



ELSEVIER

Available online at www.sciencedirect.com

 ScienceDirect

Journal of Research in Personality 40 (2006) 652–674

JOURNAL OF
RESEARCH IN
PERSONALITY

www.elsevier.com/locate/jrp

Personality perception: A developmental study

Amber Renee McLarney-Vesotski ^{a,*}, Frank Bernieri ^b,
Daniel Rempala ^c

^a *Alpena Community College, Department of Social Science, 666 Johnson St., Alpena, MI 49707, USA*

^b *Oregon State University, USA*

^c *University of Toledo, USA*

Available online 1 September 2005

Abstract

We investigated the ability of adults and children to perceive the Big Five personality traits in others from “thin slices” of the behavioral stream. Targets were videotaped in five carefully crafted situations, each designed to maximally reveal a different trait domain. In addition, each target’s personality was assessed by close friends or parents, serving as the personality criterion. Judges’ agreement with the criterion improved generally with age, but different developmental trends were observed across traits.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Personality; Impression formation; Social cognition; Personality perception; Big Five personality traits; Personality traits; Judgment; Social perception; Interpersonal sensitivity

1. Introduction

Each day we encounter and assess those we meet within minutes, perhaps even seconds. Investigators (see Funder, 1999; Hall & Bernieri, 2001, for reviews) have researched the personality perception in these brief encounters, but research involving children’s perception of personality has been more limited. This investigation

* Corresponding author.

E-mail address: vesotska@alpenacc.edu (A.R. McLarney-Vesotski).

developed a research paradigm to measure first impressions of personality operationalized by the Five Factor Model (Costa & McCrae, 1985) in both children and adults.

Adults have been shown to read accurately aspects of another's personality within seconds (Ambady, Bernieri, & Richeson, 2000; Borkenau & Liebler, 1992; Funder & Sneed, 1993; Kenny, 1994). Nonverbal cues such as facial expressions, gaze, interpersonal space, touch, and gesture (Depaulo & Friedman, 1997) are used to read others' personality (Buss & Craik, 1983; Funder & Colvin, 1991; Linkey & Firestone, 1990; Scherer, 1978). Funder and Sneed (1993), for example, examined how the Big Five traits of personality were reflected in expressive behaviors during an open-ended conversation. This study and others like it (Bem & Funder, 1978; McCrae, Costa, & Busch, 1986; McCrory, Bernieri, Grahe, & Gada-Jain, 2003) have concluded that there are indeed expressive behaviors relevant to certain personality characteristics. Personality appears to be chronically embedded within one's behavioral stream (Allport, 1937; Ambady et al., 2000).

Thin slices of the behavioral stream have been shown to reveal information about affective states (e.g., Waxer, 1976), deception (DePaulo, Lassiter, & Stone, 1982), interpersonal goals (Richeson & Ambady, 2001), and even personality (e.g., Albright, Kenny, & Malloy, 1988). For example, Dabbs, Bernieri, Strong, Campo, and Milun (2001) taped 148 participants entering a room, walking over to greet a female experimenter, and taking a seat for an interview. The first 30s of this process was then shown to observers who were asked to rate the targets on each of the Big Five traits (Costa & McCrae, 1985). Judgments of extraversion, agreeableness, openness, and conscientiousness correlated significantly with targets' psychometrically assessed traits.

A meta-analysis by Ambady and Rosenthal (1992) found that thin slices of behavior permit significantly accurate predictions. In fact, Ambady and Rosenthal (1992) reported that the level of accuracy of predictions from thin slices of behavior does not differ from the accuracy reported in such classic studies as the Office of Strategic Services (OSS) assessment study in 1948 and the Veterans Administration (VA) assessment project conducted between 1946 and 1949 (as summarized and reported in Wiggins, 1973) that were based on voluminous amounts of testing data, clinical interviews and extended behavioral evaluations. This result contrasts with our intuition that more information leads to greater accuracy.

How can judgments from thin slices be so accurate? David Funder (1995, 1999, 2001) has been developing a model of social judgment accuracy. His model makes clear that the diagnostic utility of the information decoded by the observer, not its quantity, determines judgment accuracy.

Funder (1995, 1999) described the process of reading others accurately in his Realistic Accuracy Model (RAM). Funder (1995) made explicit the obvious fact that the target person has to generate expressive cues that are both available to the observer and diagnostically relevant to the trait being assessed. The judge must then be able to infer a trait associated with these observed cues, and correctly link the behavior to its underlying trait. Much of personality perception accuracy research (reference needed here) has employed a limited number of behavioral contexts (e.g., zero acquaintance). Meeting or interacting with a stranger in a getting to know you situation seems most

relevant to the sociability traits of extraversion and agreeableness. Not surprisingly, in accordance with Funder's model, extraversion has been the trait most accurately perceived, followed by agreeableness, and conscientiousness with neuroticism and openness showing little to no accuracy (see [Kenny, 1994](#), for a review). It is important to interpret the current archive of personality perception accuracy results taking in to account the traits being assessed and the relevance of the social contexts within which these traits are being judged. For example, we would argue that judging conscientiousness from thin slice samples of targets vacuuming their bedroom would be more reasonable than assessing the judgment accuracy of extraversion in that particular nonsocial context.

The paradigm created for this study addressed explicitly the situation relevancy highlighted in the first step of RAM as it pertained to thin-slice stimuli. To better assess trait judgments based on thin-slices, we attempted to increase the trait relevancy of the situations video taped. To accomplish this, we developed a carefully structured, multi-staged, but brief experimental session where each stage provided a context to allow one or more traits to be expressed.

[Depaulo \(1991\)](#) has discussed how nonverbal cue detection develops over the course of a life span. Depaulo's discussion of nonverbal cue understanding implies that age 5 is an important time for understanding the perception of nonverbal behavior. At this age, children begin to realize that expressive behavior reveals internal states ([Eisenberg, Murphy, & Shepard, 1997](#); [Harris, Donnelly, Guz, & Pitt-Watson, 1986](#); [Losoya & Eisenberg, 2001](#)). Also during this time period, children develop a "theory of mind." At this point, children should be capable of using their ability to perceive others' personality ([Saarni, 1990](#)). Once children understand that internal states are revealed by expressive behavior: (a) they are capable of finding reasons why internal states do not always match, (b) they are capable of understanding that people are unique, and (c) they are capable of knowing that people can manipulate their nonverbal behavior to convey what they want to convey ([Gnepp & Hess, 1986](#)). Once norms and expressive behavior control are well understood by age 10 ([Gnepp & Hess, 1986](#); [Nowicki & Duke, 2001](#)), a child is equipped to interpret what they are seeing in a manner similar to adults. Of course, the more knowledge one accumulates through experience about the various traits and social norms that exist the closer to reality the interpretation should be.

Some traits are more directly (i.e., correspondingly) manifest than others ([Kenny, 1994](#)). This is supported by the Class Play procedure ([Masten, Morison, & Pellegrini, 1985](#)). Children's descriptions of their peers' sociability often matched their teachers' perceptions. These descriptions included behaviors like "who is likely to be picked on." Extraversion is more directly manifest than some other traits. For example, extraverts are described as talkative ([Costa & McCrae, 1985](#)). Judging extraversion simply involves the perception of talkativeness when in the presence of others. A child does not have to infer how talkative someone is; the person either talks or does not. Therefore, children might be able to judge successfully the level of extraversion in others at an earlier age than say the level of conscientiousness, which involves a more sophisticated understanding of task norms and standards for completion. For example, a conscientious person is someone who is dependable and responsible

(Costa & McCrae, 1985). To assess dependability and responsibility, however, a child (or adult for that matter) must know the norms and standards that constitute such constructs. The judge must understand how context influences the expected norms for the relevant behaviors.

This general knowledge about a psychological construct is acquired through experience in a given domain (Baldwin, 1992). Schemas take time and varied experiences to develop (Baldwin, 1992). Well-developed schemas are important to thin-slice judgments because the context is limited and brief. In general then, the judgment of traits involving social norms, self-presentation, and deception are likely to develop later than traits like extraversion that have a series of direct and corresponding behaviors, such as talkativeness and interpersonal proximity (Depaulo, 1991; Jones, Abbey, & Cumberland, 1998).

Assessing the accuracy of trait perception turns out to be problematic because of the “criterion problem” (e.g., Bernieri, 2001; Funder, 1999; Kenny, 1994; Kruglanski, 1989). Simply put, in order to decide whether a judgment is accurate, you need to know what the judgment should be (i.e., the reality). For trait perception accuracy, one needs to *know* a target’s trait to use it as a criterion. Thus, trait accuracy can only be assessed if: (a) the trait itself is known to exist, (b) the trait is meaningfully expressed within a child of a given age, and (c) the trait’s level of expression or intensity can be known through some kind of assessment device. The reality of many psychological constructs (e.g., moods, beliefs, emotions, and personality) is that their expression and measurement are simply not known well enough to establish a true accuracy criterion (Kruglanski, 1989). Thus, no researcher ever truly assesses trait-perception accuracy for lack of a known criterion.¹

For this reason, terms such as “interpersonal sensitivity” and “agreement” are often employed over the term accuracy because their use does not assume an objectively known criterion (Hall & Bernieri, 2001). Bernieri (2001) defined the interpersonally sensitive person as one who can “perceive or otherwise respond appropriately to the internal states of another, understand the antecedents of those states, and predict the subsequent affective, cognitive, and behavioral events that will result (p. 3).” When one further defines “appropriately” as reflecting a general consensus of judgment with relevant others (e.g., the targets themselves, or other people who either interact with targets or might otherwise have some influence over their fate), the issue of trait perception accuracy is technically an issue of observer sensitivity or agreement. The challenge for researchers is to provide a compelling justification for the criterion used in terms of it being representative of a meaningful reality (Bernieri, 2001).

The approach taken in the present investigation involved the employment of a reasonable criterion generated by people who knew the targets best (i.e., parents and friends). For children, this meant their parents provided us with their trait criterion. For college students, their trait criterion was provided by either a close friend or significant other. In other words, they were people who would claim to “know them

¹ Clever researchers have developed operational criteria (Kenny, 1994) that allow for the measurement of some types of social perception accuracy. For example, one *can* assess the accuracy with which an observer judges whether a man and a woman are brother and sister or lovers because this is a knowable fact (e.g., Costanzo & Archer, 1989). But personality perception accuracy can not be so determined.

best.” We chose this criterion over self-report because we felt it would be less biased by self-presentation issues, and the verbal competency of children would no doubt affect the validity of such self-descriptions. We also chose informed others over test-driven data because of the possible discrepancy between the construct a particular test measured and the psychological construct the observer *thought* we had in mind. We thought this would be especially problematic in the children samples. By using reports generated by informed others we had the opportunity to use items in the judgment task that were similar to the ones employed in the criterion, thus reducing conceptual ambiguity. The results reported throughout this investigation report the degree to which our participants judged the targets in a manner that was similar to the way “informed others” described them.

Although we truly are interested in documenting how children become more accurate in personality perception as they mature, and although we believe we have chosen a trait criterion that best represents the reality of the targets’ trait manifestation in relation to their own social world, we acknowledge that the term “accuracy” applies loosely here. What we are measuring is the extent to which thin-slice observers of our targets are “seeing” the same person on-screen that is known to the targets’ own parents or close friends.

This study developed a paradigm for examining personality perception in children of different ages. In phase one, eight targets from each of three different age groups were videotaped in a series of activities. Each activity was created to maximize the relevance and visibility (i.e., availability of diagnostic cues) of the Big Five characteristics described by [Costa and McCrae \(1985\)](#). In phase two, thin slice videos of these activities were shown to naïve observers of matching ages. These “judges” rated the targets on items pertaining to the Big Five traits.

We hypothesized that agreement between the judge and the criterion for the trait extraversion should be well developed in our youngest children tested (8-year-olds) due to its relatively direct link to behavioral cues ([Funder & Colvin, 1991](#); [Rholes & Ruble, 1986](#)). In addition, [Cumberland-Li, Eisenberg, and Reiser \(2004\)](#) found that children could successfully rate how nice and friendly they thought other children were. These characteristics are also linked to agreeableness ([Goldberg, 1992](#)), and authors often present agreeableness as a behavioral tendency ([Tobin, Graziano, Vanman, & Tassinari, 2000](#)). Therefore, we made similar predictions for agreeableness (e.g., the more someone nodded affirmation and approval, the more they agreed and thus the more agreeable they may be). We hypothesized that judgments of openness and conscientiousness would show less agreement than extraversion and agreeableness at younger ages because of the required understanding of norms and complex social contexts, which according to [Depaulo and Rosenthal \(1986\)](#), [Depaulo \(1991\)](#), and [Jones et al. \(1998\)](#) develops later.

No hypotheses were generated for neuroticism. However, past research has shown neuroticism to be difficult to perceive in general (see [Ambady et al., 2000](#); [Kenny, 1994](#), for reviews).

An important issue in designing this study was whether or not to match the age of the targets to the age of the judges. At least one previous study found that people read others better if they are of matching ages ([Boyatzis, Chazan, & Ting, 1993](#)). Adults may

read adults more effectively than children observers because they interact with each other most often. Adults have more experience and practice reading each other than a child does in reading adults. Also, a child probably has a different idea of how extraversion manifests itself in a peer than in an adult. Since a child has more experiences with seeing other children being told what to do (e.g., in the experimental situations) the child has a better understanding of how another child may display extraversion. We thought it would be misleading to ask a child to judge the traits of adults due to their lack of experience and perspective on adult life. Therefore, we thought it best to match the age of target and judge to best assess the children's judge-criterion agreement.

To investigate possible alternative explanations for differences found due to age matching, we had a different sample of adults watch either the 13-year-old target video or the 8-year-old target video. The level of criterion agreement achieved by adult perceivers watching the target children allowed us to determine whether differences were due to perceiver effects or target effects.

2. Method

2.1. Phase I: Development of the Personality Perception Task (PPT)

2.1.1. Targets

A total of 24 targets (4 males and 4 females participants from each age group: 8-year-olds ($M = 8.14$, $SD = .38$); 13-year-olds ($M = 12.75$, $SD = .46$); and adults ($M = 19.50$, $SD = 1.69$)), were observed and recorded. The adults were undergraduates who received partial course credit for their participation. The children participating in this study were recruited from the birth archives from an area hospital. Parents who gave birth to children 8 and 13 years prior were contacted and asked if they and their children would like to participate in a study on personality perception.

2.1.2. Personality assessment of targets

Ratings from close friends of the adults and ratings from parents of the children, by means of the California Q-sort (Bem & Funder, 1978), were used as the criterion for measuring accuracy.² Close friends and parents have been used in other studies as raters for the targets, and they have been reliable sources of personality information (Kenny, 1994, for a review).

The California Q-Sort (Block, 1961) was originally intended for expert rater use, but has since been modified by Bem and Funder (1978) for lay raters and for describing children (Van Lieshout & Haselager, 1994). Designed to measure per-

² An adult self-report personality survey based on the California Q-sort was created. Fifteen items were administered in a paper and pencil format. Three questions pertained to each of the five major personality traits. These items were the same criteria items selected from the Q-sort measure described above. The adults rated each question on a five point likert-like scale (unlike me to very much like me). Given the obvious issues of self-presentation, it was decided to employ the ratings from the knowledgeable others as the trait criterion over these self-reports.

sonality, the Q-sort contains 100 statements that were sorted in this experiment by a close friend or parent of the participant. The rater placed each of the 100 cards (according to the normal-forced-distribution for the adults while 11 statements could be placed in each category for the children except category 5, in which 12 cards could be placed) into one of nine categories ranging from least characteristic to most characteristic. Between seven and 10 questions, which correlated the highest with each of the Big Five traits based on the findings by Costa and McCrae (1985), were combined to form the criterion for that trait (see Appendix A).³ The questions were added together using the card's placement by the friend or parent in the 1–9 categories. The internal consistency scores for the items for the adults in this study were: neuroticism, $\alpha = .61$; extraversion, $\alpha = .73$; openness, $\alpha = .62$; agreeableness, $\alpha = .77$; conscientiousness, $\alpha = .73$. The internal consistency scores for the items for the children in this study were: neuroticism, $\alpha = .89$; extraversion, $\alpha = .87$; openness, $\alpha = .62$; agreeableness, $\alpha = .79$; conscientiousness, $\alpha = .73$.

2.1.3. Extraversion situation

Item 57 of the NEO-PI (Costa & McCrae, 1985), "I find it easy to smile and be outgoing with strangers," illustrates how important comfort in social situations is to the extraversion construct. To maximize extraversion, a social situation was created where the target was videotaped entering the waiting room and sitting down to a conversation between two adult confederates. The video camera was concealed behind a one-way window. The target could sit in one of five chairs to the right the confederates. Where the target sat (close or far away from the confederates) and how much the target talked with the confederates, whose behaviors were scripted, was believed to reveal how extraverted he/she was.

2.1.4. Openness situation

Item 62 of the NEO-PI (Costa & McCrae, 1985), "I am intrigued by the patterns I find in art and nature," illustrates how intellectual curiosity and an active imagination are important to the openness to experience construct. To maximize the expression of openness, targets were given an opportunity to display their curiosity and imagination. While in the waiting room, the confederates invited the target to comment on a large inkblot that was framed and hanging on a wall. What the target stated in response to this invitation was believed to reveal his/her imagination and capability to entertain novel thoughts (i.e., openness).

³ A child self-report personality survey based on the California child Q-sort items was created. The child rated each question on a five point likert-like scale (unlike me to very much like me). We felt the self-reports of children would suffer more from the validity threats of self-presentation and verbal comprehension than the self-reports of adults. Therefore, we employed the ratings from parents as the trait criterion for accuracy judgments.

2.1.5. Neuroticism situation

Item 134 of the NEO-PI (Costa & McCrae, 1985), “When I’m under a great deal of stress, sometimes I feel like I’m going to pieces,” illustrates the anxiety and vulnerability of the neurotic. To maximize neuroticism, targets were then escorted into an adjoining “game room” containing a table, two chairs and the block game, Jenga. A video camera in view in the game room recorded their entrance and subsequent task behavior. The game Jenga involves 24 pieces of wooden blocks that formed a tower. The confederate and the target were then seated (the confederate always sat to the right of the target) and told to draw a slip of paper out of the hat.

The experimenter then explained the task:

The object of this game is to take one block from the middle of the tower and place it on the top of the tower, repeating this until the tower falls or the time given for the task runs out. The two of you are partners. The person whose slip of paper says “block remover” will take out the blocks and put them on top, and the person whose slip says “block suggester” will suggest which blocks the block remover might pull out. The two of you may disagree with each other. You may discuss the best move. You have 3 min to take out as many as you can. The average person can pull out 18 in that amount of time. I will periodically call out the time left and blocks successfully placed on top.

Both slips of paper had “block remover” printed on them. The confederate’s behavior was standardized to be mildly uncooperative with the target and to choose undesirable blocks to pull.

The game was timed and a somewhat unreasonable goal was given to induce time pressure. Also, the assistant was instructed to be a bit difficult to deal with. This was done to induce more stress.

2.1.6. Agreeableness situation

Item 140 of the NEO-PI (Costa & McCrae, 1985), “I would rather cooperate with others than compete with them,” illustrates the altruistic and sympathetic nature of the agreeable person. To maximize the presentation of agreeableness, the Jenga task was created such that the target had opportunities to agree with or disagree with a confederate on every suggestion.

2.1.7. Conscientiousness situation

The final activity was designed to reveal conscientiousness. Item 15 of the NEO-PI (Costa & McCrae, 1985), “I try to perform all the tasks assigned to me conscientiously,” illustrates the importance of self-discipline to the conscientiousness construct. To maximize conscientiousness, the target was asked to help rebuild the block tower, which inevitably had crashed and scattered across the table and sometimes about the floor, during the previous activity. The confederate who was also asked to help always refused. Whether or not the target agreed to set up the tower was expected to reveal their agreeableness, but how carefully and how long they spent rebuilding it was expected to be driven by their conscientiousness.

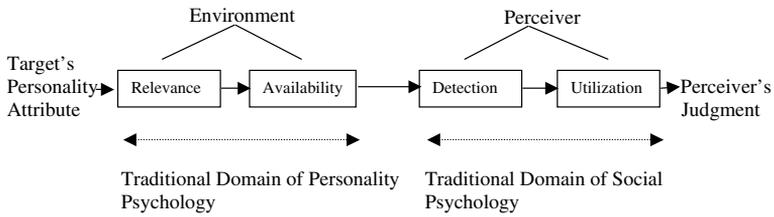


Fig. 1. A model of the process of accurate personality judgment (adapted from Funder, 1995).

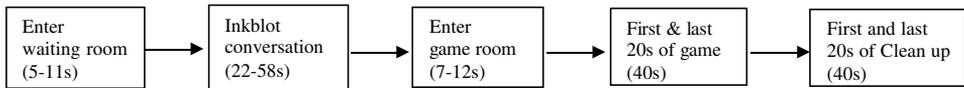


Fig. 2. Sequence of scenes that comprised a videoclip.

Funder's (1995) RAM has led to the careful consideration of the availability of behavior for the judge to rate. The tape was specifically constructed to attend the availability issue (see Fig. 1).

2.1.8. Procedure

Upon arrival, the parent/guardian of those under 18 completed the children's Q-sort relating to their child's personality. A close friend of the adults completed the adult Q-sort relating to the adult target's personality. Targets were then videotaped in a series of structured activities, each developed to maximize the expression of the behavior relevant to a given trait.

2.1.9. Stimulus tape construction

One tape was constructed for each age group of targets (8-year-olds, 13-year-olds, and adults). Each stimulus tape contained edited clips (approximately 2 min long) of 8 targets completing the previously mentioned tasks. The tape structure was similar in all aspects except for the age of the targets. Fig. 2 diagrams the sequence of scenes that comprised a clip. The target was shown entering the waiting room and sitting down. Then the inkblot conversation was included. These two parts of the tape varied in length depending on how long the target took to sit down (5–11 s), and how long the conversation lasted (22–58 s). The next sequence included the target entering the block game room and sitting down (7–12 s). Finally, the first and last 20 s. of the block game and the clean up were shown. The tapes began with a ninth target, a male, who served as practice. The judges first viewed each of the four male targets and then each of the four female targets (order was preserved for all 8 targets).

2.2. Phase II: Accuracy assessment

2.2.1. Participants

The stimulus tapes created in Phase I were shown to participants of matching ages. The 210 participants included 70 (35 females and 35 males) from each of three

age groups (8-year-olds, $M = 8.03$, $SD = .24$; 13-year-olds, $M = 12.99$, $SD = .36$; and adults, $M = 18.63$, $SD = .77$). In addition, a total of 44 adult undergraduates (22 males and 22 females) were recruited to rate the children's tapes. Half of these participants viewed the 8-year-old target tape while the other half watched the 13-year-old target stimulus tape.

Participants rated targets using a rating form based on the Big Five factors (Costa & McCrae, 1985). The questionnaires included five questions, one question relating to each trait: (a) How outgoing and animated (lively) is the person? (extraversion); (b) How untraditional and creative is the person? (openness); (c) How pleasant and positive is the person? (agreeableness); (d) How relaxed and calm is the person? (neuroticism); (e) How conventional and hardworking is the person? (conscientiousness). The items were derived from Kenny's (1994) description of the aspects of the Big Five traits in personality perception literature. The judge rated each question on a scale from 1 to 5 (unlike to very much like) according to what the judge felt the target person was actually like.

A children's version of this rating scale offered simpler terminology. Three second grade teachers from two local public elementary schools were given the five adult Big Five items and were asked to alter the words so that children as young as 8-years-old could comprehend them, while preserving the meaning of the items. The three teachers modified the items to read: (a) How outgoing and active (lively) is the person? (extraversion); (b) How different and creative is the person? (openness); (c) How likable and positive is the person? (agreeableness); (d) How relaxed and calm is the person? (neuroticism); and (e) How careful and hardworking is the person? (conscientiousness).

Prior to beginning the task, the participants were given: (a) a verbal description of the activities they would observe, (b) a practice clip, and (c) instructions pertaining to the use of the questionnaire. The judges watched one target clip (approximately 2 min long), and then rated the target. Once the rating for target one was complete, the clip of target two was shown and the processes was repeated until the final target, target eight, was rated.

3. Results

3.1. Rater agreement

The level of agreement across raters for the adults, 13-year-olds, and 8-year-olds were $\alpha = .95$, $\alpha = .70$, and $\alpha = .86$, respectively.

3.2. Judge-criterion agreement analysis

Judge-criterion agreement was computed within each trait by correlating a judge's assessment of a target's given trait with the criterion (the parent/friend judgment of each trait) across 8 targets. A judge generated one agreement correlation for each of the five traits examined. In this manner, an agreement score of $r = .00$ ($df = 6$) indicated that the perceiver's judgment was uncorrelated with the criterion trait scores collected and thus, represented the null hypothesis of no agreement. These correlations were converted to Fisher z s for analyses (Rosenthal & Rosnow, 1991) including

Table 1
Mean judge-criterion agreement across eight targets as a function of age

Trait	Age of participant and target		
	18 years, $n = 70$, r	13 years, $n = 70$, r	8 years, $n = 70$, r
Neuroticism	.12*	.01	.04
Extraversion	.57*	.39*	.42*
Openness	.18*	.13*	.02
Agreeableness	-.14*	.05	.05
Conscientiousness	.25*	.10*	.07
Median agreement	.18	.10	.05

Note. The correlations appearing in this table represent the average judge-criterion agreement correlation achieved by a group of judges. Averages were computed on Fisher z scores and then converted back into their Pearson r equivalents for presentation. Comparisons against zero (t tests) were performed on these means.

* $p < .05$.

t tests against zero and ANOVAs. For each trait, a 2 (sex of judge) \times 3 (age of judge) ANOVA was computed. Linear contrasts were conducted to determine whether judge-criterion agreement increased with age of judge. Sex of participant did not significantly effect agreement or its interaction with age in any of the analyses and will not be considered further.

Table 1 reports the mean judge-criterion agreement coefficients achieved by each group of 70 judges for each of the five traits. The values in the table have been transformed back in to their Pearson r equivalents for clarity of presentation. Adult judgments significantly agreed with the criterion on all traits except agreeableness. Agreeableness showed a significant *negative* relationship, which means the targets rated least agreeable by judges were reported as being most agreeable by their friends. The 13-year-olds agreed with all but the neuroticism and agreeableness constructs, while the 8-year-olds only agreed with the criterion on extraversion.

Linear age contrasts generated mixed results. The judge-criterion agreement coefficients of openness and conscientiousness increased linearly with age (openness: $F(1, 204) = 5.43$, $p < .05$, $\eta^2 = .03$, conscientiousness: $F(1, 204) = 8.16$, $p < .05$, $\eta^2 = .04$). Agreeableness agreement coefficients decreased with age, $F(1, 204) = 9.14$, $p < .05$, $\eta^2 = .06$. Linear trend analyses for extraversion and neuroticism were not significant. Judge-criterion agreement for extraversion was similarly high for the 8- and 13-year-olds and then jumped higher for the 18-year-olds, while agreement for neuroticism remained rather low for all age groups.

3.3. Adult-child agreement analyses

To test the alternative hypotheses that the effects observed were target effects rather than perceiver effects, we examined the level of agreement achieved by adult perceivers judging the three different age groups. The ANOVA yielded a significant Target effect for agreeableness (see Table 2), and the linear trend analyses for agreeableness was significant $F(1, 209) = 4.05$, $p < .05$, $\eta^2 = .07$.

Table 2
Mean judge-criterion agreement of adult observers as a function of target age

Trait	Target age			F(2, 110)	η^2
	Adult, n = 70, r	13 years, n = 22, r	8 years, n = 22, r		
Neuroticism	.12*	.04	.19*	1.13	.02
Extraversion	.57*	.45*	.37*	4.60*	.08
Openness	.18*	.27*	.30*	1.24	.02
Agreeableness	-.14*	.05	.07	4.05*	.07
Conscientiousness	.25*	.14*	.27*	1.37	.02

Note. The correlations appearing in this table represent the average judge-criterion agreement correlation achieved by the adult judges. Averages were computed on Fisher z scores and then converted back into their Pearson r equivalents for presentation. Comparisons against zero (t tests) were performed on these means.

* p < .05.

The post hoc comparisons revealed that adult judgments agreed more with the criterion in assessing the openness of 8-year-old targets than those from 8-year-old observers, $t(90) = 2.4, p < .05$. Adults also showed more agreement with the criterion in assessing conscientiousness in 8-year-olds $t(90) = 2.2, p < .05$.

3.4. Internal consistency of personality perception accuracy

An unresolved issue in the good judge literature is whether there exists a general competency to perceive others, or whether the skill is specific to the task or trait being observed. In other words, there is little evidence to this date showing that the reading of one trait or one internal state in one context is reliably related to that achieved for other traits, states, and contexts (Davis & Kraus, 1997).

The intercorrelations of the 5 trait agreement scores within each age group appear in Table 3. It can be seen that the intercorrelations between traits were quite low. In fact, the median intercorrelation of trait agreement coefficients were .01, .07, and .03

Table 3
Intercorrelations of trait agreement as a function of age

Trait	Age of judge											
	18 years, n = 70, r				13 years, n = 70, r				8 years, n = 70, r			
	N	E	O	A	N	E	O	A	N	E	O	A
Neuroticism	—				—				—			
Extraversion	.06				.09				.10			
Openness	-.17	.13			.03	.12			.00	.20		
Agreeableness	.05	-.14	.08		.05	.04	.20		-.33*	-.03	.17	
Conscientiousness	-.03	-.05	.00	.02	.04	.28*	-.09	-.05	.34*	.00	.21*	-.38*

Note. Correlations reflect the degree to which the judge-criterion agreement on one trait correlates with the agreement on another trait.

* p < .05.

for the 18-, 13-, and 8-year-olds, respectively. The ability to judge one particular trait was largely unrelated to the ability to judge traits in general.

4. Discussion

Personality perception was examined where friends/parents provided the personality criterion data on targets, and unacquainted participant observers assessed their personality traits from a thin slice of behavior. In phase I, eight targets from each age group performed a series of tasks. These tasks were theoretically derived to maximize the relevance and visibility of the “Big Five” characteristics described by [Costa and McCrae \(1985\)](#). According to [Funder \(1995\)](#), for judges to accurately perceive a characteristic, the target has to produce behavior that is relevant or potentially informative about his/her characteristics and these diagnostic behaviors must be available to the judge. To help ensure that the target would produce relevant and available information, we derived tasks that would enhance the relevance and visibility of the personality characteristics of interest.

In most zero-acquaintance personality perception studies, adult participants have an average Big Five agreement coefficient with the criterion of .13 ([Kenny, 1994](#)). In the above study, adults showed an average coefficient of .24. The carefully constructed thin slice employed here improved the overall agreement achieved for the traits. The increased agreement observed here suggests that in many zero-acquaintance situations, the relevance and availability of many trait related cues may be minimal. Clearly, the most relevant and diagnostic an activity is for a trait, the better the observer will be at assessing it ([Funder, 1995, 1999](#)).

It was predicted that trait agreement would differ not only over age, but also across each trait. As explained above, agreement for extraversion is likely to be the first to emerge. The trait extraversion is mostly based on physical action like talkativeness and expressiveness ([Costa & McCrae, 1985](#)). It should be easy for children to detect how talkative and expressive a person is. Consistent with this reasoning, we found that agreement for this trait at age 8 was high and remained high over these age groups.

We predicted openness and conscientiousness perception agreement to appear later. Aspects of traits involving creativity and helpfulness are more difficult to identify than aspects of extraversion such as expressiveness. The child must know that the method the target uses to rebuild the tower (i.e., slow and with precision vs. quickly and sloppy) may help in making a decision about the target’s level of conscientiousness. In addition, the child has to know when a behavior is responsible (i.e., be sensitive to context). A responsible behavior in one situation may not be responsible in another situation. Once the child knows what behavior to attend to, and whether or not the behavior is responsible, the child then has to know how to use the information in determining the trait. For instance, if a person is responsible, then they are probably conscientious, and if the person is not responsible, he/she is probably not conscientious. Thus, the child must take into account situational cues to help them determine which expressive cue reveals the internal state.

Depaulo (1991) suggested that the ability to consider both nonverbal and situational cues as needed here does not occur until around age 10. In fact, Ruble, Newman, Rholes, and Altshuler (1988) found that younger children's use of past behavior to form expectations was highly variable compared to older children. In addition, Rotenberg (1982) suggested that it was only after age 8 that children were able to reason about people in a more abstract manner such as complex traits. An understanding of openness and conscientiousness may be developing more slowly because they require the child to infer more about the behavior and the situation. Consistent with this reasoning we found that agreement for openness and conscientiousness increased with age.

The disagreement that occurred for agreeableness was a surprise. In this study, the overall agreement for agreeableness was negative. A significant negative score means that when a target's criterion score for agreeableness was high, the judges rated it as low, and if a target's criterion score was low, the judges rated it as high. Given that most studies report positive correlations for adults on this trait, it is unlikely that the children learn to read those who are agreeable as not agreeable and those who are not agreeable as agreeable.

The task created for agreeableness, the block task with the incompetent confederate, may have inadvertently misguided observer judgments of this trait. The activity and context were designed to reveal agreeableness by providing many opportunities for the target to agree or disagree with the confederate who was making poor suggestions. In fact, it was possible to count the number of times the target agreed or disagreed with the confederate explicitly while performing this task. When observer judgments of agreeableness were correlated with this overt measure of explicit agreement a different pattern of results emerged. The correlation between judged agreeableness and agreement behavior was $r = .11$ within adults, $r = .04$, within the 13-year-olds, and $r = .01$ within the 8-year-olds. In other words, the judgment of agreeableness correlated positively with agreement behavior for adults and showed an increasing, if not significant, developmental trend across ages.

The reason adult judgments were negatively correlated with the actual trait criterion may have something to do with the fact that the agreement behavior described happened to be negatively related to the trait criterion ($r = -.35$). We were correct in intuiting that the observer judgments of agreeableness would be influenced by target agreement with the confederate, but we did not anticipate that this behavior would correlate negatively with our trait criterion.

Paradoxically, the amount of agreement behavior (to the confederate) may have reflected lower trait agreeableness because of the complexity inherent in the greater social context that constituted the psychological experiment. There were two people with whom the target was interacting; the confederate and the experimenter. When the confederate begins to suggest poor actions, he undermines the goals and objectives of the experimenter. Thus, to agree with the confederate is to risk disapproval from the experimenter. An agreeable person who wants to conform to, and please, the experimenter would thus be forced to reject the suggestion of the confederate. This is supported by Tobin et al. (2000) who state that highly agreeable persons will, in an attempt to maintain smooth social relations, often be more concerned than

their peers with the emotional experience of others during an interaction. In the present case, the agreeable person may have been more concerned with pleasing the experimenter thus disagreeing with the confederate.

We suspect, therefore, that we inadvertently structured the activity with conflicting motives that fooled our observers who were given only a thin slice exposure to it. In the context provided to adult observers, the number of times a target “agreed” with the confederate actually correlated negatively with the criterion for trait agreeableness. This negative relationship between overt behavior and trait criterion, however, did not exist within the two children target samples. And consistent with this analysis, observer accuracy within children was not significantly negative. The above analysis and discussion illustrates the profound impact that the situation can have on both the expression and judgment of a given trait. Researchers need to be ever mindful of the social and interpersonal nuances within their staged settings.

An important issue in designing this study was whether or not to match the age of the targets to the age of the judges. There are three alternative explanations for the main effect of age differences predicted that could arise because target age varied along with observer age. The agreement differences could be due to differences in true target trait variance within the three age groups. For example, the 8-year-olds may agree less if the 8-year-old targets all happened to be low in conscientiousness. It would be more difficult for the children to rate conscientiousness if targets do not differ along this dimension. If the adults have a wider range of targets along this trait domain, it would make it easier to rate conscientiousness. This alternate explanation was not supported. The targets varied extensively in conscientiousness when it was measured in this study.

Another possible explanation for the results would be that the expression of traits may differ across age. Funder (1995) stated that the availability of behavior is important for trait perception. Conscientiousness for adults may include actively helping, but conscientiousness for children may include listening attentively. It would be easier for adults to assess helping behavior than it would be for children to assess listening skills. Helping behavior is more available to the judge. This could cause adult judges to more have better judge-criterion personality agreement.

A third possibility is that the children may react differently than the adults given the identical situation (Kenny, 1994). For example, all targets are asked to rebuild the tower. The adult targets may feel freer to leave the tower jumbled, whereas the child may feel intimidated by the adult experimenter. If all of the children rebuild the tower, it would be unclear if they did it to be helpful or because they felt intimidated or forced to. The first and final steps of Funder’s (1995) RAM model are important here. Funder (1995) stated that a behavior must be relevant (step 1) and the judge must be able to correctly utilize the target’s behavioral information (step 4) for an accurate judgment to occur. The child judges may incorrectly identify conscientiousness because they use the rebuilding as a sign of conscientiousness, when in fact it is intimidation that caused the children to rebuild the tower. The rebuilding behavior is not relevant to conscientiousness in this case, but the children could use this information in determining this trait. The children may be utilizing irrelevant behavior in determining a trait. The adult judges also may use the rebuilding as a sign of conscientiousness, and since there was no intimidation, they may do so correctly. The

rebuilding behavior, in this case, would be relevant to the trait, and the judges would then be correctly connecting the behavior to the trait.

To address these issues, we had a different sample of adults watch either the 13-year-old target video or the 8-year old target video. It was hypothesized that trait judgment agreement would not differ between tapes making it unlikely that the differences in agreement found between the age groups in study one were due to the above alternative explanations. The results from the adults viewing the childrens' videos suggest that the relatively low agreement of the 8-year-old judges was not due to the targets' readability. The median effect size of agreement achieved by adults reading the 8-year-olds was significantly above zero, $r = .27$, compared to the median agreement of $r = .05$ recorded for the 8-year-old judges. Likewise, the median level of agreement achieved for the adults reading the 13-year-old targets was $r = .14$, which was well above the median of $r = .10$ achieved by the 13-year-old judges. We found little evidence to suggest that the 8- and 13-year-olds were unreadable.

In addition, when comparing the results from the adults watching the childrens' videos to the children watching their own age group, adults agreed more with the criterion in reading openness and conscientiousness than did children. It may be that adults are better at rating and understanding what is meant by openness. As discussed before, the concepts of openness and conscientiousness are rather complicated. It requires more than just connecting behavior with internal states. The concepts themselves involve abstract thought. Again, Rotenberg (1982) suggests that abstract thought like those involved in trait inference does not occur until after age 8. The results for this study support this notion suggesting that adults are more skilled at perceiving more complex traits.

It is known that adults can perceive thin slice stimuli accurately (Ambady et al., 2000). This study extends this knowledge by examining how this ability improves over time trait-by-trait. Overall, the prediction that trait judgment agreement would differ depending on the trait perceived was confirmed. Judge-criterion agreement for extraversion was the first to emerge. Agreement for this trait at age 8 was high and remained high over the ages. Openness and conscientiousness perception agreement appeared later and increased over the ages. The judge-criterion disagreement for agreeableness was a surprise, but again, this pattern seems to be explained by the tasks designed for the study. No prediction was made for neuroticism, but the lack of agreement achieved replicates many previous studies (see Ambady et al., 2000; Kenny, 1994, for reviews). Future studies should extend this research by investigating whether agreement in perceiving extraversion is possible at younger ages.

There are a few more issues to consider when applying or generalizing the results of this study. First, observers only rated the targets on a single item for each of the Big Five traits, and the observers only viewed eight targets. Generally, it is preferable to use multiple items to improve both the reliability and validity of the ratings, but to lessen confusion and maximize the attention of the observers, only one question was asked for each trait. In addition, the correlations being analyzed were derived from only eight targets. Again, the attention span of an 8-year old is very short, and to have the children view the targets in more than one scenario, fewer targets had to be shown. In the future, a more comprehensive study should be conducted to examine these two issues.

Another question to be raised is whether the people who described the targets initially were of equivalent ability. In Study 1, kids' judgments were evaluated against parents' ratings whereas adult judgments were evaluated against peer ratings. For this study, the authors chose to have the person who knew the target best describe them. Obviously, having the peers of the children describe them would defeat the purpose of the study, but whether or not to use peers or parents of the adults became a difficult decision. Many of the parents of the adults in the study were miles away, and if the parents were able to evaluate their children, the adults, mostly college students, have been living out of the house for a few years with little contact with their parents. Based on previous research reviewed by [Kenny \(1994\)](#), and the fact that the parents may have been out of touch with their children's lives, it seemed most logical to use peers of the adults and the parents of the children. Although this seemed to be the best choice, some studies using peers and parents show lessened agreement about common targets ([Kenny, 1994](#), for review). In other words, the criterion data come from different sources and hence could add another unknown component to the overall results. Another study should investigate this possibility further.

Finally, as discussed earlier in the paper, three second grade teachers from two local public elementary schools were given the five adult Big Five items and were asked to alter the words so that children as young as 8-years-old could comprehend them, while preserving the meaning of the items. These items were then given to the observing children to use in their evaluation of the target children. Although the words were best matched to the words the adults used in their description, these words may have slightly different meaning to the children. The teachers did make their best effort to create words of equal meaning, but there is no way of knowing how the children really interpreted them. An additional study would be required to verify the children's actual interpretation of the words chosen, but given that each the second grade teachers have been working in the field for 25 years or more, it is likely that they could relate to the children enough to match the meaning of the words for them.

An additional result from this study addressed the unresolved issue concerning the existence of a general competency to perceive others. The idea of a "good judge" was of great interest in the early to mid 1900's (for a review see [Cline, 1964](#); [Cook, 1979](#); [Taft, 1955](#)) and has again become an active research topic ([Funder & Colven, 1997](#)). The results here do not support the "good judge" concept. Intercorrelations between the five trait perceptions were extremely low.

The tapes created for this study involved structured activities developed to maximize the expression of the behavior relevant to the given trait. Maximizing the targets' expression of behavior increased the likelihood that the behaviors relevant to the judgment are present for detection ([Funder, 1995](#)). Tapes developed to maximize relevant trait variance help make the measurement of trait perception more valid. Researchers need to ensure that there is a trait "to see" before trying to determine whether an observer can see it.

It is hoped that this study will help illustrate to person perception researchers how important it is to present targets in meaningful contexts for observers to judge. If properly chosen, the context within which targets are observed should

reveal their traits to observers even within the limits of a thin slice (Ambady et al., 2000). Furthermore, this general paradigm appears generalizable to investigate developmental trends.

Appendix A. Adult criterion Q-sort cards

A.1. Neuroticism

1. Is thin-skinned; sensitive to anything that can be construed as criticism or an interpersonal slight. (e.g., rudeness or insult). (13)
2. Is basically anxious. (Nervous, worries a lot underneath). (68)
3. Is calm, relaxed in manner. (33)
4. Is irritable. (34)
5. Is guilt prone. (47)
6. Seeks reassurance. (19)
7. Is satisfied with self. (74)
8. Is self-defensive. (12)
9. Has a clear-cut personality. (75)
10. Has fluctuating moods. (82)

A.2. Extraversion

1. Is a talkative individual. (4)
2. Keeps people at a distance; avoids close interpersonal relationships. (48)
3. Is emotionally bland; has flattened affect. (Tends not to experience strong intense emotions.) (97)
4. Has over control of impulses. (25)
5. Is gregarious. (54)
6. Is socially poised. (92)
7. Behaves assertively. (52)
8. Is skilled in play and humor. (15)
9. Has a rapid tempo. (20)
10. Is an interesting person. (57)

A.3. Openness

1. Genuinely values intellectual and cognitive matters. (Ability to achievement is not implied here). (51)
2. Tends to be rebellious and non-conforming. (62)
3. Favors conservative values in a variety of areas. (Favors preserving traditional practices, values and conditions). (7)
4. Judges in conventional terms. (63)
5. Has unusual thought processes. (39)
6. Is introspective. (16)

7. Is uncomfortable with complexities. (9)
8. Is intelligent. (8)
9. Is aesthetically reactive. (66)
10. Is moralistic. (41)

A.4. Agreeableness

1. Behaves in a sympathetic or considerate manner. (17)
2. Has warmth; has a capacity for close relationships; compassionate. (35)
3. Behaves in an assertive fashion. (Speaks up to get what he/she wants; not afraid to express opinions. This refers to how the person acts, not how he/she might feel while doing so.) (52)
4. Is critical. (1)
5. Shows condescending. (27)
6. Tries to push limits. (65)
7. Expresses hostility directly. (94)
8. Arouses liking. (28)
9. Is power oriented. (91)
10. Avoids close relationships. (48)

A.5. Conscientiousness

1. Behaves in an ethically consistent manner; is consistent with own personal standards. (70)
2. Is a genuinely dependable and responsible person. (2)
3. Various needs tend toward relatively direct and uncontrolled expression unable to delay gratification. (Has little self-control; expresses emotions impulsively; unable to postpone pleasure.). (53)
4. Is interested in opposite sex. (80)
5. Is unable to delay gratification. (53)
6. Is self-indulgent. (67)
7. Is intelligent. (8)
8. Is productive. (26)
9. Enjoys sensuous experiences. (58)
10. Has high aspiration level. (71)

Note. Q-sort item number appears in parentheses.

Appendix B. Child criterion Q-sort cards

B.1. Neuroticism

1. Self-reliant, confident. (88)
2. Fearful, anxious. (23)

3. Anxious in unpredictable situations. (60)
4. Tends to go to pieces under stress. (46)
5. Tends to brood and ruminate or worry. (24)
6. Self-assertive. (82)
7. Calm and relaxed, easy-going. (64)
8. Appears to feel unworthy. (77)
9. Inhibited and constricted. (35)
10. Has a readiness to feel guilty. (72)

B.2. Extraversion

1. Keeps thoughts and feelings to self. (8)
2. Shy and reserved. (98)
3. Likes to be by him/herself. (86)
4. Emotionally expressive. (58)
5. Expresses negative feelings openly. (18)
6. Tends to yield and give in. (44)
7. Verbally fluent. (69)
8. Withdraws under stress. (45)
9. Is a talkative child. (84)

B.3. Openness

1. Active fantasy life. (97)
2. Creative. (96)
3. Interesting and arresting child. (42)
4. Physically attractive, good-looking. (92)
5. Suspicious of others. (79)
6. Responds to humor. (73)
7. Resourceful in initiating activities. (36)
8. Admired and sought by other children. (5)
9. Cheerful. (75)

B.4. Agreeableness

1. Is considerate of other children. (2)
2. Is warm and responsive. (3)
3. Is helpful and cooperative. (6)
4. Can be trusted, is dependable. (76)
5. Shows concern for moral issues. (15)
6. Recognize feelings of others. (31)
7. Attempts to transfer blame to others. (11)
8. Tends to give, lend and share. (32)
9. Open and straightforward. (19)
10. Protective of others. (29)

B.5. Conscientiousness

1. Attentive. Able to concentrate. (66)
2. Is curious and exploring. (40)
3. Planful, thinks ahead. (67)
4. Performance standards for self-high. (47)
5. Is reflective. (99)
6. Competent, skillful. (89)
7. High intellectual capacity. (68)

Note. Q-sort item number appears in parentheses.

References

- Albright, L., Kenny, D. A., & Malloy, T. E. (1988). Consensus in personality judgments at zero acquaintance. *Journal of Personality and Social Psychology, 55*, 387–395.
- Allport, G. W. (1937). *Personality: A psychological interpretation*. New York: Holt.
- Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin, 111*, 256–274.
- Ambady, N., Bernieri, F. J., & Richeson, J. (2000). Towards a histology of social behavior: Judgment accuracy from thin slices of behavioral stream. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 32, pp. 201–271). San Diego: Academic Press.
- Baldwin, M. W. (1992). Relational schemas and the processing of social information. *Psychological Bulletin, 112*, 461–484.
- Bem, D. J., & Funder, D. C. (1978). Predicting more of the people more of the time: Assessing the personality of situations. *Psychological Review, 85*, 485–501.
- Bernieri, F. J. (2001). Toward a taxonomy of interpersonal sensitivity. In J. Hall & F. Bernieri (Eds.), *Interpersonal sensitivity: Theory and measurement* (pp. 3–20). NJ: LEA.
- Block, J. (1961). *The Q-sort method in personality assessment and psychiatric research*. Springfield, IL: Charles C. Thomas.
- Borkenau, R., & Liebler, A. (1992). Trait inferences: Sources of validity at zero acquaintance. *Journal of Personality and Social Psychology, 62*, 645–657.
- Boyatzis, C. J., Chazan, E., & Ting, C. Z. (1993). Preschool decoding of facial emotions. *Journal of Genetic Psychology, 153*, 375–382.
- Buss, D. M., & Craik, K. H. (1983). The act of frequency approach to personality. *Psychological Review, 90*, 105–126.
- Cline, V. B. (1964). Interpersonal perception. In B. A. Maher (Ed.), *Progress in experimental personality research* (pp. 221–284). New York: Academic.
- Cook, M. (1979). *Perceiving others: The psychology of interpersonal perception*. London: Methuen.
- Costa, P. T., & McCrae, R. R. (1985). *The NEO personality inventory manual*. Odessa, Florida: Psychological Assessment Resources.
- Costanzo, M., & Archer, D. (1989). Interpreting the expressive behavior of others: The interpersonal perception task. *Journal of Nonverbal Behavior, 13*, 225–245.
- Cumberland-Li, A., Eisenberg, N., & Reiser, M. (2004). Relations of Young Children's Agreeableness and Resiliency to Effortful Control and Impulsivity. *Social Development, 13*, 193–212.
- Dabbs, J. M., Bernieri, F. J., Strong, R. K., Campo, R., & Milun, R. (2001). Going on stage: Testosterone in greeting and meeting. *Journal of Research in Personality, 35*, 27–40.
- Davis, M. H., & Kraus, L. A. (1997). Personality and empathic accuracy. In W. J. Ickes (Ed.), *Empathic accuracy* (pp. 144–168). New York: Guilford Press.

- Depaulo, B. M. (1991). Nonverbal behavior and self-presentation: A developmental perspective. In R. Feldman & B. Rime (Eds.), *Fundamentals of nonverbal behavior* (pp. 351–397). New York: Cambridge University Press.
- Depaulo, B. M., & Friedman, H. S. (1997). Nonverbal communication. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 89–150). New York: McGraw-Hill.
- Depaulo, B. M., & Rosenthal, R. (1986). Measuring the development of sensitivity to nonverbal communication. In C. E. Izard (Ed.), *Measuring emotions in infants and children* (pp. 209–245). New York: Cambridge University Press.
- Depaulo, B. M., Lassiter, -G. D., & Stone, J. I. (1982). Attentional determinants of success at detecting deception and truth. *Personality and Social Psychology Bulletin*, 8, 273–279.
- Eisenberg, N., Murphy, B. C., & Shepard, S. (1997). The development of empathic accuracy. In W. J. Ickes (Ed.), *Empathic accuracy* (pp. 73–116). New York: The Guilford Press.
- Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review*, 102, 653–670.
- Funder, D. C. (1999). *Personality judgment: A realistic approach to person perception*. San Diego: Academic Press.
- Funder, D. C., & Colven, C. R. (1997). Congruence of self and others' judgments of personality. In R. Hogan, J. A. Johnson, & S. R. Briggs (Eds.), *Handbook of personality psychology* (pp. 617–647). San Diego, CA: Academic.
- Funder, D. C., & Colvin, R. C. (1991). Explorations in behavioral consistency: Properties of persons, situations, and behaviors. *Journal of Personality and Social Psychology*, 60, 773–794.
- Funder, D. C., & Sneed, C. D. (1993). Behavioral manifestations of personality: An ecological approach to judgmental accuracy. *Journal of Personality and Social Psychology*, 64, 479–490.
- Gnepp, J., & Hess, D. L. (1986). Children's understanding of verbal and facial display rules. *Developmental Psychology*, 22, 103–108.
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4, 26–42.
- Hall, J. A., & Bernieri, F. J. (2001). *Interpersonal sensitivity: Theory and measurement*. Mahwah, NJ: Erlbaum.
- Harris, P. L., Donnelly, K., Guz, G. R., & Pitt-Watson, R. (1986). Children's understanding of the distinction between real and apparent emotion. *Child Development*, 57, 895–909.
- Jones, D. C., Abbey, B., & Cumberland, A. (1998). The development of display rule knowledge: Linkages with family expressiveness and social competence. *Child Development*, 69, 1209–1222.
- Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis*. New York: Guilford.
- Kruglanski, A. W. (1989). The psychology of being "right": The problem of accuracy in social perception and cognition. *Psychological Bulletin*, 106, 395–409.
- Linkey, H. E., & Firestone, I. J. (1990). Dyad dominance composition effects, nonverbal behaviors, and influence. *Journal of Research in Personality*, 24, 206–215.
- Losoya, S. H., & Eisenberg, N. (2001). Affective empathy. In J. Hall & F. Bernieri (Eds.), *Interpersonal sensitivity: Theory and measurement* (pp. 21–46). NJ: LEA.
- Masten, A., Morison, P., & Pellegrini, D. (1985). A revised class play method of peer assessment. *Developmental Psychology*, 21, 523–533.
- McCrae, R. R., Costa, P. T., Jr., & Busch, C. M. (1986). Evaluating comprehensiveness in personality systems: The California Q-set and the five-factor model. *Journal of Personality*, 54, 430–446.
- McCrary, P., Bernieri, F. J., Grahe, J., & Gada-Jain, N. (2003). *Neuroticism observed: A thin-slice assessment of dyadic interactions over 10 weeks*. Unpublished Manuscript.
- Nowicki, S., & Duke, M. (2001). Nonverbal Receptivity: The diagnostic analysis of nonverbal accuracy scale. In J. Hall & F. Bernieri (Eds.), *Interpersonal sensitivity: Theory and measurement* (pp. 183–200). NJ: LEA.
- Richeson, J. A., & Ambady, N. (2001). Who's in charge? Effects of situational roles on automatic gender bias. *Sex Roles*, 44, 493–512.
- Rholes, W. S., & Ruble, D. N. (1986). Children's impressions of other persons: The effects of temporal separation of behavioral information. *Child Development*, 57, 872–878.
- Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis*. New York: McGraw-Hill.

- Rotenberg, K. J. (1982). Development of character consistency of self and other. *Child Development*, 53, 505–515.
- Ruble, D. N., Newman, L. S., Rholes, W. S., & Altshuler, J. (1988). Children's "naive psychology": The use of behavioral and situational information for the prediction of behavior. *Cognitive Development*, 3, 89–112.
- Saarni, C. (1990). *Emotional competence: How emotions and relationships become integrated*. Nebraska: University of Nebraska Press.
- Scherer, K. R. (1978). Personality inference from voice quality: The loud voices of extraversion. *European Journal of Social Psychology*, 8, 467–487.
- Taft, R. (1955). The ability to judge people. *Psychological Bulletin*, 52, 1–23.
- Tobin, R. M., Graziano, W. G., Vanman, E. J., & Tassinari, L. G. (2000). Personality, emotional experience, and efforts to control emotions. *Journal of Personality and Social Psychology*, 79, 656–669.
- Van Lieshout, C., & Haselager, G. (1994). The Big Five personality factors in Q-sort Descriptions of children and adolescents. In C. H. G. Kohnstamm & R. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 293–318). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Waxer, P. (1976). Nonverbal cues for depth of depression: Set versus no set. *Journal of Consulting and Clinical Psychology*, 44, 493.
- Wiggins, R. G. (1973). Differences in self-perceptions of ninth grade boys and girls. *Adolescence*, 8, 491–496.