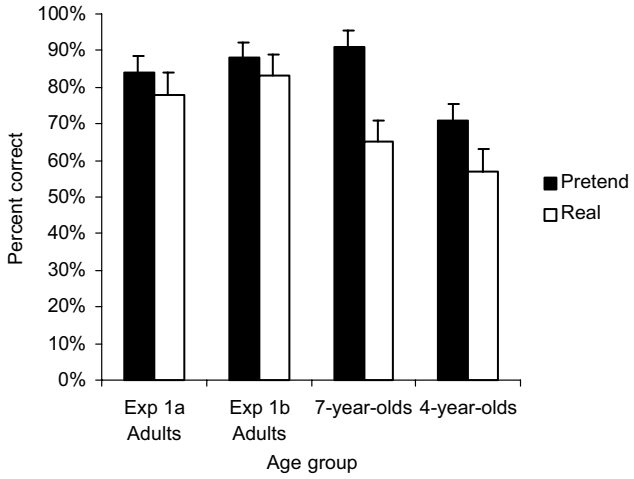

Fig. 1. Mean percent correct by age group on pretend and real clips used in Experiment 1a and Experiment 1b.

Adults were not trained. For the test phase, they were shown the 18 clips in the same random order as the children. They were tested in groups of three to five, and asked to mark on a piece of paper at the end of each clip whether they thought the event shown was pretend or real.

1.2.2. Results and discussion

Responses were coded as 1 for a correct response and 0 for an incorrect response. The percents correct for the different age groups on the pretend and real clips are shown in Fig. 1. Before comparing adult to child responses, Kolmogorov–Smirnov tests comparing binomial distributions for adults in Experiment 1a and the adults in Experiment 1b were conducted for each clip, and revealed no significant differences in responding on the clips. Thus, adult ratings on these clips were consistent across the two experiments. Data were also examined for possible response biases. Only one child in the younger group (3 years, 3 months) appeared to have a response bias, claiming all clips were pretend. This child was excluded from analysis.

Participants' responses were summed for the number of correct responses overall. Two-tailed *t* tests were conducted on the average percent of correct responding to test for chance responding. Participants responded at levels that were significantly above chance, both overall ($t[61] = 14.42$, $P < 0.001$) and within each group (4-year-olds: $t[15] = 3.47$, $P < 0.01$; 7-year-olds: $t[16] = 14.54$, $P < 0.001$; adults: $t[27] = 14.97$, $P < 0.001$). To test for a possible developmental trend, an Analysis of Variance was conducted on the number of correct responses with age group (4 s versus 7 s versus adults) as the between-subjects factor. There was a significant effect of age ($F[58] = 13.663$, $P < 0.001$). Planned contrasts revealed that the 7-year-olds ($M = 14.06$, S.D. $= 0.55$; $P < 0.01$) performed significantly better than the 4-year-olds ($M = 11.62$, S.D. $= 0.57$). There was no significant difference between the 7-year-olds and the adults ($M = 15.36$, S.D. $= 0.43$).

Finally, performance on each clip was examined to see if clips that were harder for children were also harder for adults. To examine this, Kolmogorov–Smirnov tests were conducted comparing responding on individual clips across all three age groups. There was a significant difference on only one (real) clip between the 4-year-olds and the adults ($Z = 1.72$, $P < 0.01$) and the 7-year-olds and the adults ($Z = 1.91$, $P < 0.01$). Only 29% of the 4-year-olds and 24% of the 7-year-olds responded correctly that this was a "real" clip, compared with 52% of the adults. This clip may have been more difficult for children because it contained a sound effect. Although sound effects do not appear to help toddlers interpret pretense (Lillard & Witherington, 2004), by the preschool years children may come to rely on them as a pretense. Parents very rarely make sound effects when they are engaged in real snacks.

1.3. Experiment 1c

The final step of Experiment 1 was to explore whether the percentage of correct responses for each clip varied consistently with aspects of mothers' behaviors known to vary across pretend and real behaviors (Lillard & Witherington, 2004). To this end, two extensively trained adult coders, blind to the frequency with which each clip had been correctly identified, coded the 36 clips for several criteria used by Lillard and Witherington (2004): proportion of time spent smiling, proportion of time looking to the child versus the task, duration of snack-related movements, duration of pauses between movements and the number of sound effects emitted. Coding of the durations were all highly reliable across coders: looking to the child, $r = 0.97$, $P < 0.01$, looking to the task, $r = 0.93$, $P < 0.001$, movements, $r = 0.80$, $P < 0.001$, pauses between movements, $r = 0.94$, $P < 0.001$, and smiles, $r = 0.87$, $P < 0.001$. In addition, the two raters agreed on 88% of the coded sound effects. Disagreements were resolved by discussion, and the agreements were used in the final analysis. Following coding of the clips, correlations were performed between percentage of correct responses on the clips and presence of the various behavioral cues.

1.3.1. Results and discussion

Analyses focused on the relationships between the percentage of participants in Experiment 1a who had rated a clip as pretend, and the strength of different types of cues in that clip. (Experiment 1a results were used because 36 clips provided a stronger basis for correlational analyses. Recall the difficulty ranking of the easiest 18 clips was similar for children and adults, making it likely that children used the same cues that adults did.) This provided a general measure of whether a given cue was related to participants' judgments that a particular clip was pretend or real.

These analyses revealed that rating a real snack clip as pretend was significantly more likely if the mother spent more time looking at the child, as is typical when pretending ($r[36] = 0.46$, $P < 0.01$). In contrast, if a mother spent more time looking at the task in the real condition, people more often identified real clips as real ($r[36] = 0.51$, $P < 0.01$). In addition, the percentage of participants rating a clip as pretend was also significantly correlated with the duration of time a mother held her hand at her mouth ($r[36] = 0.34$, $P < 0.05$) and the proportion of time spent making sound effects ($r[36] = 0.43$, $P < 0.01$).

Time spent smiling was not associated with pretend judgments, which is interesting in light of people's tendency to think smiling is a prime cue to pretense.

1.4. Summary

The results of Experiment 1 revealed a developmental progression in children's ability to identify whether the mother was pretending to have a snack or really having a snack. Four-year-olds were significantly less likely to correctly identify the video clips than were 7-year-olds or adults. One possible reason for the poorer performance of the younger children is that young children take significantly longer to encode information (Chi & Ceci, 1987). Perhaps longer clips would have allowed the children more time to encode the information and their performance would have improved. However, even the 4-year-olds were significantly better than chance overall at identifying the pretend and real scenarios combined. Thus, observers were proficient at identifying whether or not someone was pretending, even when obvious content cues (e.g., the presence or absence of Cheerios or water) were hidden.

In addition, despite the fact that the 4-year-olds in general performed more poorly on identifying the clips, the relative ranking of the clips was the same. Four-year-olds performed well or more poorly on the same clips, respectively, as the 7-year-olds and the adults performed well or poorly on. This suggests that children were probably not fundamentally different from adults in terms of what influenced their judgments of whether someone was pretending.

The final phase of Experiment 1 suggested that some features of pretense actions are associated with the likelihood that these actions would be labeled as pretense. More specifically, it may be that proportionally more looking to the child rather than the task, more sound effects, and conventionalized modifications in movements cue to observers that an action is pretense.

2. Experiment 2

Experiment 2 extended the findings of Experiment 1 in two ways. First, longer clips were used, to see whether viewing longer segments of behavior would assist very young children's interpretations (even though work with adults suggest that shorter clips should be as effective). Second, Experiment 2 used new clips, systematically selected for the presence of particular behaviors that other work has shown are typical of pretense versus real snacking. There are (at least) two experimental ways to implement this goal. First, one could create stimuli with well-trained actors varying their behaviors (e.g., movements, looking patterns, etc.). However, this approach risks losing something of the naturalness of a mother–child pretense interaction. The second approach, which was implemented, was to select video clips from genuine pretend and real mother–child snacking interactions based on the prevalence of particular cues.

2.1. Participants

Participants in Experiment 2 were 79 children and adults. Children were recruited from preschools and middle schools in a small university city, and the adults were recruited from

an undergraduate science group at the university. The majority of the participants were Caucasian, and a minority were African American or Hispanic. Participants were divided into three groups by age. The youngest group ranged in age from 3.1 to 5.5 ($n = 30$, $M = 4.4$, 15 males, 15 females), the middle group ranged in age from 9.0 to 10.9 ($n = 31$, $M = 9.10$, 11 males, 20 females), and the oldest group ranged in age from 15.1 to 61.8 ($n = 19$, $M = 23.2$). Some of the adult participants chose not to report their gender, so the exact breakdown is not known, but the experimenter judged the group to be about half female. Responses from one female student (age = 10.8) were dropped from analysis due to a "pretend" response bias, resulting in a middle group sample of 30 children.

2.2. Materials

2.2.1. Videotapes

Two videotapes were used for this experiment. Both videotapes contained the same twenty 20-s video clips (mean length = 20.08 s, S.D. = 0.36). The clips were in different random orders on each videotape. Longer clips were used for this experiment because, as was suggested in the discussion for Experiment 1, younger children may need longer examples of behaviors to make a correct identification. Clips were selected based on the presence or absence of different features. For each of five features (Movements, Smiling, Looking, Sound Effects, and All Features), four clips were chosen: (1) Real Good: a good example of a real eating act, (2) Real Bad: a real eating act that looked more like pretend, (3) Pretend Good: a good example of a pretend eating act, and (4) Pretend Bad: a pretend eating act that looked more like a real eating act. Clips showed eating, drinking, or pouring behaviors.

The good and bad examples of the features were chosen based on videotapes of mothers participating in our previous experiments (see Lillard & Witherington, 2004). The clips focusing on a particular cue contained at least two examples of that cue. For example, clips chosen as examples of eating movements contained at least two eating behaviors. For eating, in the clip demonstrating good examples of pretense, the length of the approach of the mother's hand from the bowl to the mouth was less than the mean observed in prior work (0.54 s) and the pause while the mother held her hand at her mouth was greater than the mean observed in prior work (0.64 s). In the good examples of real eating actions, the approach was greater than 0.76 s and the pause was less than 0.47 s (again, these are the means for these behaviors in the real condition in prior work). For poor examples, the opposite criteria were used, so the behavior in the bad-example real clips was more characteristic of the behavior normally seen in pretense.

This pattern of clip selection was applied to the other cues as well. Clips chosen as good examples of smiling in pretense had smiles that were longer than 4.57 s and that followed the mother's action; good examples of smiling in *real* interactions were less than 3.73 s and followed the child's action. In clips demonstrating good examples of pretense looking behaviors, mothers looked proportionally more often at their child than at the task. In clips demonstrating good examples of real looking behaviors, mothers looked slightly more often at the task than at the child. Clips demonstrating good examples of the use of sound effects in pretense incorporated many sound effects, whereas good examples of the use of sound effects in real interactions meant that one or no sound effect was used in the interaction.

Finally, in clips chosen for a specific feature, the salience of the other features was minimized to the extent possible. Thus, in the clips chosen for their variations in movements, the smiling and looking behaviors were average between pretend and real means so that they would not be a strong cue for identification of those 20 s of behavior. We reiterate that this was to the extent possible; being naturally derived behaviors, it was not possible to perfectly isolate each cue. To help minimize sound cues, the sound was removed from all but one of the video clips in which movement, looking, and smiling were intended to be cues. (One of these other clips did have sound, due to video editor error, but performance on this clip showed a pattern consistent with that found on other muted clips.) As in Experiment 1, video editing was used to block out the presence or absence of Cheerios and drink so that content cues would be absent. In this way, clips were selected to highlight particular cues and mute others, to the extent that this was possible given a limited selection of relatively natural mother–child snacking episodes.

2.2.2. Testing materials

Due to their different ages, the three groups of participants were given slightly different versions of the task, and thus different materials were used. Materials for the middle and older groups were a pen and paper survey on which participants were asked to mark a check beside the word "pretend" or the word "real." Materials for the younger group were two shoe boxes, one covered in red paper and one covered in yellow paper. A clear label protector was glued to the front of each box, and depending on the condition, a picture of either an empty bowl or a bowl full of Cheerios could be slid into the protector. Additionally, there were 20 pictures, one of each of the 20 mothers portrayed in the video clips. The pictures were chosen as still frames from the actual video clips shown to children, and all mothers wore a neutral expression.

2.3. Procedure

All participants watched the same 20-s video clips of mothers either pretending to have a snack or really having a snack.

2.3.1. Younger children

The younger children were invited to play a sorting game with the experimenter in a one-on-one interview. The children were shown two boxes and told, "I have two boxes here and I am going to tell you about a game that we can play with these boxes. This box is for when people are pretending [the experimenter pointed to the box with the picture of the empty bowl], and this box is for when people are doing things for real." The experimenter pointed to the box with the picture of the bowl with Cheerios in it. Children were then given the same training video clips used in Experiment 1. They were first shown a picture of a girl, Lori. Children were told, "This is Lori. We are going to watch Lori do some things, and then I am going to ask you about what Lori did. Sometimes Lori is going to be pretending, and sometimes she is going to be doing something for real. After we watch what Lori did, you get to put the picture of Lori in one of the boxes. If she was pretending, you put her picture in this box here. If she was doing something for real, you put her picture in this box here." After viewing each training clip, the experimenter paused the tape and asked the

child, "Which box should Lori's picture go in? Should it go in the pretending box or the really doing something box?" Children had to correctly sort three of the four pictures to proceed with the testing phase. Only one child was excluded for not correctly identifying three of the training clips.

To begin the testing phase, the experimenter said, "Now, I am going to show you some other women. Sometimes they are pretending to have a snack and sometimes they are really having a snack. After you watch them, you will get to put the picture of the woman into one of the boxes. If she was pretending to have a snack, then you will put her picture in this box here, and if she was really having a snack, you will put her picture in this box here." After each clip played, the experimenter paused the tape and asked the child, "Which box should her picture go in? Should it go in the 'pretending to have a snack' box or the 'really having a snack' box?"

The color of the "pretending to have a snack" box was counterbalanced across participants. In addition, the phrasing of the test questions was systematically varied within and between participants, by alternating which box was mentioned first in the question. To control for possible learning or boredom effects, half of the children saw the video clips in one random order, and the other half received the clips in an alternative random order.

2.3.2. Older children and adults

The older children watched the video together, with three teachers present. The adult participants also watched the video tape as a group. The experimenter explained to the participants that they would be watching a video with 20 clips of mothers pretending to have a snack or really having a snack. They were told to watch the video, decide whether it was a pretend or real situation, and mark their answer with a check next to "pretend" if the mother was pretending or "real" if the mother was eating for real. The experimenter emphasized that some of the clips would be hard, and that it did not matter if they were wrong or right. She also emphasized that they should not share answers and should remain quiet throughout the study. When all the clips had been shown, the experimenter collected the writing implements and told participants the correct answers so that they could compare how they did. For each clip, adults were also asked to write the basis on which they made their judgments. (These reflective responses were used in a different experiment.) Because older children and adults were run together, only one videotape could be used.

2.4. Results and discussion

Responses were coded as "1" if correct and as "0" if incorrect. Responses on the pretend and real clips were then summed for each cue, resulting in scores for each cue ranging from 0 to 2. (This scoring method was not possible in Experiment 1, since cues were discovered post hoc and were not systematically varied.) For example, for good examples of looking (to the child for pretend and to the task for real), participants could have a score from 0 to 2. This is also true for bad examples of looking (to the task for pretend and to the child for real). Responses were totaled and checked against possible chance responding. Participants responded at levels that were significantly above chance both overall ($t[78] = 11.47$, $P < 0.001$) and within each age group (young: $t[29] = 4.40$, $P < 0.001$; middle: $t[29] = 11.28$, $P < 0.001$; and adults: $t[18] = 8.46$, $P < 0.001$). A t test was conducted comparing the

child groups' responses by gender (since gender proportions were different for different age groups) and revealed no significant differences by gender.

A Repeated-Measures Analysis of Variance was conducted on responses, with age group (younger versus middle versus adults) as the between subjects variable and clip category (good versus bad) and cue (movements versus looks versus smiles versus sound effects versus loaded) as the within subjects variables. Simple planned contrasts were also included in the model to compare performance on each of the clips meant to focus on a specific cue with performance on the "loaded" clips. Repeated planned contrasts, comparing the youngest group to the middle group and the middle group to the adults, were included to test for specific age group differences.

There was a main effect for age group ($F[2, 76] = 2666.14$, $P < 0.001$). Planned contrasts indicated that the older group of children correctly identified a significantly larger percentage of the clips (74%) than did the younger group of children (57%). However, the older children did not respond significantly differently from the adults (80% correct). These findings demonstrate a similar age shift and overall level of performance as in Experiment 1. Age group did not interact significantly with any other variable, suggesting that the helpfulness of each type of cue was constant across age groups.

Importantly, there was a significant main effect for category, $F(1, 76) = 15.73$, $P < 0.001$. Participants responded correctly to a significantly higher percentage of the good clips (75%) than the bad clips (65%). In general, the clips selected as good examples of pretend or real behaviors were more likely to be identified correctly by participants, whereas participants were significantly less likely to correctly identify the clips selected as bad examples of pretend and real behaviors. This suggests that there is some clear effect of behavior alone, even when content is unavailable, for assisting people in identifying pretend and real clips. Even the bad exemplars of specific cues obtained, on average, above chance levels of correct identification. This may be because behaviors were never in isolation from other cues (since they were naturally derived clips). Still, the strength of those particular cues did have a significant impact on discrimination.

There was also a main effect for cue, $F(4, 304) = 7.61$, $P < 0.001$. The planned contrasts with responses on the "loaded" clips revealed that the percentage of participants who correctly identified the loaded clips (77%) was significantly greater than the percentage of participants who correctly identified clips focusing on smiling cues (64%; $F[1, 76] = 12.12$, $P < 0.01$) and sound effects (63%; $F[1, 76] = 12.98$, $P < 0.01$). There was no significant difference between level of performance on the loaded clips and those focusing on movements (73%) or looking (77%).

There was a significant category × cue interaction, $F(4, 304) = 5.79$, $P < 0.001$, indicating that perhaps one of the cues was not operating as predicted. Indeed, planned contrasts revealed a significant effect for the clips for which smiling was intended to be the primary cue, $F(1, 76) = 6.95$, $P < 0.01$. Smiling appeared to be having a reverse effect, as participants were more likely to respond correctly about clips meant to be poor exemplars of typical pretend/real smiling behaviors (68%) than clips intended as good exemplars of smiling (59%). Closer examination of these clips provided a possible explanation of this finding. In the clips meant to be good examples of smiling behavior, the mothers were demonstrating bad examples of the relevant looking behavior. Thus, in the case of the mother really eating Cheerios and demonstrating a good example of a "real" smile (short,

not after her own eating behavior), the mother was also demonstrating a bad example of "real" looking, by looking for longer at the child than at the task. This looking pattern is more often found in pretend interactions. Thus, it may be that smiling as a cue was not strong enough to override looking as a cue.

In summary, there are two key findings from this experiment. As in Experiment 1, the ability to correctly distinguish between pretend and real behaviors improved significantly from preschool to elementary school, and older children performed similarly to adults. However, age did not interact with category or cue, suggesting that the strategy used for determining which behaviors are pretend or real does not change across these ages. Children simply improve in their ability to use or perhaps even to detect these subtle changes in behavior.

The second key finding is that there are indeed certain behaviors that are more indicative of pretending than other behaviors. Video clips preselected for the presence of behaviors that typically vary across pretend and real snacks were more easily identified as pretend and real than were clips selected as poor examples of those behaviors.

3. General discussion

Two experiments examined how well children and adults identify pretense acts, and whether alterations in pretense behaviors uncovered in previous work (Lillard & Witherington, 2004) were related to the proficiency with which they did so.

Experiment 1 addressed two questions: (1) whether there was a developmental trend in the ability to distinguish pretense and real behaviors, when naturally occurring behavioral cues are made available to observers, and (2) whether proficiency in identifying pretend and real behaviors is related to specific behavioral cues. Adults and children were asked to distinguish between pretend and real eating in 8–12 s video splices with obvious contextual cues removed. Results revealed a developmental trend in the ability to differentiate when a mother was having a snack for pretend or for real. However, even 4-year-old children were significantly better than chance at judging whether episodes were pretense or real. In addition, results revealed significant relations between adults' proficiency to identify behaviors and the presence of specific behavioral cues. (Children's proficiency was not subjected to this analysis because the smaller number of tapes observed, and lack of preselection for specific cues, created insufficient power.) More specifically, likelihood that a particular clip was rated as pretend was correlated with proportionally more looking at the child and proportionally less looking at the task, more sound effects, and longer pauses in movements.

Since the developmental trend in Experiment 1 may have been due to the fact that younger children take longer to encode information than adults do (Chi & Ceci, 1987), Experiment 2 used longer clips with children to test whether their performance improves. In addition, to more directly address whether certain behavioral cues are indicators of pretense actions, clips for Experiment 2 were chosen based on the presence or absence of certain cues making them good or bad examples of real and pretend behaviors. Consistent with Experiment 1, there was a developmental trend in children's ability to correctly distinguish the pretend and real behaviors. Looking to the child and the conventionalized gesture of holding the hand at the mouth while eating again emerged as stronger behavioral indicators of pretense.

However, viewing longer clips and more behaviors did not improve young children's performance. It may be the case that repeated viewing of the clips would assist them. Repeated viewing is likely available in actual pretend interactions, because parents engage in more snack-related movements while pretend snacking than while really snacking (Lillard & Witherington, 2004).

In summary, there are two key findings from the experiments presented above. First, there is a significant developmental shift from preschool to elementary school age in children's ability to correctly differentiate between pretend and real behaviors. Despite the fact that the younger children in both experiments responded correctly at significantly above chance levels, they were still significantly less likely to respond correctly than children just a few years older. This age trend is interesting in light of research suggesting that children identify pretense intentions at age three (e.g., Rakoczy, Striano, & Tomasello, in press). Recent research suggests, however, that it is not until age 6 or 7 that children demonstrate a more solid understanding of the role that knowledge plays in pretense representations (Richert & Lillard, 2002). Thus, children's understanding of the connection between a pretender's intentions and actions is still developing over these years. The younger children in our experiments likely had more difficulty because our method was a stringent test of their ability to make on the spot judgments about an actor's pretend or real intentions, at an age when they are still learning to negotiate these connections.

The second key finding is that specific behaviors appeared to be associated with more correct pretense and real judgments. Specifically, where mothers were looking—at the child for pretense, and at the task for real—appeared to be associated with better judgments. In addition, the timing of mothers' movements—in pretense, faster when traversing space, but with a long conventionalized holding of the hand at the mouth—was also associated with better judgments. Although sound effects appeared to be important for adults in Experiment 1c, they were less helpful (but not unhelpful) for children. Significant effects were obtained for these behavioral indicators of pretense, despite the small number of clips in each group and despite the brevity of those clips. These experiments go beyond findings of research on children and teachers' ability to discriminate pretend and real fighting based on laughter and modified physical actions (Boulton, 1993), and have identified additional behavioral cues that may help children and adults discriminate pretending in non-aggressive scenarios.

Interestingly, in none of our experiments were patterns of smiling related to more apt judgments. This is surprising in light of people's tendency to assume that smiles are a major indicator of pretense. Indeed, we fully expected that smiling would emerge as a strong indicator to observers that an action is pretend. It is possible that the important aspect of smiling for older children and adults is not the duration and placement of smiles, but rather some quality for which we had not coded. Facial Action System coding of facial expressions in pretense is beginning to address this question. However, our other work suggests that at 18 months, a prime age for social referencing, adults' smiling just after a pretend behavior may be associated with improved understanding (Lillard et al., 2004; Nishida & Lillard, 2004).

Finally, it is worth reiterating that content cues were not available as a basis for pretense judgments in any of these experiments. One might assume that we determine whether events are pretend or real based on content (e.g., the presence or absence of Cheerios). These experiments show that behavioral cues alone can be sufficient for making pretense judgments at better than chance levels.

A few limitations to the present experiments should be noted. First, to achieve a structured experimental setting for comparison purposes, the conditions under which the pretense and real interactions occurred were tightly controlled; perhaps different cues would emerge in a more natural setting. One could argue, however, that if these cues were present in such a tightly controlled setting, in which even real snacks were somewhat "on stage" and thus pretend-like, they would be even more likely to occur in more natural settings. In addition, in an effort to achieve responses from participants before they had too much time for reflection, they were asked to rate very short video clips. The brevity of the clips, however, may have disrupted access to some of the cues (e.g., cutting smiles off) and certainly provided fewer cues for analysis. Even so, the findings of this research reveal that adults and children are proficient at discriminating between pretend and real interactions, even with short video clips of behavior. Furthermore, analysis of behavioral variation revealed that there are specific behavioral cues that appear to cue pretense to observers. Further research should explore the existence of such cues in more natural settings and with scenarios other then the snack scenarios. In addition, further studies should alter these cues in various pretense contexts to further explore if and when they facilitate children's understanding.

In conclusion, this research extends the limited body of evidence in discrimination of real and pretend fighting into a realm very relevant to young children: that of pretend and real eating. Even preschoolers correctly discriminated very short splices of behavior as pretend or real significantly more often than would be expected by chance. Further, specific behavioral changes known to accompany pretense acts were consistently associated with mistaking real video clips for pretend. This research thus breaks new ground in the search for what might enable infants to enter into the pretend realm, a first step towards their participation in imagined and hypothetical worlds.

Acknowledgments

This research was supported by Grant # RO1 HD36808 from the National Institutes of Health to the second author. We thank the children and their parents and teachers for participation, and Judy DeLoache, Lynn Heath, Lili Ma, Tracy Nishida, Amanda Robinette, Gabrielle Simcock, Amrisha Vaish, David Witherington, and many undergraduate research assistants for help with research and previous drafts.

References

Albright, L., Kenny, D. A., & Malloy, T. E. (1988). Consensus at personality judgments at zero acquaintance. *Journal of Personality and Social Psychology, 55*, 387–395.

Ambady, N., Bernieri, F. J., & Richeson, J. A. (2000). Toward a histology of social behavior: Judgemental accuracy from thin slices of the behavioral stream. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 32, pp. 201–271). San Diego, CA: Academic Press.

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin, 111*, 256–274.

Baldwin, D. A., & Baird, J. A. (1999). Action analysis: A gateway to intentional inference. In P. Rochat (Ed), *Early social cognition: Understanding others in the first months of life* (pp. 215–240). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Bateson, G. A. (1955/1972). A theory of play and fantasy. In G. A. Bateson (Ed.), *Steps to an ecology of mind* (pp. 39–51). New York: Chandler. (Reprinted from American Psychiatric Association Research Reports, 1955, II)

Bekoff, M. (1977). Social communication in canids: Evidence for the evolution of a stereotyped mammalian display. *Science*, *197*, 1097–1099.

Bekoff, M., & Allen, C. (1998). Intentional communication and social play: How and why animals negotiate and agree to play. In M. Bekoff & J. A. Byers (Eds.), *Animal play: Evolutionary, comparative, and ecological perspectives* (pp. 97–114). Cambridge: Cambridge University Press.

Bekoff, M, Byers, J. A. (1981). A critical reanalysis of the ontogeny of mammalian social and locomotor play: An ecological hornet's nest. In K. Immelmanm, G. W. Barlow, L. Petrinovich, & M. Main (Eds.), *Behavioral development: The Bielefeld interdisciplinary project* (pp. 296–337). New York: Cambridge University Press.

Biben, M., & Suomi, S. J. (1993). Lessons from primate play. In K. MacDonald (Ed.), *Parent-child play: Descriptions and implications* (pp. 185–196). Albany: State University of New York Press.

Boulton, M. J. (1991). A comparison of structural and contextual features of middle school children's playful and aggressive fighting. *Ethology and Sociobiology*, *12*, 119–145.

Boulton, M. J. (1993). Children's ability to distinguish between playful and aggressive fighting: A developmental perspective. *British Journal of Developmental Psychology*, *11*, 249–263.

Brand, R. J., Baldwin, D. A., & Ashburn, L. A. (2001). Evidence for "motionese": Modification in mother's infant-directed action. Unpublished manuscript, University of Oregon.

Bruner, J. (1979). The organization of action and the nature of the adult-infant transaction. In E. Tronick (Ed.), *Social interchange in infancy* (pp. 23–35). Baltimore: University Park Press.

Chi, M. T., & Ceci, S. J. (1987). Content knowledge: Its role, representation, and restructuring in memory development. In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 20, pp. 91–142). San Diego, CA: Academic Press.

DeLoache, J. S., & Plaetzer, B. (1985). *Tea for two: Joint mother-child symbolic play*. Paper presented at Bienniel Meeting for the Society for Research in Child Development, Toronto.

Fagen, R. (1981). *Animal play behavior*. New York: Oxford University Press.

Farver, J. (1992). An analysis of young Mexican and American children's play dialogues. In C. Howes, O. A. Unger, & C. C. Matheson (Eds.), *The collaborative construction of pretend: Social pretend play functions* (pp. 55–66). Albany, NY: State University of New York Press.

Funder, D. C., & Colvin, C. R. (1988). Friends and strangers: Acquaintanceship, agreement, and the accuracy of personality judgment. *Journal of Personality and Social Psychology*, *52*, 149–158.

Gergely, G., & Watson, J. (1999). Early social-emotional development: Contingency perception and the social-biofeedback model. In P. Rochat (Ed.), *Early social cognition* (pp. 101–136). Hillsdale, NJ: Erlbaum.

Haight, W. L. & Miller, P. J. (1993). *Pretending at home*. Albany: SUNY Press.

Harris, P. L. (2000). *The work of the imagination*. Oxford, UK: Blackwell Publishers.

Harris, P. L. & Kavanaugh, R. D. (1993). Young children's understanding of pretense. *Monographs of the Society for Research in Child Development, 58*(1, Serial No. 231).

Hill, H. L., & Bekoff, M. (1977). The variability of come motor components of social play and agonistic behavior in infant eastern coyotes. *Canis Latrans var. Animal Behavior*, *25*, 907–909.

Kavanaugh, R. D., Whittington, S., & Cerbone, M. J. (1983). Mother's use of fantasy speech to young children. *Journal of Child Language*, *10*, 45–55.

Kaye, K. (1982). *The mental and social life of babies*. Chicago: University of Chicago Press.

Lillard, A. S., Massaro, D., Vaish, A., Nishida, T. N., Ma, L., & McRoberts, J. (2004). *Mothers' vocal and behavioral signs of pretense with 15-, 18-, and 24-month-olds*. Unpublished manuscript, University of Virginia.

Lillard, A. S., & Witherington, D. (2004). Mothers' behavior modifications during pretense snacks and their possible signal value for toddlers. *Developmental Psychology*, *40*, 95–113.

Nicholich, L. M. (1977). Beyond sensorimotor intelligence: Assessment of symbolic maturity through analysis of pretend play. *Merrill-Plamer Quarterly*, *23*, 89–99.

Nishida, T., & Lillard, A. S. (2004, May). *Social referencing behavior in mother-infant pretense*. Paper presented at the International Conference on Infant Development, Chicago.

Piaget, J. (1962). *Play, dreams, and imitation in childhood*. New York: Norton.

Rakoczy, H., Tomasello, M., & Striano, T. (in press). Young children know that trying is not pretending: A test of the "behaving as if" construal of children's understanding of "pretense". *Developmental Psychology*.

Richert, R. A., & Lillard, A. S. (2002). Children's understanding of the knowledge prerequisites of drawing and pretending. *Developmental Psychology, 38*, 1004–1015.

Ruesch, J., & Bateson, G. (1951). *Communication: The social matrix of psychiatry*. New York: Norton.

Smith, P. K., & Boulton, M. J. (1990). Rough-and-tumble play, aggression, and dominance: Perception and behavior in children's encounters. *Human Development, 33*, 271–282.

Tamis-LeMonda, C. S., & Bornstein, M. H. (1991). Individual variation, correspondence, stability, and change in mother and toddler play. *Infant Behavior & Development, 14*, 143–162.

Wellman, H. M., & Hickling, A. K. (1993). Understanding pretense as pretense. *Monographs of the Society for Research in Child Development, 58*, 93–102.

Wilson, T. D., & Schooler, J. W. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality & Social Psychology, 60*, 181–192.