Journal of Cognition and Development

ISSN: 1524-8372 (Print) 1532-7647 (Online) Journal homepage: http://www.tandfonline.com/loi/hjcd20

Person Perception in Young Children Across Two Cultures

Eva E. Chen, Kathleen H. Corriveau & Paul L. Harris

JOURNAL OF COGNITION AND DEVELOPMENT, 17(3):447–467

ISSN: 1524-8372 print/1532-7647 online
DOI: 10.1080/15248372.2015.1068778

Routledge
Taylor & Francis Group

Person Perception in Young Children Across Two Cultures

Eva E. Chen
The Hong Kong University of Science and Technology, Hong Kong

Kathleen H. Corriveau
Boston University

Paul L. Harris
Harvard University

To adult humans, the task of forming an impression of another social being seems effortless and even obligatory. In 2 experiments, we offer the first systematic cross-cultural examination of impression formation in European American and East Asian preschool children. Children across both cultures easily inferred basic personality traits, such as *nice* and *mean*, about unfamiliar peers from behavioral information, whether or not they were specifically prompted to do so. Children were able to identify peers they had seen before, to remember the traits associated with these peers, and to anticipate future behaviors consistent with the traits they had attributed. Thus, for basic traits, the ability to make behavior-to-behavior predictions, via an intervening trait inference, is present in young children across diverse cultures.

Humans come to conclusions about the personality of other people rapidly and easily (Asch, 1946). We make inferences about a person's traits, such as their trustworthiness or competence, using small amounts of information based on observation or even hearsay (Ambady & Rosenthal, 1992; Anderson & Hubert, 1963; Carlston & Skowronski, 1994; Cogsdill, Todorov, Spelke, & Banaji, 2014; Hamilton, Katz, & Lierer, 1980; Jones, 1979; Jones & Nisbett, 1972; Pronin & Ross, 2006). Once a trait has been inferred, we associate it with that person across situations and time (Todorov & Uleman, 2002, 2004; Todorov, Gobbini, Evans, & Haxby, 2007). We also need not remember the original behavioral details about a person once the relevant personality traits have been inferred; instead, we rely on our robust person memory, including our memory for a person's face and the associated personality traits (Fiske & Taylor, 1991). Not only are these traits (e.g., *sincere, friendly, deceptive*) used to evaluate people (Hamilton et al., 1980), but they are also used to make predictions about future behavior (Rholes & Ruble, 1984). The attribution of traits and the prediction of future behaviors based

Correspondence should be sent to Eva E. Chen, Social Science, The Hong Kong University of Science and Technology, Room 2346, Academic Building, Clear Water Bay, Kowloon, Hong Kong. E-mail: evaechen@ust.hk
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hjcd

on those traits appear to be spontaneous and automatic among adults (see Gilbert, 1998; Todorov & Uleman, 2004).

In the present study, we examined the origins and potential universality of impression formation. We asked whether the rapid processing and retention of trait information emerges early in development and whether such processing and retention are robust across two distinct cultures. In the following sections, we review relevant research in two pertinent domains: a) young children's ability to form impressions, and b) cross-cultural evidence on trait attribution.

The Development of Impression Formation in Children

Early research on impression formation has suggested that young children's processing of trait information is limited. Livesley and Bromley (1973) reported that when faced with unfamiliar individuals, young children are sensitive to their physical appearance and social roles (e.g., parent vs. child) but do not regularly attend to psychological characteristics until they are around 7 years of age. Similarly, Rholes and Ruble (1984) found that although children younger than 9 years can infer traits correctly after hearing a behavioral vignette (e.g., labeling an individual as *generous* after learning that he voluntarily shares his lunch), they may subsequently have difficulty in using the inferred trait information to predict behavior (e.g., anticipating that the individual will continue to be generous in the future).

Other research, by contrast, offers a more positive portrait of preschoolers' ability to process trait information. The ability to talk about certain traits (e.g., *nice*) emerges in toddlerhood (Bretherton & Beeghly, 1982). Indeed, by 3 years of age, children prefer to use trait labels over physical appearances to make inferences about the characteristics of story characters (Heyman & Gelman, 2000). By 4 years of age, they can rely on personality trait labels to make predictions about the behavior of unfamiliar individuals (Liu, Gelman, & Wellman, 2007), as well as infer traits (e.g., *nice, mean*) from behavioral descriptions (Boseovski & Lee, 2006; Liu et al., 2007). Finally, by 6 years of age, there is some evidence that children can use past behaviors to make predictions regarding future behavior (e.g., anticipating that a character will be helpful to his sister after learning that the character helps elderly people down the stairs, shows new students around the school, and helps set the table; Gnepp & Chilamkurtu, 1988).

Nevertheless, the use of past behaviors to make behavioral predictions via the attribution of a trait is quite difficult for young children. Liu et al. (2007) found that across three age groups (ages 4, 5, and 7 years), children were better at making trait-consistent predictions (e.g., predicting that a mean character would laugh at a child who fell and cried) when they were provided with trait labels (e.g., *mean*) compared with when they were only provided with trait-consistent behaviors (e.g., the character pulling on the family cat's tail). These results suggest that behavior-to-behavior predictions may be a process involving two components—a behavior-to-trait inference component and a trait-to-behavior prediction component, with younger children having trouble coordinating both components (Liu et al., 2007).

In this study, we investigated these processes further with preschool children. Specifically, we sought to confirm whether it is indeed difficult for preschool-aged children to make behavior-to-behavior predictions. To increase the likelihood that preschoolers might make such behavior-to-behavior predictions, we deliberately created behavioral vignettes showing the protagonist acting in either a nice or mean fashion. The findings of Bretherton and Beeghly (1982) demonstrated that by the time children reach 2 years of age, they are capable of using words that could be understood as trait words (e.g., *nice, naughty, mean*), indicating that these traits are understood and produced very early in development. In addition, we examined person perception processes across two distinct cultures—one individualistic, one collectivistic—to assess whether the use of past behavioral information to infer traits and predict future behaviors varies across these two cultural groups.

The Influence of Culture

Previous studies have shown that culture has a considerable impact on the manner in which traits are conceptualized and inferred. In particular, differences have been found between more individualistic (e.g., American) and more collectivistic (e.g., Asian) cultures. For instance, East Asian adults are more likely than European American adults to consider dispositions as flexible and susceptible to contextual factors (Choi, Nisbett, & Norenzayan, 1999). They are also more optimistic than Americans about the possibility of certain traits (e.g., *unkind*) changing during the life span (Lockhart, Nakashima, Inagaki, & Keil, 2008). American adults are more likely to attribute behaviors (e.g., falsifying tax information) to general dispositions or personality characteristics, whereas Indian adults tend to attribute such behaviors to context (Miller, 1984). Additionally, European American adults are more likely, as indexed by reaction time data as well as electroencephalogram data, to spontaneously infer trait information from behavioral descriptions compared with Asian Americans (Na & Kitayama, 2011).

Research has indicated that some of these differences may have resulted from enculturation; that is, as they age into adulthood, individuals shift their focus to either personality characteristics or contextual factors in explaining others' behaviors, depending on whether they are immersed in a more individualistic (Western) culture or a more collectivistic (non-Western) culture, respectively (Miller, 1984, 1986). This enculturation process suggests that cultural differences may be less apparent among young children (Miller, 1986). For instance, younger children (ages 5–6 years) from both American and Japanese cultures anticipated similar amounts of change in certain psychological traits over time; by contrast, older American children (ages 8–10 years) were more inclined than Japanese children to treat these traits as stable features (Lockhart et al., 2008). Similarly, as compared with their Latino peers, European American fifth graders were more likely to predict consistency in an individual's behavior and were less likely to attend to situational factors (Newman, 1991). However, more recent work has suggested that some cultural differences in person perception emerge even at a young age. Kuwabara, Son, and Smith (2011) found that Japanese 3-year-olds were more likely than their American counterparts to adjust their judgments of a character's emotions in the light of contextual information. Thus, young children from a more collectivistic culture, in attending to context, might be more reluctant to attribute trait information to unfamiliar characters, to remember this trait attribution over time, and to use it to make predictions of future behavior.

To our knowledge, this study is the first to provide a systematic cross-cultural comparison of young children's person perception abilities. We focused on children aged 4 years to 7 years old because previous work has demonstrated that by the time they are in preschool, children are able to reliably make psychological inferences based on trait label information (e.g., Heyman & Gelman, 2000). Additionally, across different cultures (e.g., American and Japanese), preschool children are able to distinguish between psychological and physical traits and can reason about the stability of psychological traits across the life span (Lockhart et al., 2008). We compared children from an individualistic culture and children from a collectivistic culture to assess differences in their abilities to infer, retain, and subsequently utilize traits in making behavioral predictions.

In Experiment 1, we compared the ability of European American and Taiwanese children to attribute traits on the basis of behavioral vignettes and to subsequently retain information about both traits and behaviors. In Experiment 2, we asked how far Taiwanese children make trait attributions and trait-based predictions, with and without external prompting.

EXPERIMENT 1

Method

Participants

One hundred and four preschool-age children (49 girls; M_{age} = 5;5, age range = 4;2–6;6) from two cultures were recruited. In line with previous trait research with preschool children (e.g., Liu et al., 2007; Livesley & Bromley, 1973), we targeted children starting from 4 years of age to 7 years of age. Fifty-nine (33 girls; M_{age} = 5;5, age range = 4;2–6;6) European American children were recruited from a local museum in Boston, Massachusetts. Because of museum policy, no formal socioeconomic-status information could be collected about the participants' family background, but prior visitor surveys conducted by an external evaluator of the museum indicated that the majority of the parents of visiting families had college or graduate degrees (Soren, 2009). Thus, the participants were likely from middle- and upper middle-class families. Additionally, 45 (16 girls; M_{age} = 5;4, age range = 4;4–6;3) Taiwanese children were recruited from a local preschool in Taipei, Taiwan. Because the preschool was affiliated with a well-known public university, the majority of the participants' families were highly educated and therefore were considered to be middle or upper middle class. For all participants, written permission was first obtained from a legal guardian, and verbal consent was obtained from the participant before the experiment began.

Materials. The experiment was conducted in three phases: the *trait inference* phase, the *face recognition* phase, and the *trait memory* phase. In the *trait inference* phase, eight cards featuring photographs of children were used (see Figure 1). Half of the cards featured girls, and the other half featured boys; additionally, half of the children were European American (White), and the other half were East Asian (Asian). Photographs of the children were chosen so that the faces and facial expressions in each of the four groups (White boy, White girl, Asian boy, and Asian girl) were physically similar to one another to ensure that participants were less able to use a unique facial feature to remember a particular character. Each photograph was associated with a vignette describing the card character's behavior. The vignettes (see Appendix A) were modeled after Liu et al. (2007) and Todorov et al. (2007). Half depicted positive behaviors (e.g., helping to set the table for dinner), whereas the other half depicted negative behaviors (e.g., pulling on a dog's tail). Photographs were counterbalanced across participants, such that a stimulus face that belonged to a nice child for one group of participants belonged to a mean child for the other group of participants, and vice versa. We chose these two trait labels to increase the

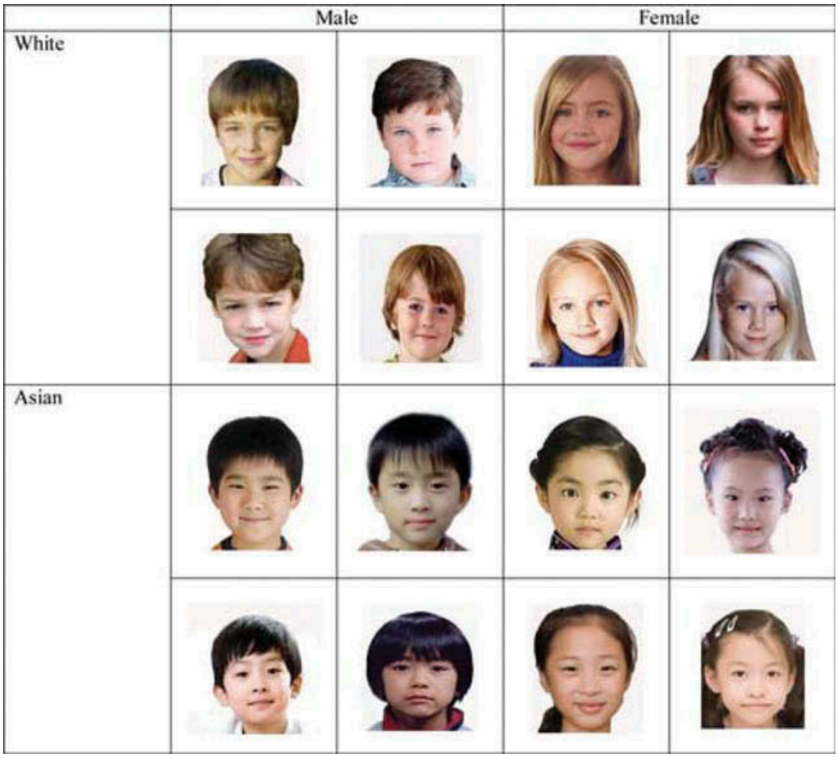

FIGURE 1. Stimuli used in both experiments. Participants saw 8 of the photographs (2 from each quadrant) in Phase 1, all 16 faces in Phase 2, and the original 8 for Phase 3.

likelihood that even the youngest participants in the two experiments could complete the task successfully (e.g., Heyman & Gelman, 1999).

In the *face recognition* phase, eight more cards were added to the original photographs from the trait inference phase. As in the original set of pictures, half of the cards featured girls and the other half featured boys; in addition, half of the cards featured White children and the other half featured Asian children. The 16 photographs were counterbalanced across participants such that each stimulus face was either an *old* face that they had seen before or a *new* face.

Finally, in the *trait memory* phase, the original eight pictures were re-presented.

For the Taiwanese children, experimenter instructions and behavioral vignettes were translated into Mandarin Chinese by a native Taiwanese speaker, unaware of the experimental premises, and then were back-translated into English by another bilingual Mandarin–English speaker.

Procedure

In the trait inference phase, participants were asked to infer traits from behavioral descriptions of fictional child characters. An experimenter showed each participant eight cards, one card at a time. While showing each card, the experimenter read a vignette describing the card character's behavior and then asked: “Do you think this child was nice or mean?” Once the experimenter had posed the question, the card was given to the participant, who placed it in either a nice category (represented by a smiling cartoon face) or a mean category (represented by a frowning cartoon face).

Once the trait inference phase was complete, children were given a face recognition task to check that they could reidentify the faces they had seen before and distinguish them from new faces. The experimenter gathered the first set of 8 cards and shuffled them with 8 new cards. Participants were then shown the 16 cards, now randomly shuffled, 1 at a time; and for each card, they were asked to indicate whether they recognized the pictured child as one they had previously seen or as one who was new to them.

Finally, in the trait memory phase, participants were shown the original eight cards from the initial trait inference phase. After reshuffling the cards, the experimenter gave participants the cards, one at a time, and without repeating the associated vignette, asked participants to sort the cards into the nice category or the mean category. Once participants had placed a given card in one of the two categories, they were asked if they could recall the behavioral vignette about that card character. Participants’ responses were recorded by the experimenter. Together, the three phases assessed: a) whether participants were able to infer traits from behavioral descriptions of fictional children, b) whether participants were able to recognize faces they had seen previously and distinguish them from faces they had not seen before, and c) whether participants were able to remember i) the traits and ii) the behavioral vignette associated with each face.

Results

Key results were as follows. Participants excelled at inferring who was nice and who was mean based on the behavioral vignettes. In the subsequent face recognition phase, children were quite accurate at identifying both old and new faces. Finally, in the trait memory phase, participants often remembered who had been nice or mean even if they frequently failed to recall the behavior that had led them to that inference. In all three phases, the pattern displayed by the European American and Taiwanese children was very similar. In the following paragraphs, we document these results in more detail.

Phase 1: Trait Inference Phase

Overall, participants were very good at sorting the characters into the two categories based on the behavioral vignettes. There was no significant difference in the number of correct trait inferences between the European American and Taiwanese children, $t(102) = 0.50$, $p = .62$. Four of the 59 European American participants (7%) and 2 of the 45 Taiwanese participants (5%) were correct in seven out of eight trials in the trait inference phase. The remaining 55 European American (93%) and 43 Taiwanese (95%) participants correctly inferred the traits from all eight behavioral descriptions.

Phase 2: Face Recognition Phase

Next, we investigated the ability of participants to recognize the faces they had seen in the previous trait inference phase (old faces) and to distinguish them from the faces they had never seen before (new faces). On average, Taiwanese children were slightly better than their European American counterparts at correctly recognizing old faces (Taiwanese, $M = 6.98$, $SD = 1.25$;

European American, *M* = 6.27, = 2.11), *t*(102) = 2.00, *p* = .05, *d* = 0.40, and also at identifying new faces (Taiwanese, *M* = 7.67, *SD* = 0.77; European American, *M* = 7.05, *SD* = 1.99), *t*(102) = 1.97, *p* = .05, *d* = 0.39. Despite these differences, both European American and Taiwanese participants scored significantly above chance (i.e., 50%, or correctly on four out of eight trials) for both the old faces—European American, *t*(58) = 8.28, *p* = .001, *d* = 2.17; Taiwanese, *t*(44) = 15.95, *p* = .001, *d* = 4.81—and the new faces—European American, *t*(58) = 11.80, *p* = .001, *d* = 3.10; Taiwanese, *t* (44) = 32.00, *p* = .001, *d* = 9.65—indicating that on average, their recognition of the faces they had previously seen was robust.

Phase 3: Trait Memory Phase

Participants' ability to remember the traits associated with the original eight faces they had seen in the trait inference phase was examined. This analysis was conducted in three steps. First, we examined how often participants correctly remembered the faces as nice or mean. Second, we asked how often they were able to recall the relevant behavioral vignette for a given face. Finally, we examined the relationship between memory for traits and memory for behavioral vignettes more closely. More specifically, we asked how well participants were able to remember the traits both when they were also able to remember the relevant vignette and also when they were unable to do so.

To anticipate the overall pattern of results, participants a) were good at remembering whether a given face was nice or mean, b) often failed to recall the relevant behavioral vignette, but c) still remembered whether a face was nice or mean even when they could not recall the relevant vignette. Nevertheless, correct recall of the vignette was associated with better trait memory. This pattern emerged for both European American and Taiwanese children.

Remembering Faces as Nice *Versus* Mean. To provide an overall analysis of children's performance, we conducted a four-way repeated-measures analysis of covariance (ANCOVA), with participant gender (male, female) and participant race (White, Asian) as between-subjects factors; stimulus gender (male, female) and stimulus race (White, Asian) as within-subjects factors; and arcsine square root of the proportion of correct memory judgments as the dependent variable. Given that the proportion of correct answers for each of four categories (White girl, White boy, Asian girl, and Asian boy) was our dependent variable, we conducted arcsine transformations on the proportions to account for the non-Gaussian distribution of our results. Participant age (in months) was included as a covariate. Results revealed only a significant main effect of age, *F*(1, 99) = 4.16, *p* = .04, η^2_{par} = .04. To examine the effect of age further, we divided the children into two groups based on the median age (63.50 months [5;4]). Older children (aged 63.50 months [5;4] and older; *n* = 52) performed (*M* = 5.63, *SD* = 1.33) marginally better than younger children (*n* = 52; *M* = 5.10, *SD* = 1.73) in recategorizing the characters correctly, *t*(102) = 1.78, *p* = .08. Nevertheless, comparisons to 50% chance (= 4 faces categorized correctly) confirmed that both older and younger children sorted the faces systematically: older children, *t*(51) = 8.87, *p* = .001, *d* = 2.48; younger children, *t*(51) = 4.57, *p* = .001, *d* = 1.28.

Behavior Recall. Participants' recall of the behavioral vignette associated with each stimulus face was recorded during the experiment. The responses of the Taiwanese children were translated into English by a bilingual experimenter. Subsequently, two independent coders who were not involved in the study and who did not know which stories were affiliated with each of the faces coded the recall of a subset of the European American participants. Children were not often able to recall any given vignette in a verbatim fashion. Accordingly, to assess children's ability to correctly recall the gist of a vignette, the coders were given a list of the eight behavioral vignettes and were asked to match the recall of the participants to one of the vignettes; failure to generate any story and generation of a completely new story were also coded accordingly. Agreement between the two coders was 98% (Cohen's K = .94); all disagreements were resolved through discussion between the two coders and the first author. Participants' recall responses, based on the independent coders' judgments, were divided into three categories: a) *correct recall* (i.e., the generation of a story that the coders correctly matched to the story associated with the stimulus face); b) *incorrect recall* (i.e., the generation of a story that the coders did not correctly match to the story associated with the stimulus face); and c) *no recall* (i.e., a failure to generate any story for the stimulus face).

The pattern of recall (i.e., correct vs. incorrect vs. no recall) varied by cultural group. On average, European American participants (1.36 times out of a total of 8, or 17% of the time) were significantly more likely than Taiwanese participants (0.49 times out of 8, or 6.13% of the time) to recall a correct story, $t(102) = 3.14$, $p = .002$, $d = 0.62$. However, European American children were also more likely to recall an incorrect story (0.83 times out of 8, or 10.38%) compared with Taiwanese children (0.20 times out of 8, or 2.50%), $t(102) = 2.92$, $p = .004$, $d = 0.58$. Thus, Taiwanese children (7.31 times out of a total of 8 times, or 91.37%) were significantly more likely than European American children (5.81 times out of a total of 8, or 72.62%) to not recall any story, $t(102) = 3.99$, $p = .001$, $d = 0.79$.

Relationship Between Trait Memory and Behavior Recall. The results so far indicate that children performed quite well in remembering faces as nice or mean but performed poorly at recalling the relevant behavior. Next, we examined the relationship between these two measures. We compared participants' accuracy in categorizing a face for which they recalled the correct story to their accuracy in categorizing a face for which they failed to recall a story. For instance, if a participant correctly recalled the vignette for two out of the eight faces, we calculated the proportion of the two faces that were correctly categorized as nice or mean; then, if the participant failed to recall a story for five out of eight faces, we calculated the proportion of the five faces that were correctly categorized.

Of the participants, 33 European American children and 9 Taiwanese children a) were able to recall a correct story for at least one stimulus face, and b) failed to recall a story for at least one stimulus face. For the European American children, a paired-samples t test revealed that children were more accurate in categorizing the faces for which they could recall the vignette ($M_{\text{proportion}} = 0.97$, $SD = 0.10$), compared with the faces for which they recalled no vignette ($M_{\text{proportion}} = 0.57$, $SD = 0.30$), $t(32) = 7.67$, $p = .001$, $d = 2.71$. Similarly, a paired-samples t test revealed that the Taiwanese children were more accurate in categorizing faces for which they could recall the correctly matched vignette ($M_{\text{proportion}} = 1.00$, $SD = 0.00$) compared with faces for which they recalled no vignette ($M_{\text{proportion}} = 0.81$, $SD = 0.17$), $t(8) = 3.41$, $p = .009$, $d = 2.41$.

In summary, analyses of children's trait memory revealed that although they accurately remembered the faces as nice or mean, children failed to recall the relevant vignette for the majority of faces. However, children's categorization of the faces as nice or mean was more accurate when they were able to recall the relevant vignette. This pattern of findings emerged for both European American and Taiwanese children, although European American children were more likely to recall behavioral information from the vignettes—both correct and incorrect.

Discussion

Experiment 1 showed that European American and Taiwanese children can infer nice and mean traits from behavioral descriptions of characters they have not seen before. Subsequently, they can recognize the faces of these characters and differentiate them from new faces. Finally, they are able to remember the traits associated with particular faces. In the final trait memory phase, children were unable to recall the majority of the behavioral descriptions but were still able to remember who was nice and who was mean. In Experiment 2, we probed person perception abilities in more detail. Because few cultural differences emerged in the first experiment—children in both cultures performed quite similarly in inferring and remembering trait information about unfamiliar individuals—we focused solely on Taiwanese children in the second experiment.

We probed three aspects of children's ability to infer and remember trait information in Experiment 2. Recall that in the initial trait inference phase of Experiment 1, children were asked, following each story, to categorize the story character as either nice or mean. It is possible that this prompt to categorize each story character played a major role in children's relatively good performance both during the initial trait inference phase and also in the final trait memory phase. This speculation is especially pertinent for the Taiwanese children because cross-cultural findings with adults have suggested that people of Asian descent are less likely than those of European American descent to infer traits from behavioral descriptions unless prompted to do so (Na & Kitayama, 2011). Accordingly, participants were given one of three sets of instructions in the trait inference phase. Some children were given a *systematic prompt*. After each vignette, they were asked to explicitly say whether the character had been nice or mean (just as in Experiment 1). By contrast, a second group of children were given only one *generic prompt* before they listened to any of the vignettes. Children were briefly introduced to the two familiar personality trait labels (e.g., Bretherton & Beeghly, 1982) but were not instructed to make trait inferences. Finally, a third group of children received *no prompt* prior to listening to the vignettes. They were neither provided with trait labels nor with instructions to infer traits from the behavioral stories. We anticipated that the absence of systematic prompting following each vignette might attenuate trait inference and retention in the generic and no prompt conditions.

In Experiment 2, we also examined children's ability to make forward-looking trait predictions about the character's future behavior. Thus, in the third phase of the study, children were not asked to recall whether the story character was nice or mean. Instead, they were presented with two potential behaviors—one nice and one mean—and were invited to predict in which behavior the character would engage. Prior research by Liu et al. (2007) has suggested that 4-, 5-, and 7-year-olds have difficulty in making behavior-to-behavior predictions when such predictions call for an intervening trait inference. Experiment 2 provided an opportunity to examine whether children display the same difficulties if, at an earlier point in time, they had been given a systematic prompt to state the intervening trait, only a generic prompt, or no prompt at all.

Finally, Experiment 2 was designed to explore a plausible explanation for the considerable gap between trait recall and behavioral recall observed in Experiment 1. In that experiment, the trait recall task was a forced-choice measure. Children only had to recategorize the characters as nice or mean to demonstrate their memory for the characters' traits. However, the behavioral recall task was open-ended. It required children to recall the behaviors of the characters to demonstrate their retention of what the character had done. In Experiment 2, the behavioral recall task was a forced-choice task, as described in more detail in the next section.

EXPERIMENT 2

Method

Participants

Seventy-five (31 girls; M_{age} = 5;3, age range = 3;10–6;6) Taiwanese children were recruited from two local preschools in Taipei, Taiwan. Children from both preschools were from families of comparable socioeconomic status to those families in Experiment 1. As in the first experiment, written permission was obtained from one of the participant's legal guardians and verbal consent was obtained from participants before the experiment began. The youngest participant (age = 3;10) was unable to complete the experiment; thus, analyses were run on the remaining 74 (31 girls; M_{age} = 5;3, age range = 4;3–6;6) participants.

Materials and Procedure

The experimenter instructions and behavioral vignettes were created in English and translated into Mandarin Chinese by a native Taiwanese speaker, unaware of the experimental premises, and then were back-translated into English by another bilingual Mandarin–English speaker. As with the previous experiment, Experiment 2 consisted of three phases, of which the second (face recognition) was the same. However, several changes were made to the other two phases. First, to test whether explicitly categorizing the behavior of each story protagonist as *nice* or *mean* in the initial phase would lead to better performance at later phases, we randomly assigned participants to one of three groups. There were no significant differences in participants' ages between the three groups, $F(2, 73) = 0.70$, $p = .50$. The first group of participants (n = 26; M_{age} = 5;4, SD = 7.33 months) completed the same trait inference phase task as those in Experiment 1 (the *systematic prompt* group); that is, after hearing each of the eight behavioral vignettes (see Appendix A), they were explicitly asked to categorize the character in that particular vignette as either nice or mean. The second group of participants (n = 26; M_{age} = 5;2, SD = 8.15 months) was given only a generic prompt (the *generic prompt* group). Before they heard any of the stories, they were prompted to think about the characters (i.e., "In these stories, some children are very nice, but some children are very mean. I want you to think carefully about them!"), but they were not asked to do anything else after listening to each individual story. For the third group of participants (n = 22; M_{age} = 5;2, SD = 5.02 months), the experimenter simply asked children to listen closely to the stories and pay attention to the protagonists' faces (the *no prompt* group); the words *nice* and *mean* were not mentioned, and children were not asked to do anything else after hearing each story. In all three conditions, children were given time to respond. The experimenter handed the card depicting the character to the children to categorize in the systematic prompt group and simply paused after each story in the generic and no prompt groups; if the children in the latter two groups responded, the experimenter thanked them afterward, regardless of the content, and moved to the next character and behavioral vignette.

We also made two changes to the final phase. First, instead of recalling whether the eight original characters were nice versus mean after the face recognition phase, participants received a forced-choice *behavioral prediction* task (i.e., they were asked to predict how each character would act when meeting a new gender-matched character, Dodo ["Little Bean" in Mandarin Chinese]). For instance, if Dodo was holding a lot of toys and could not open the door, participants were asked to predict whether the character would open the door for Dodo or leave the door shut. In other words, participants not only had to remember whether each character had previously been nice or mean, but they also needed to use that trait information to predict whether the character would behave in a nice or mean way in a new situation. Effectively, children were invited to make behavior-to-behavior inferences that bridged the three phases of the experiment. See Appendix B for the full list of behavioral prediction questions.

Second, instead of having to recall the previous behavior of each character with no supportive cues, participants received a forced-choice *behavioral recall* task: They received two options of the same valence, one containing the behavior in which the character previously engaged and one containing a different behavior in which another character had engaged (e.g., "Did this child pull on the dog's tail, or did she/he kick her/his brother's blocks?"). See Appendix B for the full list of behavioral recall questions. For both behavioral prediction and behavioral recall, participants were asked to verbalize their choice between the two options.

Finally, to check whether children's performance on either the behavioral prediction task or the behavioral recall task was influenced by performance on the other task, the order of these two tasks was counterbalanced across participants. Thus, six subgroups of children were tested in total: a) children in a *systematic prompt* subgroup who completed the *behavioral prediction* task followed by the *behavioral recall* task; b) children in a *systematic prompt* subgroup who completed the *behavioral recall* task followed by the *behavioral prediction* task; c) children in a *generic prompt* subgroup who completed the *behavioral prediction* task followed by the *behavioral recall* task; d) children in a *generic prompt* subgroup who completed the *behavioral recall* task followed by the *behavioral prediction* task; e) children in a *no prompt* subgroup who completed the *behavioral prediction* task followed by the *behavioral recall* task; and finally, f) children in a *no prompt* subgroup who completed the *behavioral recall* task followed by the *behavioral prediction* task.

Results

In the initial trait inference phase, children in the systematic prompt group easily and accurately categorized the characters as nice or mean based on the behavioral vignettes. Unexpectedly, most of the children in the generic prompt group and even some children in the no prompt group spontaneously stated who was nice and who was mean, even though they had not been asked to do so. All participants differentiated the previously seen faces from new faces with little difficulty. Finally, participants were able to identify the affiliated behavior in the final trait memory phase. Participants also performed well on the behavioral prediction task. In the following paragraphs, we report these findings in more detail.

Phase 1: Trait Inference Phase

In the systematic prompt condition, 25 of 26 children (96%) were correct on all eight trials; only one participant made a mistake and sorted seven out of the eight characters correctly.

In the generic prompt condition, 24 of 26 participants responded to each of the behavioral vignettes, even without prompting from the experimenter to categorize the characters in the vignettes according to the inferred traits. Of these participants, 19 children (79%) spontaneously generated an evaluative label that was consistent with the behavior described in the story for all eight stories (14 children used the terms *nice* and *mean*; the 5 other children used terms that were similar in valence, e.g., *right* and *wrong, good* and *not good, happy* and *angry*). Three children responded nonverbally: One child used hand gestures, correctly signaling an "O" for the four nice characters and an "X" for the four mean characters, and 2 children sorted the cards into two groups, 1 doing so correctly for all eight characters (i.e., putting the four nice characters in one group and the four mean characters in a separate group) and 1 doing so correctly for six of the characters. One child generated a wrong label (saying "good" after hearing the story for the mean White boy), and 1 child responded with "I don't know" after each of the eight vignettes. The remaining 2 children made no detectable response. In sum, despite receiving only a generic, initial prompt, the majority of the children in the generic prompt group made an overt response to each of the stories they heard, and their responses were almost always consistent with the valence of the trait implied by the behavior described.

Finally, in the no prompt condition, 4 of 22 participants (18%) generated labels in response to each of the behavioral vignettes without any prompting from the experimenter. One participant correctly labeled all of the nice characters and mean characters as *right* and *wrong*, respectively, while sorting them into two separate groups. Another participant commented at the end of the first phase that some of the children did good things and were good-hearted people whereas other children did bad things and were bad-hearted people. Two participants responded to select card characters; 1 participant responded with "Great!" for three of the card characters (nice Asian girl, nice Asian boy, and nice White boy), while 1 participant nodded or shook her head appropriately for four of the card characters (nice Asian girl, mean Asian girl, mean Asian boy, and mean White girl) and responded with, "Not good" for one character (mean White boy).

Phase 2: Face Recognition Phase

In Phase 2, we examined participants' ability to recognize the faces they had previously seen and to distinguish them from a set of eight new faces (maximum score for each stimuli type = 8 faces categorized correctly). One participant was removed due to experimenter error. We ran a repeated-measures ANCOVA, with stimuli type (old face, new face) as the within-subjects variable, participant gender (male, female) and prompt type (systematic, generic, none) as the between-subjects variables, and number of correct responses as the dependent variable. Age (in months) was included as a covariate. Our analysis revealed only a main effect of age, $F(1, 66) = 11.85$, $p = .001$, $\eta^2_{par} = .15$. Separating participants into two groups based on the median age (63.00 months [5;3]), we found that older children ($M = 14.45$, $SD = 1.36$) were better at correctly categorizing faces compared with younger children ($M = 12.70$, $SD = 2.91$), $t(71) = 3.39$, $p = .001$, $d = 0.80$. Regardless of age, participants were significantly above 50% chance in correctly identifying the old faces ($M = 6.78$, $SD = 1.32$), $t(72) = 18.07$, $p = .001$, *d* = 4.26, and the new faces (*M* = 6.88, *SD* = 1.97), *t*(72) = 12.47, *p* = .001, *d* = 2.94. There was no significant main effect of stimuli type, *F*(1, 66) = 2.18, *p* = .14, nor of prompt type, *F*(2, 66) = 0.44, *p* = .64. Thus, the experimenter's explicit request to state the trait label for each story in the trait inference phase did not enhance (or undermine) children's subsequent recognition of the faces they had previously seen.

Phase 3: Trait Memory Phase

All participants completed two tasks in Phase 3: the behavioral prediction task and the behavioral recall task. Overall, participants scored significantly above chance (i.e., 50%, or correctly on four out of eight trials) for both the behavioral prediction (*M* = 5.04, *SD* = 1.29), *t*(73) = 6.96, *p* = .001, *d* = 1.63, and behavioral recall tasks (*M* = 4.96, *SD* = 2.27), *t*(73) = 3.64, *p* = .001, *d* = 0.85. Figure 2 offers a more detailed comparison of performance in the behavioral prediction and behavioral recall tasks by showing the mean number of: a) correct responses, b) incorrect responses, and c) "don't know" responses for each task. Inspection of Figure 2 shows that although the mean number of correct responses did not differ by task, the pattern of error did vary. Children were significantly more likely to make an incorrect response for the behavioral prediction task (*M* = 2.78, *SD* = 1.20) than the behavioral recall task (*M* = 1.54, *SD* = 1.56), *t*(73) = 5.54, *p* < .001, *d* = 1.30, whereas they were significantly less likely to make a "don't know" response for the behavioral prediction task (*M* = 0.18, *SD* = 0.98) than the behavioral recall task (*M* = 1.50, *SD* = 2.08), *t*(73) = 5.92, *p* < .001, *d* = 1.39.

To explore children's performance across both tasks, we first conducted a repeated-measures ANCOVA with prompt (systematic, generic, none), task order (behavioral recall first, behavioral prediction first), and participant gender as between-subjects variables and task (behavioral recall, behavioral prediction) as the within-subjects variable on the arcsine square root of the total number of correct answers, with participant age (in months) included as the covariate. This analysis revealed a main effect of age, *F*(1, 61) = 4.93, *p* = .03, η^2_{par} = .08, and of task, *F*(1, 61) = 5.07, *p* = .03, η^2_{par} = .08, as well as an Age × Task interaction, *F*(1, 61) = 5.47, *p* = .02, η^2_{par} = 08. No other effects or interactions were found.

To examine the Age × Task interaction further, we separated participants into two groups based on the median age (63.00 months [5;3]). See Figure 3 for a depiction of the results. In the behavioral prediction task, older children (*M* = 5.25, *SD* = 1.15) displayed levels of correct prediction similar to those of younger children (*M* = 4.79, *SD* = 1.41), *t*(72) = 1.53, *p* = .13. By contrast, in the behavioral recall task, older children performed significantly better than younger children (older children, *M* = 5.58, *SD* = 1.87; younger children, *M* = 4.24, *SD* = 2.50), *t*(72) = 2.64, *p* = .01, *d* = 0.62. Paired-sample *t* tests revealed no significant differences in performance on the two tasks for both younger children, *t*(33) = 1.30, *p* = .20, and older children, *t*(39) = 1.14, *p* = .26. However, children younger than 63 months (5;3) scored above chance level (= 4 behaviors selected correctly) in the behavioral prediction task, *t*(33) = 3.29, *p* = .002, *d* = 1.15, and performed no better than chance on the behavioral recall task, *t*(33) = 0.55, *p* = .59, whereas older children performed above chance level on both the behavioral prediction task, *t*(39) = 6.88, *p* = .001, *d* = 2.20, and the behavioral recall task, *t*(39) = 5.34, *p* = .001, *d* = 1.71. Children's performance did not vary depending on the order of the two tasks. As in Phase 2, whether or not participants were prompted to offer a trait label in the first phase made no difference to their performance on either task type. As a further check on

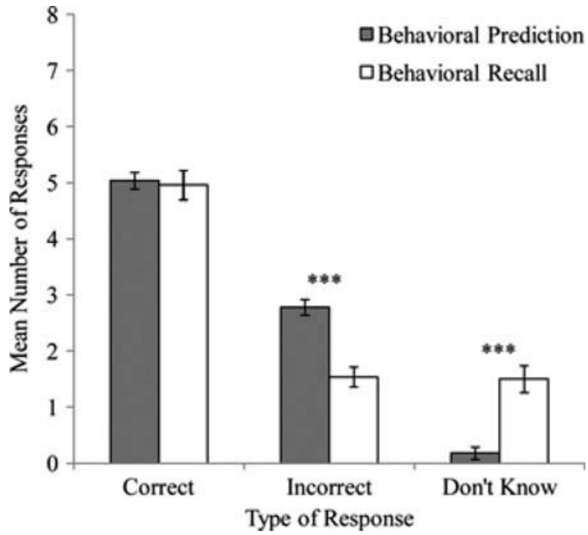

FIGURE 2. Comparison of the types of responses for the behavioral prediction and behavioral recall tasks in Phase 3 of Experiment 2. ***$p < .001$.

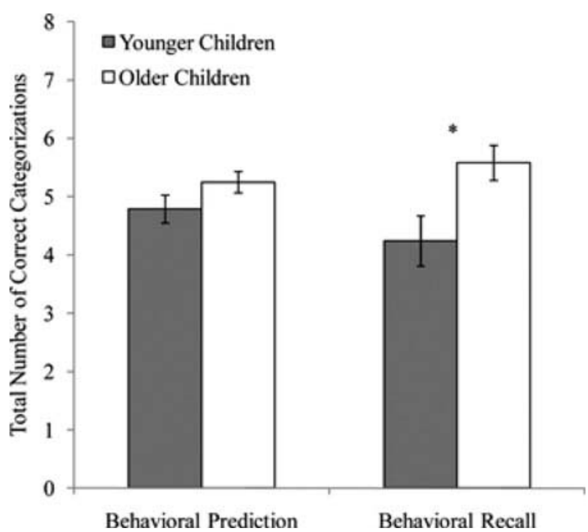

FIGURE 3. Comparison of the total number of correct categorizations for younger and older children in the behavioral prediction and behavioral recall tasks in Phase 3 of Experiment 2. *$p < .05$.

this latter result, we compared the proportion of children obtaining a relatively high score (i.e., six, seven, or eight correct) across the three prompt groups. Chi-square tests revealed no significant difference among the three groups for either the behavioral prediction task, $\chi^2(2) = 4.12$, $p = .12$, or the behavioral recall task, $\chi^2(2) = 2.79$, $p = .25$.

Discussion

The findings in Experiment 2 replicated two of the main findings in Experiment 1. First, children demonstrated a robust ability to infer nice and mean traits from descriptions of characters' behavior. Second, children could also easily distinguish the faces of the original characters from the faces of new characters. Experiment 2 also revealed several new findings. First, although children in the generic prompt and the no prompt groups received no systematic trait prompting, a number of them spontaneously made inferences about the vignette characters. They often said that the character was nice or mean or dichotomized the characters in a valenced fashion (e.g., right vs. wrong or good vs. not good). Indeed, the majority of children in the generic prompt group and even a few of the children in the no prompt group engaged in this overt pattern of judgment after every story despite the absence of any systematic prompting from the interviewer. By implication, young children readily adopt an evaluative stance by judging characters either positively or negatively.

Children's systematic responding in the behavioral prediction task of the final phase of the study shows that they are also capable of making behavior-to-behavior inferences (Liu et al., 2007). Not only were children able to infer a trait in the initial phase, but they were able to infer a future behavior consistent with that trait in the final phase. Moreover, this inferential ability was not affected by the instructions that children had received in the initial phase. Whether they received systematic story-by-story prompting to make a trait inference, a single generic prompt at the outset, or no prompt at all, children were able to predict a future behavior consistent with the trait inferred in the original vignette.

Finally, when the difficulty of the behavioral recall task was reduced (i.e., it involved a forced choice between two behavioral descriptions rather than the recall of a particular vignette), children, especially older children, proved able to remember the previous behaviors of the characters. Moreover, children's level of accuracy was equivalent across the behavioral prediction and behavioral recall tasks. These findings suggest that the gap observed in Experiment 1 between children's retention of the story character's trait as compared with their recall of the story character's behavior in the vignette is best interpreted in terms of the difficulty children faced when trying to recall and narrate a given vignette in the absence of any cues or reminders other than the character's face.

The results from Experiments 1 and 2 suggest that children are able to infer basic traits from behavioral information and can do so even when no systematic prompting is provided. These findings support the notion that memory for the behaviors, and presumably for the inferred traits, as well as accuracy in making predictions of future behavior do not rely on any prior overt statement of the relevant trait terms. Consistent with the proposal that children readily make evaluative trait attributions, children went on to make behavior-to-behavior inferences no matter what type of prompt they had received in the initial phase.

GENERAL DISCUSSION

The ability to attribute traits to unfamiliar informants was examined across three different phases in each of two experiments. Specifically, we investigated whether young children are able to infer trait information from behavioral descriptions, remember the faces they have seen, and retain the information associated with those stimulus faces at a later point in time when asked to recall what behavior the story character had displayed previously and what behavior the character might display in the future. Finally, we checked for potential cultural differences and similarities in trait attribution across two separate cultures in the first experiment and examined these person perception abilities further with Taiwanese children in the second experiment.

Overall, we found few differences between the two groups of children in Experiment 1. Only one notable difference was apparent in the participants' recall of stories during the trait memory phase: Although both groups of participants recalled very few of the stories they had heard in the initial trait inference phase, the Taiwanese participants generated fewer stories than the European American children.

The reasons for this cultural difference call for further investigation. One plausible interpretation is that there are cross-cultural differences in patterns of recall. European American mothers tend to elaborate on their children's recall and to actively encourage their children to continue the conversation (Fivush, Haden, & Reese, 2006; Nelson & Fivush, 2004). By contrast, Chinese mothers prefer to focus on repetition and direct their children's responses. In response, 3-year-old European American children are more likely to provide memory information and elaborative responses compared with their Chinese peers (Wang, Leichtman, & Davies, 2000). This pattern was echoed in Experiment 1. In the trait memory phase at the end of the experiment, children were invited to recall the vignette they had heard for each of the faces presented in the initial trait inference phase. The European American children generated more vignettes—both correct and incorrect—than did Taiwanese children. Thus, Taiwanese children were more likely than European American children to fail to generate any vignette—whether correct or incorrect.

In sum, the degree to which children generate a story or justify their answers may vary depending on the culture (e.g., European American, East Asian) in which they are raised; furthermore, results from Experiment 2 indicate that this difference may stem more from variation in children's willingness to generate stories rather than in their ability to remember the original behaviors. But aside from this difference, the person perception tendencies of Taiwanese and European American children were strikingly similar.

In the initial trait inference phase, across both cultures and both experiments, young children were remarkably good at inferring traits based only on descriptions of the characters' behavior and sorted the stimuli into the appropriate categories with near-perfect accuracy. Unexpectedly, in Experiment 2, a majority of the children in the generic prompt condition and some children in the no prompt condition spontaneously generated traits that were consistent with the valenced traits used by children in the systematic prompt condition. Taken together, these results confirm the earlier findings of Liu et al. (2007) and show that young children are able to infer traits from behavioral descriptions. The results go beyond those earlier findings, however, by showing that this pattern of inference is found among young Taiwanese children as well as American children. Indeed, contrary to expectation (Na & Kitayama, 2011), the Taiwanese children tested in Experiment 2 spontaneously inferred such traits even when they were not given a systematic prompt to do so. Across both experiments, children were also remarkably accurate in the face memory phase, distinguishing between the stimulus faces to which they had been previously exposed and the faces shown to them for the first time. Finally, in the last phase, participants in Experiment 1 were able to re-sort the original stimuli into the correct trait categories even after a time lapse. Both older and younger participants in Experiment 2, regardless of the prompt type they received, systematically predicted likely future behaviors and older participants also systematically recognized past behaviors. In the following paragraphs, we discuss two notable findings in more detail: a) children's relatively poor memory for the behavioral vignettes in Experiment 1, especially as compared with their memory for the traits that they inferred and b) children's accurate predictions of future behaviors in Experiment 2.

Participants were good at inferring and retaining trait information about unfamiliar characters. Nevertheless, a considerable gap emerged in Experiment 1 between children's retention of the trait category and their recall of the behavior that had led to that categorization. In both the United States and Taiwan, the participants were often unable to generate any story at all for a given face and yet they accurately categorized the face as belonging to a nice or mean child. These results resemble those of previous studies with adults showing a dissociation between memory for an inferred trait and memory for the behavioral source of that trait, among both healthy adults (Anderson & Hubert, 1963; Todorov & Uleman, 2002) and patients with amnesia (Johnson, Kim, & Risse, 1985; Tranel & Damasio, 1993).

One aim of Experiment 2 was to explore a possible reason for the gap between the ability to categorize previously seen faces and the ability to recall more detailed behavioral information related to those faces: the difference in complexity between recognition and recall. Researchers have suggested that explicit recollection is more complex than recognition, and as a result, accurate reproduction of previously learned facts is rarer than accurate recognition (Bartlett, 1995; Kintsch, 1974; Kintsch & van Dijk, 1978). Similarly, recognizing a face as belonging to either a nice or mean person may be easier than recalling the behavior from which the trait was inferred. In sum, the two levels of information (the behavioral details and the trait information) have different degrees of staying power—the details get lost, but the gist remains.

Accordingly, in Experiment 2, we asked participants to respond in a forced-choice behavioral recall task, using a format that was comparable to the forced-choice behavioral prediction task. When task difficulty was equated in this way, no differences were found between the performances on these two tasks for older or younger children. However, children younger than 63 months (5;3) were able to make behavioral predictions systematically but did no better than chance in the behavioral recall task. It is possible that the ability to recall previous behaviors requires more time for children to master. Another possibility is that perhaps the behavioral recall task, which required children to choose from two behaviors they had heard of before (albeit in connection with different faces; see Appendix B), was more difficult than the behavioral prediction task, which required children to pick either a positive behavior or a negative behavior. Regardless, the type of task (behavioral prediction vs. behavioral recall) overall had no effect on children's responses. These results demonstrate that a) older preschool children are capable of remembering prior behavioral information, provided their memory is probed via a forced-choice recognition task rather than via open-ended recall; and b) younger children may find it more challenging to recollect the original behaviors of the characters than to predict future behavior based on the inferred traits of the characters.

Turning to the second notable finding, the children tested in Experiment 2 successfully predicted future behavior on the basis of the traits that they had previously inferred. Moreover, they did so whether they had been given a systematic prompt after each story (i.e., to categorize the protagonist as either nice or mean), a generic prompt before any of the stories, or no prompt at all. These results go against a relatively long-standing consensus that preschool children do not make behavior-to-behavior predictions by means of an intervening trait inference (Rholes & Ruble, 1984). They are also inconsistent with the findings from Liu et al. (2007): When 4-, 5-, and 7-year-olds were invited to make behavior-to-behavior predictions, they performed at chance even though all three age groups proved able to infer traits from behavior and to predict future behavior from traits. By contrast, the preschoolers here were relatively sophisticated in their ability to make behavior-to-behavior inferences.

Why were preschoolers successful in Experiment 2, contrary to results from previous research? One line of explanation builds on a suggestion by Liu et al. (2007). They noted that when a trait label is generated, it is likely to serve as a placeholder that calls children's attention to the predictability of future behaviors. Building on this insight, it is plausible that whether or not they are prompted to overtly verbalize a trait label, young children spontaneously generate an inference about a protagonist based on the behavior presented. These results show that the trait-processing ability of preschool children has been underestimated; even when left to their own devices, they can make behavior-to-behavior predictions.

However, in this connection, it is important to note a limitation of both experiments. Children's attribution of only one trait dimension—nice versus mean—was measured. This dimension was deliberately chosen because the trait terms in question are likely to be known to preschool children, including 3- and 4-year-olds. It is possible that spontaneous attributions like those observed in Experiment 2 are especially likely if the behavior in question invites a global, positive evaluation or global, negative evaluation. Spontaneous attributions may be less likely to occur if the behavior invites a more specific evaluation—for example, an evaluation of strength or smartness—or is relatively neutral in terms of evaluation. Accordingly, in future research, it will be informative to assess whether the same pattern of findings emerges when different traits are included and also when children younger than 3 years of age are examined.

Overall, our findings show that the ability to make trait attributions emerges at a young age. Across two different cultures, children demonstrated similarly robust trait attribution abilities. First, children readily infer trait labels from behavioral descriptions of unfamiliar individuals. Second, they not only recognize these unfamiliar individuals at a later point, but they also remember the trait that was originally attributed to those individuals. Yet if asked to recall the behavior on which the trait attribution was based, children are often unable to recall the particular behavior displayed by an individual. Finally, and especially noteworthy in relation to previous findings, young children can successfully predict an individual's future behavior on the basis of the trait that they have attributed.

ACKNOWLEDGMENTS

The authors thank Mahzarin R. Banaji for comments on earlier drafts of the article. We also thank Yi-ting Chen, Catherine Glennon, Martha Jane Jenkins, Jennifer Zinser, and Cynthia Zwicky for assistance in data collection.

FUNDING

This research was supported by an Initiation Grant from The Hong Kong University of Science and Technology.

REFERENCES

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, *111*, 256–274. doi:10.1037/0033-2909.111.2.256

Anderson, N. H., & Hubert, S. (1963). Effects of concomitant verbal recall on order effects in personality impression formation. *Journal of Verbal Learning and Verbal Behavior*, *2*, 379–391. doi:10.1016/S0022-5371(63)80039-0

Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology*, *41*, 258–290. doi:10.1037/h0055756

Bartlett, F. C. (1995). *Remembering: A study in experimental and social psychology.* New York, NY: Cambridge University Press.

Boseovski, J. J., & Lee, K. (2006). Children's use of frequency information for trait categorization and behavioral prediction. *Developmental Psychology*, *42*, 500–513. doi:10.1037/0012-1649.42.3.500

Bretherton, I., & Beeghly, M. (1982). Talking about internal states: The acquisition of an explicit theory of mind. *Developmental Psychology*, *18*, 906–921. doi:10.1037/0012-1649.18.6.906

Carlston, D. E., & Skowronski, J. J. (1994). Savings in the relearning of trait information as evidence for spontaneous inference generation. *Journal of Personality and Social Psychology*, *66*, 840–856. doi:10.1037//0022-3514.66.5.840

Choi, I., Nisbett, R. E., & Norenzayan, A. (1999). Causal attribution across cultures: Variation and universality. *Psychological Bulletin*, *125*, 47–63. doi:10.1037/0033-2909.125.1.47

Cogsdill, E. J., Todorov, A. T., Spelke, E. S., & Banaji, M. R. (2014). Inferring character from faces: A developmental study. *Psychological Science*, *25*, 1132–1139. doi:10.1177/0956797614523297

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition*. New York, NY: McGraw-Hill.

Fivush, R., Haden, C. A., & Reese, E. (2006). Elaborating on elaborations: Role of maternal reminiscing style in cognitive and socioemotional development. *Child Development*, *77*, 1568–1588. doi:10.1111/j.1467-8624.2006.00960.x

Gilbert, D. T. (1998). Ordinary personology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., pp. 89–150). New York, NY: McGraw Hill.

Gnepp, J., & Chilamkurti, C. (1988). Children's use of personality attributions to predict other people's emotional and behavioral reactions. *Child Development*, *59*, 743–754. doi:10.2307/1130573

Hamilton, D. L., Katz, L. B., & Leirer, V. O. (1980). Organizational processes in impression formation. In R. Hastie, T. M. Ostrom, E. B. Ebbesen, R. S. Wyer, Jr., D. L. Hamilton, & D. E. Carlston (Eds.), *Person memory: The cognitive basis of social perception* (pp. 89–119). Hillsdale, NJ: Lawrence Erlbaum Associates.

Heyman, G. D., & Gelman, S. A. (1999). The use of trait labels in making psychological inferences. *Child Development*, *70*, 604–619. doi:10.1111/1467-8624.00044

Heyman, G. D., & Gelman, S. A. (2000). Preschool children's use of trait labels to make inductive inferences. *Journal of Experimental Child Psychology*, *77*, 1–19. doi:10.1006/jecp.1999.2555

Johnson, M. K., Kim, J. K., & Risse, G. (1985). Do alcoholic Korsakoff's syndrome patients acquire affective reactions? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 22–36. doi:10.1037/0278-7393.11.1.22

Jones, E. E. (1979). The rocky road from acts to dispositions. *American Psychologist*, *34*, 107–117. doi:0.1037//0003-066x.34.2.107

Jones, E. E., & Nisbett, R. E. (1972). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 79–94). Morristown, NJ: General Learning Press.

Kintsch, W. (1974). *The representation of meaning in memory.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*, 363–394. doi:10.1037//0033-295x.85.5.363

Kuwabara, M., Son, J. Y., & Smith, L. B. (2011). Attention to context: U.S. and Japanese children's emotional judgments. *Journal of Cognition and Development*, *12*, 502–517. doi:10.1080/15248372.2011.554927

Liu, D., Gelman, S. A., & Wellman, H. M. (2007). Components of young children's trait understanding: Behavior-to-trait inferences and trait-to-behavior predictions. *Child Development*, *78*, 1543–1558. doi:10.1111/j.1467-8624.2007.01082.x

Livesley, W. J., & Bromley, B. D. (1973). *Person perception in childhood and adolescence*. London, England: Wiley.

Lockhart, K. L., Nakashima, N., Inagaki, K., & Keil, F. C. (2008). From ugly duckling to swan? Japanese and American beliefs about the stability and origins of traits. *Cognitive Development*, *23*, 155–179. doi:10.1016/j.cogdev.2007.08.001

Miller, J. G. (1984). Culture and the development of everyday social explanation. *Journal of Social and Personality Psychology*, *46*, 961–978. doi:10.1037//0022-3514.46.5.961

Miller, J. G. (1986). Early cross-cultural commonalities in social explanation. *Developmental Psychology*, *22*, 514–520. doi:10.1037//0012-1649.22.4.514

Na, J., & Kitayama, S. (2011). Spontaneous trait inference is culture-specific: Behavioral and neural evidence. *Psychological Science*, *22*, 1025–1032. doi:10.1177/0956797611414727

Nelson, K., & Fivush, R. (2004). The emergence of autobiographical memory: A social cultural developmental theory. *Psychological Review*, *111*, 486–511. doi:10.1037/0033-295X.111.2.486

Newman, L. S. (1991). Why are traits inferred spontaneously? A developmental approach. *Social Cognition*, *9*, 221–253. doi:10.1521/soco.1991.9.3.221

Pronin, E., & Ross, L. (2006). Temporal differences in trait self-ascription: When the self is seen as an other. *Journal of Personality and Social Psychology*, *90*, 197–209. doi:10.1037/0022-3514.90.2.197

Rholes, W. S., & Ruble, D. N. (1984). Children's understanding of dispositional characteristics of others. *Child Development*, *55*, 550–560. doi:10.2307/1129966

Soren, B. J. (2009). A participatory model for integrating cognitive research into exhibits for children: Summative evaluation final report. Unpublished manuscript, Museum of Science, Boston, MA.

Todorov, A., Gobbini, M. I., Evans, K. K., & Haxby, J. V. (2007). Spontaneous retrieval of affective person knowledge in face perception. *Neuropsychologia*, *45*, 163–173. doi:10.1016/j.neuropsychologia.2006.04.018

Todorov, A., & Uleman, J. S. (2002). Spontaneous trait inferences are bound to actors' faces: Evidence from a false recognition paradigm. *Journal of Personality and Social Psychology*, *83*, 1051–1065. doi:10.1037/0022-3514.83.5.1051

Todorov, A., & Uleman, J. S. (2004). The person reference process in spontaneous trait inferences. *Journal of Personality and Social Psychology*, *87*, 482–493. doi:10.1037/0022-3514.87.4.482

Tranel, D., & Damasio, A. R. (1993). The covert learning of affective valence does not require structures in hippocampal system or amygdala. *Journal of Cognitive Neuroscience*, *5*, 79–88. doi:10.1162/jocn.1993.5.1.79

Wang, Q., Leichtman, M. D., & Davies, K. I. (2000). Sharing memories and telling stories: American and Chinese mothers and their 3-year-olds. *Memory*, *3*, 159–177. doi:10.1080/096582100387588

Appendix A

Behavioral vignettes used in Experiment 1 (similar vignettes were used in Experiment 2).

Nice:

1.This child is at his class's show-and-tell. He is showing his friends his new stuffed bear. One of his friends asks if she can hold the bear. This child shares his new bear with his friend.
2.This child is playing with her toys. Her mommy says it is time for dinner. She cleans up right away and helps her mommy set the table for dinner.
3.This child is walking in his neighborhood. He sees an old woman is having a lot of trouble carrying a big box. This child carries the box for the old woman to her front door.
4.This child is shopping at the grocery store. She is buying some candy. A boy accidentally knocks over some cereal boxes. This child helps him pick up the boxes and put them back on the shelf.

Mean:

1.This child is watching TV. His dog is also in the room. This little boy pulls the dog's tail very hard, and the dog runs away.
2.This child is going up the stairs at home. Her brother left his blocks in the stairway. This little girl kicks his blocks down the stairs and walks to her room.
3.This child is walking down a hall. A little girl nearby trips and drops her books. This little boy laughs and points at her.
4.This child is playing in the sandbox. A little boy is filling his bucket with sand. This little girl grabs the bucket without asking and starts playing with it herself.

Appendix B

Behavioral recall and prediction questions used in Experiment 2.

Behavioral recall questions:

1. Did this child share his/her bear with a friend, or did he/she help set the table?
2. Did this child help his/her mommy set the table, or did he/she carry a box for an old woman?
3. Did this child carry a box for an old woman, or did he/she pick up boxes that fell off a shelf?
4. Did this child help pick up cereal boxes, or did he/she share his/her bear with a friend?
5. Did this child pull his/her dog's tail, or did he/she kick his/her brother's blocks?
6. Did this child kick his/her brother's blocks, or did he/she laugh and point at a little girl?
7. Did this child laugh and point at a little girl, or did he/she grab a little boy's bucket?
8. Did this child grab a little boy's bucket without asking, or did he/she pull his/her dog's tail?

Behavioral prediction questions:

1. It is Dodo's first day of school. Do you think this little boy/girl will go up to say hello to Dodo or ignore Dodo instead?
2. Dodo is holding a lot of toys and cannot open the door. Do you think this little boy/girl will open the door for Dodo or leave the door closed?
3. It is snack time, and Dodo is eating snacks alone. Do you think this little boy/girl will sit with Dodo at snack time or sit somewhere else?
4. It is naptime, and Dodo needs help spreading out the sleeping bag. Do you think this little boy/girl will help Dodo with the sleeping bag or will he/she just spread his/her own sleeping?
5. At recess, the whole class goes outside to play. Do you think this little boy/girl will play with Dodo on the playground or leave Dodo to play alone?
6. The teacher asks children to read storybooks together. Do you think this little boy/girl will let Dodo pick the storybook or will he/she not let Dodo pick the storybook?
7. This little boy/girl has brought a toy from home. Do you think he/she will share this toy with Dodo or play with the toy by himself/herself?
8. It is the end of the school day, and everyone is getting ready to go home. Do you think this little boy/girl will go and say goodbye to Dodo or will he/she go home without saying goodbye to Dodo?