Child Development, December 1998, Volume 69, Number 6, Pages 1556–1576

Assessing Young Children's Views of Their Academic, Social, and Emotional Lives: An Evaluation of the Self-Perception Scales of the Berkeley Puppet Interview

Jeffrey R. Measelle, Jennifer C. Ablow, Philip A. Cowan, and Carolyn P. Cowan

In a prospective, longitudinal study we examined the psychometric properties of the self-perception scales of the Berkeley Puppet Interview (BPI). A total sample of 97 young children were assessed with the BPI at 3 time points: preschool, kindergarten, and first grade. The BPI assesses young children's self-perceptions of their school adjustment in 6 domains: academic competence, achievement motivation, social competence, peer acceptance, depression-anxiety, and aggression-hostility. Results showed that 4½- to 7½-year-olds possess a multidimensional self-concept that can be reliably measured and that the BPI is sensitive to normative changes and individual differences in young boys' and girls' views of themselves. Support for the method's validity was derived from consistent and meaningful patterns of convergence between children's self-perceptions and ratings by adult informants—mothers, fathers, and teachers—as well as standardized test scores. In fact, in this study, the concordance between young children's self-reports and parallel ratings by teachers or mothers were consistently as strong as if not stronger than the concordance between mothers' and teachers' ratings.

INTRODUCTION

Current theories of children's academic competence and motivation (Dweck & Leggett, 1988; Eccles, Wigfield, Harold, & Blumfeld, 1993), social competence (Crick & Dodge, 1994; Ladd, Price, & Hart, 1988), and psychological adjustment (Cicchetti, 1993; Rutter, 1989) posit that children's self-perceptions play an active role in shaping their behavior. Most research on children's self-perceptions has been conducted with children ages 8 or older (Byrne, 1996; Harter & Pike, 1984; Wylie, 1989). Very little is known about the nature and behavioral significance of younger children's self-perceptions.

This gap in our knowledge stems in part from a lack of measures appropriate to assess young children's self-perceptions. In their recent critiques of self-concept instrumentation, both Byrne (1996) and Wylie (1989) concluded that only two measures—the Pictorial Scale of Perceived Competence and Social Acceptance for Young Children (PSPCSA; Harter & Pike, 1984) and the Joseph Pre-School and Primary Self-Concept Screening Test (JPPSST; Joseph, 1979)—demonstrated adequate psychometric properties with children ages 4–7. Of these, the PSPCSA is the most widely used self-concept measure with young children. Concern, however, has been raised about the PSPCSA's lack of criterion validity (Vaughn, 1995), and more recent evidence suggests that preschoolers have difficulties understanding the instrument's administration format (Fantuzzo, McDermott, Holliday Manz, Hampton, & Alvarez Burrdick, 1996). In short, both Byrne (1996) and Wylie (1989) conclude that the options available to researchers interested in young children's self-perceptions are limited.

This article reports on a new instrument, the Berkeley Puppet Interview (BPI; Ablow & Measelle, 1993), an age-appropriate method that uses puppets to interview young children about their self-perceptions of key aspects of their lives. In addition to describing how the BPI can help to overcome difficulties in obtaining self-reports from young children, we present construct validity data on six BPI scales designed to measure 4½- to 7½-year-olds' perceptions of their academic competence, achievement motivation, social competence, acceptance by peers, depression-anxiety, and aggression-hostility. We begin by considering unresolved issues in research on young children's self-perceptions. In particular, questions persist about (1) the structure of young children's self-conceptions, (2) age changes and gender differences in children's self-perceptions, and (3) the patterns of agreement between children's self-reports and the ratings of adult informants such as teachers and parents.

Gaps in Knowledge about Young Children's Self-Concept

Despite evidence that the self-concept becomes more differentiated with age (Harter, 1990; Marsh,

Craven, & Debus, 1991; Shavelson, Hubner, & Stanton, 1976), the age at which children possess more than a unidimensional view of themselves is unclear. Harter and Pike (1984) developed the PSPCSA to measure four a priori self-concept scales. Because factor analysis of the PSPCSA provided support for just two factors, a competence factor and a social acceptance factor, Harter and Pike concluded that the self-concept is not well differentiated in young children.

Recent evidence has begun to contradict this assertion. Using their Self-Description Questionnaire (SDQ-I), Marsh et al. (1991) found that kindergarten, first-, and second-grade children held reasonably differentiated perceptions of themselves in eight different self-concept domains, including academic, social, and physical dimensions of self. Eccles and her colleagues (Eccles et al., 1993) found that even first-grade children can distinguish between their own academic competence and achievement motivation. Research by Ladd (1990, 1996) revealed that kindergartners and first graders differentiate between how they are perceived by peers and how they view their own social efficacy. On the whole, these studies suggest that even kindergarten-age children hold more differentiated concepts of themselves than previously thought. However, clear evidence for the dimensionality of preschool children's self-concept has yet to be established.

Studies with elementary school-age children have typically found that children's self-appraisals become less positive with development and new experiences (Marsh et al., 1991; Stipek & Daniels, 1988; Wigfield et al., 1997). Researchers have attributed this decline to developmental factors, such as enhanced reasoning skills (Case, 1988) and greater use of self-other comparisons when evaluating one's own competence (Harter, 1990), as well as ecological explanations, such as the increasing importance of grades as children progress through school (Stipek, 1992).

Much less is known about the age at which the decline in positive self-appraisal begins. Harter and Pike (1984) found no age differences in their sample of preschool through second-grade children. In contrast, Marsh et al. (1991) found that three of the SDQ-I's eight self-concept scales (physical appearance, peer relationships, and general school self) showed modest declines in their cross-sectional study of kindergarten, first-, and second-grade children. However, Marsh et al. (1991) did not discuss that the kindergarten and first grade means were not significantly different and that in absolute terms, the mean self-perceptions of first graders were generally slightly higher than those of the kindergarten children. Although there appears to be consistent evidence that children's self-perceptions become less positive *after* first grade (Wigfield et al., 1997), longitudinal studies of young children are needed to determine whether this decline has already begun *prior* to school entry.

Research is also needed to understand more fully early gender differences in young children's self-perceptions. Studies by Eccles et al. (1993) and Marsh and Craven (1991) provide important evidence for the presence of counterbalancing differences in the patterns of boys' and girls' self-perceptions. For example, Eccles et al. (1993) found that first-grade boys had higher sports and math self-concepts whereas girls had higher reading self-concepts. In their sample of young children, Marsh and Craven (1991) found that boys had higher self-concepts of their physical ability whereas girls had higher self-concept scores on physical appearance and reading competence. In other words, researchers have begun to reveal stereotypic gender differences in the self-perceptions of elementary school children, but the age at which these differences first appear remains unclear.

Finally, researchers interested in psychosocial adaptation emphasize the importance of assessing children's functioning and impairment from different perspectives, including children's own self-reports (Achenbach, McConaughy, & Howell, 1987; Stanger & Lewis, 1993). In practice, however, this usually means adult perspectives only, regardless of how little agreement is typically found among mothers', fathers', and teachers' ratings of internalizing (r = .13) and externalizing behavior problems (r = .32) (Achenbach et al., 1987; Hinshaw, Han, Erhart, & Huber, 1992). Although researchers have come to view low agreement among informants as evidence of situational effects, the self-reports of young children are rarely accorded the same latitude because the level of agreement between children and adults tends to hover around .20 (Achenbach et al., 1987). We argue that because children are often the best informants about their own internal feelings of distress, more attention must be paid to the patterns of association between child self-report and the more commonly utilized assessments by adult informants.

Overcoming Obstacles to the Assessment of Young Children's Self-Perceptions Using Puppets

The task of eliciting reliable self-reports from young children is complicated by a variety of developmental factors. First, children have short attention spans, especially when working with paper and pencil measures (Irwin, 1985). Despite attempts to make these approaches more appealing through pictures and quantity icons, young children tend to favor toys

and life-like props for self-expression (Greenspan & Greenspan, 1991). Second, the language abilities of younger children are a concern, given that expressive skills tend to develop more slowly than receptive abilities (Foster, 1990). Methods that rely exclusively on children's verbal abilities are less likely to work with inhibited or language-delayed children. Third, as with most self-report measures, the validity of children's self-perceptions is often questioned because of possible response or motivational biases such as social desirability (Harter, 1990; Paulus, 1991), acquiescence and over-compliance (Garbarino & Stott, 1992), and defensive idealization (Measelle, 1995; Stipek, Recchia, & McClintic, 1992).

Finally, young children may have difficulty responding honestly in situations where adults are reading questions to them. However, when children are engaged in more age-appropriate activities, they talk openly about their experiences (Ceci & Bruck, 1993). Researchers who use dynamic techniques report that children of varied clinical and socioeconomic backgrounds are capable of relating both positive and negative aspects of their experience when they are involved in doll play or storytelling with actual figures (Bretherton, Ridgeway, & Cassidy, 1990; Mize & Ladd, 1988). Indeed, young children display enhanced verbal competence when engaged with a listener who can adjust for and find meaning in their expressions (Bogg & Eyberg, 1990; Schaefer & O'Conner, 1983). Researchers who use children's own lexicon and syntactic structure to inquire about concrete, behaviorally observable characteristics find that even 3½-year-olds can provide reliable descriptions of themselves (Eder, 1989).

Because they involve a dynamic medium of conversation, puppets can relate to children on their own terms and can respond flexibly depending on children's individual needs and language. Early studies showed that puppets could keep 4- to 7-year-old children from clinical and economically disadvantaged samples engaged, even when discussing anxiety-producing topics, such as death (Bernhart & Prager, 1985), sadness (Irwin, 1985; Schaefer & O'Conner, 1983), and aggressive behavior with peers (Woltmann, 1952). In more recent applications, puppets have been used effectively to assess young children's perceptions of themselves. Mize and Ladd (1988) found that 4- and 5-year-olds' enactments of their perceptions using puppets were better predictors of peer relationships than were the responses they gave through a picture medium. Eder (1989, 1990) has used puppets extensively in her research with 3½- to 7½-year-old children. Using two hand puppets and a structured-alternative format similar to Harter and Pike's (1984), Eder (1990) found that even the youngest children maintained psychologically meaningful conceptions of themselves. Unfortunately, Eder did not investigate the validity of her method by examining the associations between children's self-descriptions and external criteria such as equivalent parent or teacher ratings. Nevertheless, her findings provide some of the most direct support for the use of puppets as a systematic method to assess individual differences in young children's self-perceptions.

The Berkeley Puppet Interview

The BPI builds on the creative work of Eder (1990) and other investigators who have used puppets (Irwin, 1985; Schaefer & O'Conner, 1983; Woltmann, 1952) in studies of young children. The BPI method blends structured and clinical interviewing techniques to elicit children's self-perceptions in that the exchange between the children and the puppets is shaped by each child's own mode of communication (e.g., verbal or nonverbal) and style of interaction (e.g., inhibited or uninhibited). The goal of this method is to promote a peer-like exchange between a child and two puppets.

Like Harter and Pike's (1984) PSPCSA, the BPI presents bipolar options (both positive and negative poles) as a way to maximize comprehension while minimizing socially desirable responding. This approach has the added advantage of allowing children to hear possible and acceptable responses. However, in contrast to the PSPCSA's or the SDQ-I's (Marsh et al., 1991) use of explicitly stated and forced-choice response formats, the BPI encourages children to respond in a way that is most comfortable for them, be it verbal or nonverbal. Although most 4½- to 7½-year-olds interviewed with the BPI respond verbally, shy or less verbal children use limited verbal responses, such as naming a puppet to indicate their response, or nonverbal responses, such as pointing or placing objects in the puppets' mouths. In sum, by allowing each child to respond in his or her own idiosyncratic fashion, the BPI method attempts to promote a natural and unself-conscious dialogue between a child and the puppets.

The BPI adopts a multidimensional approach to the measurement of young children's self-perceptions (Shavelson et al., 1976) and builds on the theoretical idea that children can provide domain-specific descriptions and judgments of themselves when asked about salient aspects of their lives (Damon & Hart, 1988; Markus & Wurf, 1987). However, rather than attempt to assess all known facets of the young child's self-concept, our goal was to develop an age-

appropriate instrument that could be used easily to measure young children's perceptions of their school adjustment.

Along with others (Ladd, 1996; Sameroff & Haith, 1996), we define school adjustment as the degree to which children become engaged, interested, and successful in the academic and nonacademic aspects of school. Scholastically, children must contend with new demands on their cognitive capabilities while being challenged to cultivate an orientation toward school that will lead to achievement (Alexander & Entwisle, 1988; Eccles et al., 1993). Socially, children must negotiate new peer relationships while maintaining a sense of acceptance (Ladd, 1996). Underlying their academic and social adaptation is children's ability to regulate themselves emotionally and behaviorally, in particular, sad and anxious feelings that can lead to withdrawal and loneliness, and aggressive or hostile feelings that can lead to antisocial conduct (Ladd et al., 1988). In short, the degree to which children hold positive versus negative perceptions of themselves in different academic, social, and emotional domains is one important marker of school adjustment (Ladd, 1996).

In constructing individual scales, we choose not to develop item sets that reflected a single theory. Rather, the goal was to assess each domain broadly by drawing on theoretical perspectives that are not mutually exclusive and that have demonstrated strong validity in the literature. For example, the BPI's achievement motivation scale consists of items that tap children's school liking and valuing (Eccles et al., 1993; Ladd, 1996) as well as different facets of intrinsic motivation (Harter & Connell, 1984; Nicholls, 1989).

In the academic domain, given persuasive evidence that perceived academic competence and achievement motivation are distinct aspects of school adjustment (Eccles et al., 1993; Wigfield et al., 1997), separate measures of each were developed. Academic competence was defined as the combination of cognitive abilities and effective academic behaviors such as learning and good school work. Achievement motivation was defined as the valuing of school, preference for challenge, persistence, and a view of school as important and enjoyable.

In the social domain, separate measures of perceived social competence and of peer acceptance were developed, given evidence that children who experience peer difficulties may or may not express negative appraisals of their social abilities (Boivin, Poulin, & Vitaro, 1994). Social competence was defined as the ability to engage effectively in social tasks, such as making friends, asserting oneself in a socially appropriate manner, and seeking engagement over isolation. Peer acceptance was defined by a child's perception of himself or herself as liked and included by other children and as not ignored, excluded, or teased.

Finally, children's depressed-anxious feelings and aggressive-hostile feelings were measured separately, based on empirical support for the distinction between internalizing and externalizing symptoms in children (Achenbach et al., 1987). Drawing on evidence for the overlap between sadness, loneliness, and anxiety in young children (Achenbach et al., 1987; Kovacs, 1985), the BPI's depression-anxiety scale was defined by children's descriptions of themselves as sad, lonely, nervous, worried, and irritable. The BPI's aggression-hostility scale was defined by children's descriptions of their overtly hostile behavior toward peers and by their lack of empathy.

In sum, this study examines the psychometric properties of the self-perception scales of the BPI within and across three separate assessments of the same children: when they were in their final preschool year, kindergarten, and first grade. We then compared children's self-perceptions on the BPI with parents' and teachers' ratings of the child in similar domains, and with their scores on standardized achievement tests.

METHOD

Participants

Children were participants in a longitudinal prevention study that followed 97 two-parent families as their oldest child made the transition from preschool to kindergarten and first grade. Families, recruited to join the study through preschools, day-care programs, and local media, were predominantly lower-middle-class to upper-middle-class residents of the Greater San Francisco Bay area. Of the families in this community sample, 21% were of African American, Hispanic American, or Asian American ethnicity and 79% were Anglo-American. Of the 44 girls and 53 boys in the total sample, 27 served as a screening sample for the purpose of instrument development during the initial (preschool) year of the study. Therefore, in the present investigation, self-report data are presented on 73 children when they were in preschool, 97 when they were in kindergarten, and 95 when they were in grade 1. Children's mean ages were 4.6 years (SD = .47) prior to entering school, 5.9 years (SD = .38) at the end of kindergarten, and 6.9 years (SD = .35) at the end of first grade.

Procedures

Families interested in participating in a research project on children's transition to kindergarten entered the study during their oldest child's final year of preschool. Once enrolled in the study, 61 (63%) of the study's couples were randomly assigned to 16 week couples groups (three to four couples per group) *prior* to their child's entry into kindergarten. Over a 4 month period, group couples met with two licensed clinicians weekly to discuss marital or parenting issues. The goal of the couple's groups was to combine psychoeducation with insight-oriented discussions to enhance families' adaptation during this period of change (see Cowan & Cowan, 1997, for a thorough description of the prevention design). Parents not randomly assigned to one of the groups were offered one consultation per year with the project's clinical staff.

During each of the summers that families participated in the study—prior to kindergarten, following kindergarten, and following first grade—children were visited in their homes, where the puppet interview and standardized achievement testing were conducted. All three home visits were structured similarly. Upon arrival at a home, the experimenter and the child went to a room of the family's choosing. The visit began with 10–15 min of free play to establish rapport. When comfortable, children then participated in the first of two counterbalanced puppet interview modules (e.g., the self-perception interview or the perceived family environment interview). Children next completed some portion of a standardized achievement test: during the preschool visit, the entire Peabody Picture Vocabulary Test (PPVT) was administered; during the postkindergarten and post–grade 1 visits, the first three subtests of the Peabody Individual Achievement Test (PIAT) were administered. Testing was followed by a 15 min snack and free play break. After the snack, preschool children completed the second puppet interview; kindergarten and grade 1 children completed the last two PIAT subtests followed by the second puppet interview. Home visits ended with 10–15 min of free play. On average, home visits took about 2 hours to complete.

Parents independently complete mailed questionnaires about children's competencies and adjustment during the preschool period and while children were in kindergarten and first grade. Teachers' ratings of children's competencies and adjustment were collected during the fall and spring of kindergarten and first grade. In all but six cases, children's kindergarten and first grade teachers were different people. To insure that teachers were blind to the identity of the study's participants, teachers agreed to complete ratings on every child in their classroom.

The Berkeley Puppet Interview (BPI)

The BPI was developed to provide researchers with a semi-structured way to assess 4½- to 7½-year-old children's perceptions of (1) their academic, social, and emotional self-concept and (2) their family environment, including their parents' marital conflict and parent-child and sibling relationship quality. As the focus of the present investigation is on children's self-perceptions, a full description of the family environment module is omitted to save space (see Ablow & Measelle, 1995). The BPI Self-Perception Scales consist of six separate scales: two academic self-perception scales (academic competence and achievement motivation), two social self-perception scales (social competence and peer acceptance), and two symptom-related scales (depression-anxiety and aggression-hostility).

Item generation for these scales proceeded in three phases. First, after reviewing the self-concept literature, the first author (JM) generated a list of items inductively to correspond with self-perception domains. Initially this list included items in the academic and social domains only. After two of the three preschool waves had been interviewed, additional items were developed to measure children's appraisals of their aggressive-hostile and depressed-anxious feelings, including an adaptation of the short, 11 item version of the Child Depression Inventory (CDI; Kovacs, 1985). Because fewer data are available, children's preschool views of their symptomatology are not included in the present investigation. Second, a panel of experts attempted an independent classification of all items (once for the original academic and social items and later with the symptomatology items). With few exceptions, item classification by the second group was highly consistent (greater than 95% agreement) with the first author's original classification. Third, based on pilot interviews with 27 preschool children, a total of seven items were dropped, nine items were added, and a number of wording changes made to increase children's comprehension. This third step resulted in a final pool of 60 self-perception items.

Each BPI item consists of a pair of opposing, bipolar statements designed to reflect the positive and negative ends of different behaviors and attributes. Thirty-two of the original items focused on positive attributes and behaviors (e.g., "I'm good at making friends"/"I'm not good at making friends") and 28 of the original items focused on negative behaviors

and attributes (e.g., "I tease other kids"/"I don't tease other kids"). Items were counterbalanced so that children first heard the positive half of an item as often as they first heard the negative half. Positive and negative attributes were divided up equally between the two puppets so that children would not identify with either puppet.

Ten items were originally developed to assess perceived academic competence. Items in this scale were designed to ask children about their cognitive abilities, learning efficacy, and good school work. Nine items were designed to assess children's achievement motivation. Items in this scale asked children to describe their mastery orientation, motivation to learn, and interest in school. The social competence scale consisted of seven items, which ask children to assess how well they performed social tasks such as making friends and initiating play with other children. Nine items were developed to assess children's acceptance by peers. These items taped perceived affiliation, acceptance by peers, and rejection or victimization by peers. Symptomatology items were worded to use emotional state and trait terms that preschoolers have used to describe themselves in earlier research (Ridgeway, Waters, & Kuczaji, 1985). The depression-anxiety scale consisted of 16 items, including items that assess children's feelings of sadness, loneliness, self-worth, anxiety, and irritability. The aggression-hostility scale consisted of seven items, including measures of perceived conflict, absence of empathy, and antisocial behavior.

Procedures. Interviewers were upper-level undergraduate or graduate students who participated in a 2 day (16 hr) BPI training workshop that utilized a manual and training videotapes. Training emphasized the importance of creating and maintaining an open and natural dialogue between the puppets and the child. Considerable attention is paid to (1) building rapport between the puppets and the child, (2) insuring that children comprehend the method through the use of practice items or open-ended questions, (3) maintaining a neutral yet engaging interviewing style, (4) adjusting the interview process to children's verbal and nonverbal response styles, and (5) standardized use of rule-based techniques for eliciting codable responses.

During the second day of the training, trainees conducted two practice interviews which were then critiqued from videotape. Following the training workshop and prior to collecting actual data, interviewers conducted additional pilot interviews which were also critiqued by the first two authors. All interviewers trained in the method were ready to collect data after three to five pilot interviews. Of the 13 interviewers trained in the BPI during the method's development, only one was considered ineffective; those data were removed and the children (six cases) reinterviewed.

The BPI does not use a puppet theater or window, as we have found that the rapport between child and puppets is more natural and peer-like when they are face to face. Most interviews took place on the floor in children's bedrooms with the child sitting directly in front of the puppeteer's extended arms. Interviews were videotaped to allow the child experimenter to concentrate exclusively on his or her role as puppeteer.

Children are interviewed by two identical hand puppets, tan-colored puppy dogs that we named "Iggy" and "Ziggy," as indicated on their visible name tags. The interview begins with Iggy and Ziggy explaining that each will say something about themselves to the child and then, "We want to learn about you." Three to four neutral practice items (e.g., "I like/don't like pizza") are administered to acclimate the child to the BPI method. During the interview, one puppet might say "I have lots of friends at school," the second puppet would say "I don't have lots of friends at school," and then the child would be asked by one of the puppets, "How about you [child's name]?"

The BPI method does not use a forced-choice response format or recognition task as is the case with most other self-perception measures for children this age (Eder, 1990; Harter & Pike, 1984; Marsh et al., 1991). Rather, children are allowed to respond in whatever way is comfortable to them. Most responded verbally or made partial verbal statements that paralleled or modified one of the item-halves offered by the puppets. A few of the younger children responded nonverbally by pointing to a puppet or by shaking their head in agreement or disagreement. If a child's nonverbal manner of responding was ambiguous, the puppeteer tried to establish a clearer means of communication with the child. For example, if a child nodded or shook his or her head "yes" or "no," the puppets encouraged the child to point as well.

Scoring. All interviews were scored by two coders from videotape. Training in the BPI's scoring system takes approximately 2 hr of formal didactics and an additional 3–5 hr of practice scoring. Based on the degree to which children's responses parallel one of the puppet's statements, responses are coded on a 7 point (1–7) Likert-style scale where the endpoints represent very negative (1) and very positive (7) self-perceptions. On the BPI's 7 point scale, the numbers 2 and 6 represent responses that are equivalent to one

of the bipolar statements made by either puppet. For example, if a child responds negatively by indicating "I'm a dumb kid too" this would be scored a 2 because the response directly reflects what one of the puppets said. However, if a child were to amplify on the original negative statement by indicating "I'm *really* dumb," this response would be scored a 1 to represent a response that was more negative than the puppet's original statement. If the child expressed a less emphatic negative self-perception (e.g., "I'm *kind of* dumb") or felt the need to qualify or to circumscribe the response (e.g., "I'm kind of dumb in counting"), such responses would be scored a 3. Positive responses are coded on the 5 to 7 range of the scale, where 5 is a less emphatic positive response than a 6 response (e.g., "I'm *pretty* smart" versus "I'm smart") and 7 represents a more positive response than the puppet's original statement (e.g., "I'm super smart"). If children indicate that "both" options pertain to them or that they are "in the middle," their responses are coded a 4. Although children's responses can be complex at times, agreement among coders is high: 97.6%, 94.7%, and 98.2% agreement, respectively, for children's preschool, kindergarten, and first-grade responses.

The BPI asks young children a number of potentially anxiety-provoking questions. Nevertheless, the clear majority of children interviewed enjoyed the experience and were able to complete the interview. Even children who reported negative self-perceptions, including sad and lonely children, were able to complete the BPI. In the nearly 300 interviews on which the BPI was developed, only five children were unable to complete the BPI. Indeed, families repeatedly expressed that the BPI method was an event that children anticipated with excitement. Moreover, the method has proven to be an effective way to collect data with less verbal and/or inhibited children because of the medium (e.g., a "puppet show") and because it is a dynamic process that can be adjusted for individual differences in children.

Additional Measures

Ratings of children's behavior by adult informants. Teachers, mothers, and fathers completed the Child Adaptive Behavior Inventory (CABI). To Schaefer and Hunter's (1983) 60 item Child Behavior Inventory (CBI), a total of 31 items were added from a downward extension of the Quay-Peterson Behavior Problem Checklist (O'Donnell & VanTuinen, 1979), and from Achenbach and Edelbrock's (1983) Child Behavior Checklist (CBCL), with additional items designed specifically for the present study to cover problems among peers. Responses on the CABI are scored on a 4 point Likert scale anchored at 1 ("not at all like") and 4 ("very much like").

In the present investigation six multiscale CABI scores were created to parallel the six BPI scales. These composite scales and their component, multiitem scales were rationally constructed scales, the same for teachers and parents: (1) Academic Competence was defined by two scales: cognitive competence (e.g., "Is smart for his/her age") and academic creativity (e.g., "Shows creativity in his/her work"; "Thinks up interesting things to do"); (2) Achievement Motivation had only one scale that consisted of items tapping task- and learning-oriented behaviors (e.g., "Works carefully and does his/her best"); (3) Social Competence was defined by three scales: social efficacy (e.g., "Makes friends easily"), extroversion (e.g., "Has an outgoing personality"), and prosocial behavior (e.g., "Gets along well with other kids"); (4) Peer Acceptance was defined by two scales: social acceptance (e.g., "Is liked by other kids") and social victimization (reversed) (e.g., "Is picked on by other kids"); (5) Internalizing was defined by four scales: depression (e.g., "Is sad a lot"), somatic complaints (e.g., "Often complains about not feeling well"), anxiety (e.g., "Worries a lot"), and irritability-oppositionality (e.g., "Is stubborn or irritable"); and (6) Externalizing was defined by three scales: antisocial behavior (e.g., "Tends to disobey or break rules"), hostility (e.g., "Has a hot temper"), and empathy (reversed) (e.g., "Is concerned about others' feelings"). Across all scales, teachers' (alpha *range* = .79–.95), mothers' (alpha *range* = .82–.91), and fathers' (alpha *range* = .80–.93) ratings of children's behaviors demonstrated acceptable internal consistency.

To index the validity of parents' perceptions, we computed the average correlations between mothers' and fathers' CABI ratings across each domain during the preschool, $r(69) = .40$, $p < .01$, kindergarten, $r(88) = .52$, $p < .001$, and first-grade years, $r(88) = .51$, $p < .001$. These data are encouraging in that parents appear to overlap in their perceptions of children's competence and well-being when rated with the CABI. However, their views are not so highly correlated as to suggest redundancy. As such, in the analyses performed in this study, we examined parents' ratings separately.

Teachers in this study agreed to use the CABI to describe every child in their class and to remain blind to the one or, occasionally, two pupils from their classroom in this study. Although teacher CABI ratings were collected on approximately 1,500 children, rounded, only the ratings of the project's 97 partici-

pants were analyzed in the present study. Given that teachers' ratings came from approximately 93 kindergarten and 95 first-grade classes, scores on each teacher-rated item were standardized within each classroom and by child gender so as to represent participants' behavior relative to their classmates of the same gender. Standardizing teacher's ratings of a child within his or her classroom controlled for the fact that teachers (*a*) utilize rating scales differently, (*b*) often form impressions of individual children by comparing them to other children in the classroom, and (*c*) in different classrooms have different proportions of children with behavior and emotional problems. Standardizing teacher's ratings by gender within classrooms controlled for the possibility that teachers might provide highly gendered ratings in certain domains, for example, internalizing and externalizing behavior problems.[1]

In the validational analyses that follow, children's self-perceptions were compared with mothers' and fathers' ratings collected concurrently in all 3 years. Because not all children were in the same type of preschool and because we did not collect ratings from preschool teachers, the following correlational strategy was used when examining child-teacher agreement. Children's preschool self-perceptions, obtained during the summer just before kindergarten, were correlated with kindergarten teachers' fall ratings. Children's kindergarten self-perceptions, obtained during the summer following kindergarten, were correlated with kindergarten teachers' spring ratings. Similarly, children's first-grade self-perceptions, obtained during the summer following first grade, were correlated with first-grade teachers' spring ratings. In short, we attempted to validate the BPI using parent and teacher ratings closest in time.

Standardized achievement test scores. Following kindergarten and first grade, children's academic achievement was assessed using the Peabody Individual Achievement Test—Revised (PIAT; Markwardt, 1989). The PIAT is a widely used achievement test that assesses students' progress in six standardized domains. In the present study, only children's math and overall reading test scores were analyzed.

1. As suggested by one reviewer, standardization of teachers' ratings by classroom and child gender might affect the degree of correspondence between teacher-child or teacher-parent ratings. As a preliminary check, we correlated child self-perceptions and parent ratings with teachers' ratings when they were standardized by classroom only and when they were standardized by classroom and gender (as reported in this article). In general, the differences between sets of coefficients were not statistically significant. However, one pattern did seem to emerge: the correlation between mothers' ratings of depression and teachers' ratings of depression were .06 larger each year, on average, when we used teachers' ratings that were standardized by classroom and gender. Although beyond the scope of this study, the standardization of teachers ratings by classroom and gender may affect differently the degree of correspondence with parent ratings depending on the construct of interest.

Preliminary Data Analyses

These data were first analyzed to determine whether there were any intervention effects present in the child and informant data. The intervention was examined in three ways: (1) as a correlate of the child and adult data (e.g., an effect code of 1 or 0 for the intervention and nonintervention conditions); (2) as a control in partial correlations (e.g., the intervention term was controlled when examining BPI consistency coefficients); and (3) as a control in the repeated-measure ANOVAs (e.g., when examining mean-level changes over time). In all three analytic approaches, significant intervention effects emerged in less than 5% of the analyses. The results of this study are therefore presented *without* controlling for the intervention conditions. The fact that very few intervention effects emerged in the child self-perception data in particular was not unexpected. In other analyses (Cowan, Cowan, Heming, & Boxer, 1995), there is consistent evidence that the project's couples intervention had its most direct effect on the marriage or on mothers' and fathers' parenting. Although there are likely to be a number of indirect ways in which preventive interventions with parents will help to shape children's adaptation, it is also possible that a longer time period will be needed to detect differences in children's adjustment.

BPI data were also examined for interviewer effects. None of these analyses proved significant, suggesting that individual differences in children's responses on the BPI were not attributable to differences in the effectiveness or style of the systematically trained interviewers used in this study.

RESULTS

Construction of the BPI Scales

Because the dimensions we hoped to measure with the BPI were known beforehand, scale development proceeded in a rational, two-step manner. First, preliminary scales were created using all theoretically relevant items. Each item was assigned to a single scale. Only the academic competence, achievement motivation, social competence, and peer acceptance items were administered to the full sample in all 3 years. Thus, data on children's perceptions of their depressed-anxious and aggressive-hostile

Table 1 Psychometric Analyses of Rationally Constructed Academic, Social, and Emotional Self-Concept Scales on the BPI

Scales	Items	Alpha	Mean *rij*	*M*	*Range*
Academic competence:					
Preschool	6	.76	.32	.49	.23–.65
Kindergarten	6	.75	.33	.48	.29–.65
Grade 1	6	.70	.28	.44	.26–.79
Scale *M*	6	.74	.31	.47	.26–.70
Achievement motivation:					
Preschool	7	.74	.29	.45	.22–.87
Kindergarten	7	.68	.23	.37	.24–.53
Grade 1	7	.73	.28	.43	.18–.63
Scale *M*	7	.72	.27	.42	.21–.68
Social competence:					
Preschool	5	.65	.27	.40	.26–.62
Kindergarten	5	.63	.25	.39	.23–.59
Grade 1	5	.62	.24	.38	.26–.49
Scale *M*	5	.63	.25	.39	.25–.57
Peer acceptance:					
Preschool	8	.68	.23	.38	.15–.55
Kindergarten	8	.75	.27	.43	.32–.50
Grade 1	8	.70	.21	.34	.17–.45
Scale *M*	8	.71	.24	.38	.21–.50
Depression-anxiety:					
Preschool	N.A.				
Kindergarten	10	.76	.31	.49	.20–.65
Grade 1	10	.78	.26	.44	.38–.52
Scale *M*	10	.77	.29	.47	.29–.59
Aggression-hostility:					
Preschool	N.A.				
Kindergarten	5	.70	.28	.40	.24–.53
Grade 1	5	.70	.31	.43	.31–.52
Scale *M*	5	.70	.30	.42	.29–.53
M	7	.71	.27	.42	.25–.60

Note: Rationally developed items in the internalizing and externalizing domains were not added to the BPI until two-thirds of the sample had completed preschool.

feelings are reported for kindergarten and first grade only. The coefficient alphas of the preliminary rational scales indicated that our conceptually derived scales demonstrated modest, albeit encouraging internal consistency during the last preschool year (*range* = .59–.72), kindergarten (*range* = .63–.73), and first-grade years (*range* = .62–.75). To increase the homogeneity of each scale, we then eliminated a total of 18 items whose item-total correlations averaged less than .20 across all 3 years. The final or surviving items are presented in the Appendix.

The outcome of this two-step procedure is presented in Table 1. Each of the six scales demonstrated acceptable levels of internal consistency, with all but the social competence scale demonstrating internal consistencies that exceeded .70 across all 3 years. These data suggest that the items designed to reflect a common dimension appeared to hang together reasonably well when children were in preschool, kindergarten, and first grade.

Next we conducted a factor analytic check of the data to evaluate further the reliability and dimensionality of young children's self-perceptions on the BPI. Separate principal-component analyses (PCA) of the item-level data conducted by year supported the number and type of scales indicated by our preliminary analyses.[2] Results of the PCAs and of the

2. To determine the optimal number of factors needed to represent children's BPI responses, three different statistical criteria were used to guide decisions. First, the eigenvalue distribution (i.e., Cattell's scree test) indicated a four-factor solution in conjunction with the preschool data and six-factor solutions in conjunction with the kindergarten and grade 1 data. Second, parallel analysis was used to compare the empirically derived eigenvalues-values against the eigenvalues generated by a series of Monte Carlo simulations of our data (see Zwick & Velicer, 1986). In three separate Monte Carlo runs, our empirically derived eigenvalues consistently exceeded or approximated that of the largest random factor. Finally, we used the interpretability of oblique-rotated solutions as added criteria. Note that a rerotation of these factors to a VARIMAX solution produced simi-

oblique-rotated solutions (Oblim) are presented in Table 2. We present the oblique solutions given previous evidence that children's self-perceptions tend to be correlated meaningfully in different domains. As shown in Table 2, items had moderate to high factor loadings on their designated factor with one consistent exception. In kindergarten and first grade, three peer acceptance items loaded highly, if not more strongly on the depression scale. The eigenvalues for children's kindergarten and first-grade perceptions of their acceptance by peers were .87 and .93, respectively, slightly lower than the recommended 1.0. Nonetheless, given the overall support provided by the factor analytic check, we decided to retain the rationally constructed scales as they were originally constructed.

Table 3 presents the means, standard deviations, and scale intercorrelations for the four preschool scales and six kindergarten and six first-grade scales. Scale scores were computed by taking the unweighted average of children's responses to the items in each domain. Consistent with earlier studies (Eccles et al., 1993; Marsh et al., 1991; Stipek et al., 1992), scale means were skewed in the positive direction, reflecting young children's tendency to see themselves positively. The standard deviations suggest, however, that there was a fair degree of variability in children's responses on the BPI.

As anticipated, scales were correlated within each year, although not so highly as to suggest redundancy. Although the year-to-year differences in mean intercorrelations were not statistically significant (Steiger, 1980), the average correlation among scales declined each year, consistent with the idea that children's self-perceptions become more differentiated with age (mean $r = .36$, .31, and .28 during preschool, kindergarten, and first grade, respectively).

Consistency over Time

The 1 and 2 year consistency coefficients are given in Table 4. Although not equivalent to short-term reliability, these data provide an estimate of the BPI's long-term consistency. Across all domains, children's views demonstrated only modest consistency across the 2 year period from preschool through first grade (mean $r = .29$). The 1 year consistency coefficients were uniformly stronger than the 2 year coefficients, especially the 1 year coefficients from kindergarten to first grade (mean $r = .51$). On the whole, these data suggest that young children do have coherent and relatively consistent conceptions of themselves across the period of their lives when they are going from being at home to becoming elementary school children. On the other hand, the modest nature of these stability coefficients, especially children's perceived academic competence, also suggests that the perceptions of many young children do shift as they begin their academic careers.

Gender and Grade Differences in the Level of Young Children's Self-Perceptions

A series of repeated-measures ANOVAs was conducted to examine the possible effects of gender and grade (preschool, kindergarten, and grade 1) on children's self-perceptions in each BPI domain. The results of these analyses are presented in Table 5. Except for perceived academic competence, $F(2, 136) = 3.22$, $p < .05$, none of the gender $\times$ grade interaction terms was significant. Boys' perceived academic competence remained stable across the three time periods, whereas girls' sense of their academic competence tended to decline significantly. Gender effects were not present in the achievement motivation scale or the two social scales of the BPI. However, in stereotypic ways, gender, but not grade level, was statistically related to mean-level differences in the symptomatology scales. Specifically, girls' mean score on the depression-anxiety scale was significantly higher than boys', whereas boys' mean score on the aggression-hostility scale was significantly higher than girls'.

Rather than decline as predicted, boys' and girls' self-reported achievement motivation, social competence, and peer acceptance scores became more positive with age, Fs(2, 136) = 8.38, $p < .001$, 4.65, $p < .01$, and 5.58, $p < .01$, respectively. Planned t tests showed that children's achievement motivation, $t(67) = 2.17$, $p < .05$, and peer acceptance scores, $t(67) = 2.47$, $p < .05$, increased significantly between preschool and kindergarten but *not* between kindergarten and first grade. In the social competence domain, the only significant difference was between children's preschool and first-grade scores, $t(67) = 2.33$, $p < .05$, suggesting a gradual but significant increase across this 2 year period.

How should we interpret the fact that children's self-perceptions on the BPI increased in several do-

lar results. In addition to the hypothesized four- (preschool year) and six-factor solutions (kindergarten and first grade), we forced additional solutions as added tests of dimensionality. Overall, these procedures produced solutions that closely approximated our hypothesized BPI dimensions. The one exception was the peer acceptance factor, whose items tended to load on the depression factor, when five factors were rotated, or to lose the majority of its items to the depression factor, when seven factors were rotated.

Table 2 Factor Pattern of Children's Preschool, Kindergarten, and Grade 1 Responses of the BPI Self-Perception Scales

Scale and Item[a]	Preschool Factors				Kindergarten Factors						First Grade Factors					
	1	2	3	4	1	2	3	4	5	6	1	2	3	4	5	6
Academic competence:																
4. I'm a smart girl/boy		.75					.66						.76			
5. I do a good job in school	.48	.63					.77						.68			
11. I'm not dumb		.68					.72						.81			
24. I learn things well		.73					.69						.72			
30. I'm smarter than other kids		.56					.55	.39					.51			
33. I don't make mistakes a lot		.69					.62				.41		.55			
Achievement motivation:																
1. I like school work that's hard	.75					.60						.66				
7. When things are hard, I keep trying	.59	.40				.60	.41					.78				
15. Don't give up if work is hard	.69					.71						.70				
37. School's important	.81					.73						.68				
38. I like school	.70					.78						.69				
40. I try my best at school	.78					.61	.49					.58				
48. When school is hard, I try my best	.51					.72						.74				
Social competence:																
3. I'm not shy			.68					.74							.69	
14. Not hard to make friends	.37		.62					.66	.35						.71	
34. I ask kids to play			.59					.69							.66	
42. Easy to get kids to like me			.61					.72							.73	
50. If kids are playing, I ask to play			.66					.73							.71	
Peer acceptance:																
12. I have lots of friends			.54	.61						.65	−.41					.60
20. Kids like me				.72						.59						.81
21. Kids pick me for their games			.48	.57	−.46					.44						.69
23. Kids don't tease me				.74						.66						.75
29. Kids are nice to me				.72						.73	−.48					.62
35. Kids ask me to play				.58	−.48					.56	−.53					.48
49. I'm a good friend to have				.71	−.51					.48					.39	.53
53. Kids don't say mean things to me				.59						.69						.60
Depression-anxiety:																
2. I'm not a happy kid					.71						.60					−.46
9. I'm sad a lot					.73						.72					
16. I get cranky a lot					.58			−.44			.51					
19. I worry a lot					.68						.77					
25. I cry a lot					.73						.84					
27. I'm lonely a lot					.66					−.39	.69					−.54
31. I get nervous if teacher calls on me					.64						.68					
41. Bad things are going to happen to me					.76						.52					
43. I get nervous at school					.60						.67					
58. I get mad when I make mistakes					.52				.39		.61					
Aggression-hostility:																
6. If someone is mean, I hit them									.69					.75		
8. It's funny when a friend gets in trouble									.58					.63		
22. I don't feel bad after I fight									.67					.53	.37	
45. I tease other kids									.74					.71		
56. I pick on other kids									.76					.72		

Note: Only cross-loadings >.35 printed. Items that load higher on nondesignated factors are underlined.

[a] Items are listed in an abbreviated form; see Appendix for actual item wording.

Table 3 Means and Intercorrelations among BPI Self-Perceptions Scales by Year of Assessment

BPI Scales by Year	*M*	*SD*	Academic Competence	Achievement Motivation	Social Competence	Peer Acceptance	Depression-Anxiety
Preschool:							
Academic competence	5.48	.94	—				
Achievement motivation	4.91	1.21	.43**				
Social competence	4.49	1.17	.24*	.30*			
Peer acceptance	5.01	1.02	.49***	.24*	.44**		
Kindergarten:							
Academic competence	5.36	1.32	—				
Achievement motivation	5.28	1.41	.27**				
Social competence	4.52	1.63	.20	.30**			
Peer acceptance	5.32	1.22	.36**	.35**	.25**		
Depression-anxiety	2.70	1.27	−.38***	−.27**	−.32**	−.58***	
Aggression-hostility	2.61	1.20	−.02	−.27**	−.28**	−.40***	.41***
Grade 1:							
Academic competence	5.49	1.18	—				
Achievement motivation	5.45	1.70	.29**				
Social competence	4.76	1.55	.25**	.30**			
Peer acceptance	5.45	1.01	.31**	.29**	.28**		
Depression-anxiety	2.56	1.02	−.24**	−.30**	−.28**	−.42***	
Aggression-hostility	5.40	1.38	.00	−.24**	−.25**	−.36**	.44***

Note: Symptomatology items were not asked while children were in preschool. $n = 70$ (preschool); 94 (kindergarten); 96 (grade 1).
* $p < .05$; ** $p < .01$; *** $p < .001$; one-tailed.

mains rather than decrease as reported in earlier studies (Eccles et al., 1993; Marsh & Craven, 1991)? It is possible that these mean changes are due to the continued effects of overly optimistic self-assessments by young children. Because boys' and girls' perceived academic competence was either stable or declined across the preschool through first-grade period, age-related forms of optimism may not accurately describe young children's self-perceptions in all domains. Alternatively, given the emphasis placed on developing school readiness skills during the preschool and kindergarten years (Meisels, 1996), these mean-level increases may reflect real behavioral change about which children are aware.

We examined this behavioral change hypothesis by testing for mean-level changes in mothers' and fathers' ratings of children's behavior in each of the BPI domains.[3] Consistent with children's self-reports, mothers' and fathers' ratings of children's achievement motivation, social competence, and peer acceptance increased over time, $Fs(2, 136) = 5.42$, 3.97, and 4.23, respectively, $ps < .01$, suggesting that the changes reported by children may reflect actual changes in their behavior. In contrast to the child

3. In the present research, teachers' ratings were not used to examine mean level changes, as these ratings were standardized within classrooms by gender. Thus, age effects were removed from teachers' data.

Table 4 Temporal Consistency of BPI Self-Perception Scales

BPI Scales	PreK to Grade 1[a] (2 Years)	PreK to K[a] (1 Year)	K to Grade 1[b] (1 Year)	*Mean r*
Academic competence	.24	.29	.32	.28
Achievement motivation	.25	.41	.54	.40
Social competence	.30	.46	.51	.39
Peer acceptance	.31	.49	.55	.42
Depression-anxiety	...	...	.58	.58
Aggression-hostility	...	...	.56	.56
Mean r	.28	.41	.51	

Note: One-tailed tests of significance. All $rs > .25$, $p < .05$. All $rs > .28$, $p < .01$.
[a] $n = 69$.
[b] $n = 92$.

Table 5 Sex and Grade Effects in the BPI Self-Perceptions Scales

BPI Scale and Sex	Grade			Test Statistic (*F*)		
	PreK	Kindergarten	Grade 1	Sex	Grade	Sex × Grade
Academic competence:						
Boys	5.37	5.43	5.32	.37	3.65*	3.22*
Girls	5.60	5.20	5.08			
Achievement motivation:						
Boys	5.07	5.28	5.53	.57	8.38***	.60
Girls	4.81	5.30	5.42			
Social competence:						
Boys	4.51	4.60	4.75	.27	4.65**	.77
Girls	4.35	4.47	4.86			
Peer acceptance:						
Boys	4.90	5.38	5.40	.32	5.58**	1.34
Girls	5.07	5.13	5.38			
Depression-anxiety:						
Boys	...	2.52	2.41	3.42*	2.54	.68
Girls	...	2.88	2.87			
Aggression-hostility:						
Boys	...	2.84	2.76	3.37*	1.34	.92
Girls	...	2.48	2.40			

Note: $N = 70$ in these analyses, as only 70 children were available for the preschool analyses. We analyzed the same 70 children when analyzing the two symptomatology scales.
* $p < .05$; ** $p < .01$; *** $p < .001$.

data, neither the main effects of gender, grade nor the child gender × grade interaction were significant in mothers' or fathers' ratings of their children's academic competence. Although daughters' academic self-perceptions declined with time, parents' ratings of academic competence were affected neither by their child's gender nor age. In contrast to the gender differences found in children's reports of their symptomatology, gender was not a significant factor in parents' ratings of internalizing or externalizing behavior problems. At this age and on these dimensions, gender appears to play a role in children's self-perceptions but not in parents' perceptions of their child.

Correlations between Children's Self-Perceptions and Informants' Ratings

Next we examined the correspondence between children's self-perceptions on the BPI with their teachers', mothers', and fathers' ratings in similar domains. We also correlated teachers' and mothers' ratings and then compared this pair of cross-informant coefficients to pairs of informants that included children's self-reports (e.g., teacher-mother versus child-teacher). Although the field tends to discount the self-reports of young children in favor of pairs of adult informants, in particular, mothers and teachers (Achenbach et al., 1987), we hypothesized that the level of agreement between teachers and mothers would not exceed significantly the level of agreement between each child-adult combination. A summary of these correlations is presented in Table 6.

In the school domain, children's perceived academic competence and achievement motivation were consistently correlated with their teachers' ratings in all 3 years. Children who saw themselves as academically competent and achievement oriented tended to be described by teachers as more competent and more mastery oriented in school. Maternal ratings of academic competence and achievement orientation showed little concordance with children's self-perceptions during preschool, but mothers who saw their children as academically competent and motivated tended to have children who saw themselves similarly during kindergarten and first grade. Fathers' and children's ratings in these domains were unrelated until grade 1, at which point they were modestly associated in the predicted direction.

In the social domain, children's perceived acceptance by peers was significantly related to their teachers' ratings of peer acceptance during preschool, kindergarten, and first grade and with mothers' ratings of peer acceptance once children had entered elementary school. In other words, children reporting positive peer acceptance on the BPI tended to be seen by their teachers and mothers as positively engaged with peers. Fathers' ratings of their children's acceptance by peers were unrelated to the children's self-report from preschool through first grade. Of the six

Table 6 Summary of Cross-Informant Ratings

BPI Scales by Year	Cross-Informant Ratings[a]			
	Child-Teacher	Child-Mother	Child-Father	Teacher-Mother
Preschool:[b]				
Academic competence	.24	.09	.10	.16
Achievement motivation	.32*	.20	.02	.21
Social competence	.29*	.21	.10	.26*
Peer acceptance	.35*	.11	.09	.40**
Mean r	.30*	.16	.09	.26*
Kindergarten:[c]				
Academic competence	.34**	.25*	.16	.14
Achievement motivation	.42***	.33**	.18	.18
Social competence	.19	.10	.00	.29**
Peer acceptance	.31**	.33**	.12	.39***
Depression-anxiety	.33**	.31**	.24	.26*
Aggression-hostility	.42***	.28**	.29**	.33**
Mean r	.34**	.27**	.17	.27*
Grade 1:[c]				
Academic competence	.31**	.28**	.25*	.15
Achievement motivation	.44***	.36**	.26*	.28**
Social competence	.23	.14	.09	.24*
Peer acceptance	.38***	.34**	.01	.44***
Depression-anxiety	.35**	.34**	.22	.29**
Aggression-hostility	.40***	.30**	.31**	.34**
Mean r	.35**	.30**	.19	.29**

Note: Mean *r*s were computed using the Fischer *r* to *z* transformations.

[a] Teachers, mothers, and fathers each rated children in domains that were conceptually similar to the six BPI scales.

[b] $n = 69$.

[c] $n = 88$.

* $p < .05$; ** $p < .01$; *** $p < .001$, two-tailed.

BPI scales, only children's perceived social competence scores consistently demonstrated little or no correspondence with adult informants' ratings of their social competence.

Given that little concordance is usually found among informants' reports of children's behavioral and emotional problems (Achenbach et al., 1987; Stanger & Lewis, 1993), the BPI's symptomatology scales demonstrated encouraging criterion validity. Perhaps most notable in these results, children who saw themselves as aggressive and hostile were typically rated by all three adult observers as more antisocial during kindergarten and first grade. During kindergarten and first grade, children's perceptions of their depressed-anxious feelings were significantly related to teachers', mothers', and, to a lesser extent, fathers' ratings of their internalizing problems.

Next, we compared the correlations between children's views and those of each of the adult informants with the mother-teacher correlations. As shown in Table 6, child-teacher *r*s exceed mother-teacher *r*s in all 3 years, except in the peer acceptance domain in which agreement between teachers and mothers was consistently better. Although the differences between these correlation coefficients were not statistically significant, a sign test (Loftus & Loftus, 1988) showed that the child-teacher *r*s exceeded the mother-teacher *r*s (in 11 of 16 comparisons) significantly more often than would be predicted by chance ($p = .05$). Child-mother *r*s were smaller than the mother-teacher *r*s during preschool but tended to be comparable in size during kindergarten and first grade; none of the differences was statistically significant. Child-father *r*s were consistently less than mother-teacher *r*s, especially in the peer acceptance domain in which testing revealed significantly lower child-father agreement. These data suggest that agreement between children's self-reports and ratings by teachers or mothers tend to be as strong as if not stronger than the level of agreement between adult informants.

Correlations between Children's Self-Perceptions and Achievement Test Scores

As a final test of validity, children's kindergarten and first-grade self-perceptions were correlated with

Table 7 Correlations between BPI Self-Perception Scales and Standardized Test Scores

Test Scores	Academic Competence	Achievement Motivation	Peer Acceptance	Social Competence	Depression-Anxiety	Hostility-Aggression
Kindergarten self-perceptions:						
Math achievement score	.19	.26**	.11	.13	−.34**	−.16
Reading achievement score	.19	.30**	.07	.01	−.22*	−.13
Grade 1 self-perceptions:						
Math achievement score	.24*	.35**	.03	−.03	−.30**	−.14
Reading achievement score	.23*	.33**	.03	−.08	−.16	−.13

* $p < .05$; ** $p < .01$, two-tailed.

their standardized achievement test scores in math and reading. These data are presented in Table 7. Results indicate that children's perceived achievement motivation was associated with their scores in math and reading following kindergarten and first grade. Children who described themselves as more academically motivated tended to earn higher scores in math and reading than children who reported lower motivation. Although children's perceived academic competence and achievement test scores were unrelated at the end of kindergarten, they were modestly related to their math and reading test scores following first grade.

Except for the depression-anxiety scale, none of the other BPI scales was significantly correlated with children's math and reading test scores. The general lack of association between children's test scores and their nonacademic self-perceptions provides evidence of the BPI's specificity. The significant association between the BPI depression-anxiety scale and the two PIAT scores was unexpected, yet it is consistent with competency-based models of depression (Cole, 1991). Children reporting higher levels of depression and anxiety at the end of kindergarten and first grade earned lower math and reading scores at the end of kindergarten and lower scores in math after first grade.

DISCUSSION

The goal of this study was to examine the psychometric properties of the Berkeley Puppet Interview's (BPI) Self-Perception Scales. Given the absence of self-concept research with young children, particularly longitudinal studies of the young self, this study was also designed to examine (1) the dimensionality of children's self-concept during the preschool and early elementary school years, (2) the consistency and stability of boys' and girls' self-perceptions, and (3) the levels of agreement between children's self-reports and those of adult informants' as well as between children's self-perceptions and their scores on standardized achievement tests.

The results from this study suggest that young children's self-perceptions can be measured reliably with the BPI. During preschool, kindergarten, and first grade, children produced internally consistent responses that typically exceeded .70 alpha in each BPI domain. Although not as reliable as the self-reports of adolescents or adults, reliability estimates of this magnitude have not typically been found with children as young as 4½ years (Hughes, 1984; Wylie, 1989). A factor analytic check of the data further confirmed that the a priori dimensions were reliably measured and that the factor structure of children's responses on the BPI was reasonably well defined and invariant across three points in time. In sum, these data add to the evidence (Eccles et al., 1993; Ladd, 1996; Marsh et al., 1991) that young children's self-perceptions are more differentiated than originally suggested by Harter and Pike's (1984) two-factor solution. Indeed, when assessed with age-appropriate methods such as the BPI, our results suggest that this differentiation may begin even earlier than kindergarten and first grade, perhaps even during the preschool period in children as young as 4½ years.

Across time, changes in the intercorrelations among the BPI scales provided only minimal support for the hypothesis that children's self-perceptions become more differentiated with age (Marsh et al., 1991; Shavelson et al., 1976). Whereas the average correlation among the BPI scales did decline within each year, the differences across years were not statistically significant. Despite evidence from cross-sectional studies that children's self-perceptions become more differentiated across the elementary school years (Eccles et al., 1993; Marsh & Craven, 1991), these longitudinal data suggest that the process of differentiation during the preschool through grade 1 period may not be as pronounced. In fact, Marsh and his colleagues (1991) found little differ-

ence between the differentiation of kindergarten and first-grade children, whereas second-grade children demonstrated greater differentiation in their self-perceptions than the kindergarten and first-grade children.

Consistency and Stability

Given its longitudinal design, this study was in a unique position to examine individual differences in the consistency of young children's self-perceptions as well as normative age and gender trends during their transition to elementary school. As predicted, children's academic, social, and symptom self-perceptions demonstrated significant, albeit modest, 2 year consistency and considerably stronger 1 year consistency, especially between the end of kindergarten and the end of first grade. In keeping with studies that have documented the stability of early personality traits (Block, Block, & Keys, 1988; Caspi, 1987), competencies (Alexander & Entwisle, 1988), and the persistence of emotional and behavioral problems in children (Kellam, Rebok, Mayer, Ialongo, & Kalodner, 1994), these data suggest that young children have conceptions of themselves that are reasonably consistent across time. On the other hand, the size of these continuity estimates also indicates that a good deal of change takes place during this time in children's lives. Because the preschool to early elementary school period is marked by considerable developmental and contextual change for most children (Cowan et al., 1994), data suggesting shifts in children's views of themselves seem realistic.

It is generally thought that children's self-perceptions become increasingly less positive with age (Eccles et al., 1993; Marsh & Craven, 1991; Stipek & Daniels, 1988), yet the research is considerably less clear with regard to young children, in those not yet in first grade. In the present investigation, boys and girls generally reported positive scholastic and social self-perceptions and low levels of depression-anxiety and hostility-anger before they entered kindergarten and *increasing* achievement motivation, social competence, and peer acceptance as they completed kindergarten and grade 1. The fact that mothers' and fathers' ratings in conceptually similar domains also increased significantly in a positive direction suggests that these trends in children's self-reports may reflect actual changes in competence following the transition to school. Indeed, one of the primary goals of kindergarten is to encourage children's interest in learning as well as to increase their level of socialization through cooperative work and play (Ladd, 1996). It is possible that the trend toward increased negative self-evaluation frequently found in other studies does not set in until the more rigorous and structured primary grades.

The findings with regard to academic competence results are complex. Recall that during preschool, boys and girls reported comparably high levels of academic competence. Of the six BPI scales, only the academic competence scores declined steadily with age and solely among girls. It cannot be determined from this study whether this decline in girls' academic self-perceptions is attributable to their sensitivity to evaluative feedback (Roberts, 1991), to differential effects of family or school socialization practices (Block, 1983; Stevenson & Newman, 1986), or to other explanations entirely. Nonetheless, these data are consistent with studies of older children and adolescents that often find gender differences in instrumental aspects of self-concept, many of which tend to favor boys (Roberts, 1991).

The two symptomatology scales were also significantly affected by gender but not age in this study in somewhat stereotypic ways. Girls' mean depressed-anxious scores in kindergarten and first grade were significantly greater than boys', while at the same time boys rated themselves higher in aggression and hostility than did girls. Clear gender differences in the level of young children's internal distress are not usually found, possibly because most studies have relied on adult informants who, by most accounts, tend to be less aware of children's internal distress (Achenbach et al., 1987; Keenan & Shaw, 1997). Perhaps more alarming is the fact that this gender difference among young children is consistent with the preponderance of depressive symptoms that emerge most clearly in girls during their early adolescence.

The finding of greater aggression in boys is consistent with other studies of adults' ratings of preschool and early elementary school children (Hinshaw et al., 1992; Keenan & Shaw, 1997). The gender differences found here in boys' and girls' reports of their externalizing behaviors may be due to the fact that the BPI's aggression-hostility scale contains measures of overt aggression only. Crick (1995) has found that when measures of relational aggression are included, boys and girls often report similar levels although different forms of aggression. The inclusion of items tapping relational aggression (Crick, 1995) might eliminate the gender difference in overt aggression reported in this study.

Child-Adult Correspondence

Is the BPI a valid measure of children's preschool, kindergarten, and grade 1 self-perceptions? Results

from three sets of analyses provide encouraging evidence that it is. First, we found consistent convergence between children's self-reports and ratings by teachers and mothers on scales that were conceptually relevant to each of the BPI's self-concept scales. Moreover, the average level of agreement between children and teachers was statistically significant at all three grade levels and statistically significant between children and mothers once children entered kindergarten and first grade. The correlations between children's reports and fathers' were not significant during preschool and kindergarten. However, the fact that there were consistent child-teacher and child-mother associations as well as the fact that there was significant child-father agreement during first grade suggests that fathers may become more accurate informants the longer their children are in school.

Second, the average level of agreement between children and teachers and between children and mothers consistently exceeded the average level of agreement between mothers and teachers. Although the differences were not statistically significant, these findings support the idea that young children can provide meaningful information about their competencies and problems and that reports by older informants need not substitute for young children's self-reports, as is typically the case in much research.

Finally, young children's self-perceptions of their achievement motivation and academic competence were related to their math and reading achievement test scores. Researchers consider the prediction of real-world variables essential for evaluating the validity of most behaviorally relevant constructs (Robins, John, & Caspi, 1994). Indeed, in previous studies, children's performances on standardized achievement tests have been shown to be important correlates of academic, social, and psychological adaptation during the preschool to early elementary school period (Cowan et al., 1994; Hinshaw, 1992).

Limitations in Generalizing the Current Findings

Caution should be exercised when generalizing these results. First, replication with a larger and more diverse sample is clearly warranted. Second, although 20% of our sample were members of racial and ethnic minority groups—not far from Bay Area representation—the sample consisted primarily of Anglo-American children from middle-class backgrounds. Before it can be concluded that the BPI is a developmentally appropriate method that works for most children, its utility must be investigated in culturally and socioeconomically diverse samples. We believe that the puppet method and the fact that nonverbal answers are permitted (pointing) makes the BPI useful for a wide range of young children, but this is a matter now under investigation by researchers in this and other laboratories. Third, attention should be paid to the BPI's social competence scale given its lack of criterion validity. Items in the social competence scale were intended to discriminate between children who feel able or unable to make friends. Given that only five items were retained, expanded coverage of social competence could improve the scale's validity. Ablow and Measelle continue to look for ways to improve the performance of each BPI item (e.g., better wording and age-appropriate syntax) as well as to add other relevant domains. For example, we have expanded the number of symptomatology scales to cover a greater range of psychopathology (e.g., attention deficits, separation anxiety, enuresis) and the usefulness of these newer scales is currently being assessed in a large sample of clinic and nonclinic referred children. In short, the BPI's development is an ongoing process.

Conclusion: Costs versus Benefits of the BPI

The BPI is a more labor-intensive instrument than most standardized child self-report scales in interviewer training, data-gathering, and response coding phases. The interviewers must have sensitivity and clinical skills and be trained to be appropriately engaging with the child and to be able to modify the interview flexibly in response to a child's style of engagement and responses. Although the BPI's authors have found that a variety of people can function as skilled puppeteers, we have also seen how data can be compromised when interviewers are not sufficiently trained. Time is also a factor in administration (about 25 min on the average) and the preparation of data (responses must be coded from videotapes after the interview). An obvious question remains: Do the benefits of using the instrument outweigh the costs?

Using the BPI, we have found that young children possess a multidimensional self-concept that can be reliably measured. It is possible that the sensitivity of the method has helped to reveal age-related characteristics of the young self, in particular the self-perceptions of preschoolers, that have gone undetected or misinterpreted with other instruments. Overall, our results showed that the BPI is sensitive to normative changes and individual differences in 4½- to 7½-year-olds' views of their academic and social self-perceptions as well as their affective and behavioral adjustment. The BPI's validity draws support from the consistent pattern of associations be-

tween children's self-perceptions and ratings by adult informants. Indeed, our findings suggest that the concordance between young children's self-reports and parallel ratings by teachers or mothers are *at least* as strong as the concordance between mothers' and teachers' ratings. The present findings are encouraging. A precise evaluation of costs and benefits can come only in future studies of young children that compare the BPI directly with other self-perception instruments.

ACKNOWLEDGMENTS

Work on this project was supported by National Institute of Mental Health grants to Jeffrey R. Measelle (1 F3 MH10816), Jennifer C. Ablow (1 F3 MH11038), and Philip A. Cowan and Carolyn P. Cowan (MH-31109). Work on the development of training materials for the Berkeley Puppet Interview was supported by a grant to Jeffrey Measelle and Jennifer Ablow from the MacArthur Foundation Research Network on Psychopathology and Development (PR-10004). We are indebted to Rhona Weinstein, Stephen P. Hinshaw, Oliver P. John, and Helena C. Kraemer for their helpful comments on earlier drafts of this article. We also gratefully acknowledge the contributions of Joan Kaplan during the early development of the Berkeley Puppet Interview, the many research assistants who collected and coded data, in particular, Kaly Nguyen, and the families participating in the Schoolchildren and Their Families Project.

ADDRESSES AND AFFILIATIONS

Corresponding author: Jeffrey R. Measelle, Department of Psychiatry and Behavioral Sciences, Child Division, Stanford University, 401 Quarry Road, Stanford, CA 94305-5719; e-mail: measelle@leland.stanford.edu. Jennifer C. Ablow is at Stanford University, and Philip A. Cowan and Carolyn P. Cowan are at the University of California, Berkeley.

APPENDIX

ITEMS USED TO ASSESS CHILDREN'S SELF-PERCEPTIONS ON THE BERKELEY PUPPET INTERVIEW

Academic Competence

4. I'm smart./I'm not smart.
5. I do a good job in school./I don't do a good job in school.
11. I'm dumb./I'm not dumb.
24. I don't learn things well./I learn things well.
30. I'm good at letters and writing./I'm not good at letters and writing.
33. I make mistakes in school a lot./I don't make mistakes in school a lot.

Achievement Motivation

1. I like school work that's hard./I don't like work that's hard.
7. When things are hard for me, I keep trying./When things are hard for me, I do something else.
15. When I can't figure something out, I give up./When I can't figure something out, I don't give up.
37. I think it's important to do well in school./I don't think it's important to do well in school.
38. I like school./I don't like school.
40. I try my best at school./I don't try my best at school.
48. When school is hard, I try my best./When school is hard, I don't try my best.

Social Competence

3. I'm shy when I meet new people./I'm not shy when I meet new people.
14. It's not hard for me to make new friends./It's hard for me to make new friends.
34. I ask kids to play with me./I don't ask kids to play with me.
42. It's hard for me to make new friends./It's easy for me to make new friends.
50. If kids are playing together, I watch them./If kids are playing together, I ask if I can play.

Peer Acceptance

12. I have lots of friends at school./I don't have lots of friends at school.
20. Kids don't like me./Kids like me.
21. Kids let me play games with them./Kids don't let me play games with them.
23. Kids at school tease me./Kids at school don't tease me.
29. Kids are nice to me./Kids are not nice to me.
35. Kids at my school don't ask me to play with them./Kids at my school ask me to play with them.
49. Kids think I'm a good friend to have./Kids do not think I'm a good friend to have.
53. Kids say mean things to me./Kids do not say mean things to me.

Depression-Anxiety

2. I'm not a happy (boy/girl)./I'm a happy (boy/girl).
9. I'm sad a lot./I'm not sad a lot.
16. I get cranky a lot./I don't get cranky a lot.
19. I worry a lot./I don't worry a lot.
25. I cry a lot./I don't cry a lot.
27. I'm lonely a lot./I'm not lonely a lot.

31. I don't get nervous if my teacher calls on me./I get nervous if my teacher calls on me.

41. Bad things are going to happen to me./Bad things are not going to happen to me.

43. I get nervous and scared at school./I don't get nervous and scared at school.

58. I get mad when I make mistakes./I don't get mad when I make mistakes.

Aggression-Hostility

6. If someone is mean to me, I hit them./If someone is mean to me, I don't hit them.

8. I think it's funny when a friend of mine gets in trouble./I don't think it's funny when a friend gets into trouble.

22. I feel bad after I fight with kids./I don't feel bad after I fight with kids.

45. I tease other kids./I don't tease other kids.

56. I pick on other kids at school./I don't pick on other kids at school.

REFERENCES

Ablow, J. C., & Measelle, J. R. (1993). *Berkeley Puppet Interview: Administration and Scoring System Manuals.* Berkeley: University of California.

Ablow, J. C., & Measelle, J. R. (1995). The Berkeley Puppet Interview: Children as informants on family and self. Paper presented at the biennial meeting of the Society for Research in Child Development, Indianapolis.

Achenbach, T. M., & Edelbrock, C. S. (1983). *Manual for the Child Behavior Checklist and Revised Child Behavior Profile.* Burlington, VT: Author.

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101,* 213–232.

Alexander, K. L., & Entwisle, D. R. (1988). Achievement in the first two years of school: Patterns and processes. *Monograph of the Society for Research in Child Development, 53*(2, Serial No. 218).

Bernhart, B. R., & Prager, S. B. (1985). Preventing child suicide: The elementary school death education puppet show. *Journal of Counseling and Development, 63,* 311–312.

Block, J. H. (1983). Differential premises arising from differential socialization of the sexes: Some conjectures. *Child Development, 54,* 1335–1354.

Block, J., Block, J. H., & Keys, S. (1988). Longitudinally foretelling drug use in adolescence: Early childhood personality and environmental precursors. *Child Development, 59,* 336–355.

Bogg, S. R., & Eyberg, S. (1990). Interview techniques and establishing rapport. In A. M. La Greca (Ed.), *Through the eyes of the child: Obtaining self-reports from children and adolescents* (pp. 85–108). Boston: Allyn & Bacon.

Boivin, M., Poulin, F., & Vitaro, F. (1994). Depressed mood and peer rejection in childhood. *Development and Psychopathology, 6,* 483–498.

Bretherton, I., Ridgeway, D., & Cassidy, J. (1990). Assessing internal working models of the attachment relationship: An attachment story completion task for 3-year-olds. In M. T. Greenberg, D. Cicchetti, & E. M. Cummings (Eds.), *Attachment in the preschool years: Theory, research, and intervention* (pp. 273–308). Chicago: University of Chicago Press.

Byrne, B. M. (1996). Measuring self-concept across the life span: Issues and instrumentation. Washington, D.C.: American Psychological Association.

Case, R. (1988). The whole child: Toward an integrated view of young children's cognitive, social, and emotional development. In A. Pellegrini (Ed.), *Psychological bases for early education.* Chichester: Wiley.

Caspi, A. (1987). Personality in the life course. *Journal of Personality and Social Psychology, 53,* 1203–1213.

Cicchetti, D. (1993). Developmental psychopathology: Reactions, reflections, projections. *Developmental Review, 13,* 471–502.

Ceci, S. J., & Bruck, M. (1993). Suggestibility of child witness: A historical review and synthesis. *Psychological Bulletin, 113,* 403–439.

Cole, D. A. (1991). Preliminary support for a competency-based model of depression in children. *Journal of Abnormal Psychology, 100,* 181–190.

Cowan, C. P., & Cowan, P. A. (1997). Working with couples during stressful transitions. In S. Dreman (Ed.), *The family on the threshold of the 21st century.* Hillsdale, NJ: Erlbaum.

Cowan, C. P., Cowan, P. A., Heming, G., & Boxer, C. C. (1995). *Preventive interventions with parents of preschoolers: Effects of marriage, parenting, and children's development.* Paper presented at the biennial meeting of the Society for Research in Child Development, Indianapolis.

Cowan, P., Cowan, C. P., Schulz, M., & Heming, G. (1994). Prebirth to preschool family factors in children's adaptation to kindergarten. In R. D. Parke & S. G. Kellam (Eds.), *Exploring family relationships with other social contexts* (pp. 75–113). Hillsdale, NJ: Erlbaum.

Crick, N. R. (1995). Relational aggression: The role of intent attributions, feelings of distress, and provocative type. *Development and Psychopathology, 7,* 313–322.

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin, 115,* 74–101.

Damon, W., & Hart, D. (1988). *Self-understanding in childhood and adolescence.* New York: Cambridge University Press.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95,* 256–273.

Eccles, J., Wigfield, A., Harold, R. D., & Blumfeld, P. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development, 64,* 830–847.

Eder, R. A. (1989). The emergent personologist: The structure and content of 3½-, 5½-, and 7½-year-olds' concepts

of themselves and other persons. *Child Development, 60,* 1218–1228.

Eder, R. A. (1990). Uncovering young children's psychological selves: Individual and developmental differences. *Child Development, 61,* 849–863.

Fantuzzo, J. W., McDermott, P. A., Holliday Manz, P., Hampton, V. R., & Alvarez Burrdick, N. (1996). The Pictorial Scale of Perceived Competence and Social Acceptance: Does it work with low-income urban children? *Child Development, 67,* 1071–1084.

Foster, S. H. (1990). *The communicative competence of young children.* London: Longman.

Garbarino, J., & Stott, F. M. (1992). *What children can tell us: Eliciting, interpreting, and evaluating critical information from children.* San Francisco: Jossey-Bass.

Greenspan, S. I., & Greenspan, N. T. (1991). *The clinical interview of the child* (2d ed.). Washington, DC: American Psychological Association.

Harter, S. (1990). Developmental differences in the nature of self-representations: Implications for the understanding, assessment, and treatment of maladaptive behavior. *Cognitive Therapy and Research, 14,* 113–142.

Harter, S., & Connell, J. P. (1984). A model of children's achievement and related self-perceptions of competence, control, and motivational orientation. In J. Nicholls (Ed.), *The development of achievement motivation* (pp. 33–56). Greenwich, CT: JAI.

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

Hinshaw, S. P. (1992). Externalizing behavior problems and academic underachievement in childhood and adolescence: Causal relations and underlying mechanisms. *Psychological Bulletin, 111,* 127–155.

Hinshaw, S. P., Han, S. S., Erhart, D., & Huber, A. (1992). Internalizing and externalizing behavior problems in preschool children: Correspondence among parent and teacher ratings and behavior observations. *Journal of Clinical Child Psychology, 21,* 143–150.

Hughes, H. (1984). Measures of self-concept and self-esteem for children ages 3–12 years: A review and recommendations. *Clinical Psychology Review, 4,* 657–692.

Irwin, E. C. (1985). Puppets in therapy: An assessment procedure. *American Journal of Psychotherapy, 39,* 389–399.

Joseph, J. (1979). *Joseph Pre-School and Primary Self-Concept Screening Test.* Wood Dale, IL: Stoelting.

Keenan, K., & Shaw, D. (1997). Developmental and social influences on young girls' early behavior problems. *Psychological Bulletin, 121,* 95–113.

Kellam, S. G., Rebok, G. W., Mayer, L. S., Ialongo, N., & Kalodner, C. R. (1994). Depressive symptoms over first grade and their response to developmental epidemiologically based preventive trails aimed at improving achievement. *Development and Psychopathology, 6,* 463–481.

Kovacs, M. (1985). The Children's Depression Inventory (CDI). *Psychopharmacology Bulletin, 21,* 995–1124.

Ladd, G. W. (1990). Having friends, keeping friends, making friends, and being liked by peers in the classroom: Predictors of children's early school adjustment? *Child Development, 61,* 1081–1100.

Ladd, G. W. (1996). Shifting ecologies during the 5 to 7 year period: Predicting children's adjustment during the transition to grade school. In A. J. Sameroff & M. M. Haith (Eds.), *The five to seven year shift: The age of reason and responsibility* (363–386). Chicago: University of Chicago Press.

Ladd, G. W., Price, J. M., & Hart, D. (1988). Predicting preschooler' peer status from their playground behaviors. *Child Development, 59,* 986–992.

Loftus, G. R., & Loftus, E. F. (1988). *Essence of statistics* (pp. 152–197). New York: Knopp.

Markus, H., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology, 38,* 299–337.

Markwardt, F. C. (1989). *Peabody Individual Achievement Test—Revised.* Circle Pines, MN: American Guidance Service.

Marsh, H. W., & Craven, R. G. (1991). Self-other agreement on multiple dimensions of preadolescent self-concept: Influences by teachers, mothers, and fathers. *Journal of Educational Psychology, 83,* 393–404.

Marsh, H. W., Craven, R. G., & Debus, R. (1991). Self-concept of young children 5 to 8 years of age: Measurement and multidimensional structure. *Journal of Educational Psychology, 83,* 377–392.

Measelle, J. (March, 1995). *Positive illusions and mental health? Reformulating the question in terms of young children.* Poster presented at the biennial meeting of the Society for Research in Child Development, Indianapolis.

Meisels, S. J. (1996). Performance in context: Assessing children's achievement at the outset of school. In A. J. Sameroff & M. M. Haith (Eds.), *The five to seven year shift: The age of reason and responsibility* (407–434). Chicago: University of Chicago Press.

Mize, J., & Ladd, G. W. (1988). Predicting preschoolers' peer behavior and status from their interpersonal strategies: A comparison of verbal and enactive responses to hypothetical social dilemmas. *Developmental Psychology, 24,* 782–788.

Nicholls, J. G. (1989). *Competence and accomplishment: A psychology of achievement motivation.* Cambridge, MA: Harvard University Press.

O'Donnell, J. P., & VanTuinen, M. (1979). Behavior problems of preschool children: Dimensions and congenital correlates. *Journal of Abnormal Child Psychology, 7,* 61–75.

Paulus, D. L. (1991). Measurement control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of social psychological attitudes: Vol. 1. Measures of personality and social psychological attitudes* (pp. 17–59). San Diego, CA: Academic Press.

Ridgeway, D., Waters, E., & Kuczaji, S. A. II. (1985). Acquisition of emotion-descriptive language: Receptive and productive vocabulary norms for ages 18 months to 6 years. *Developmental Psychology, 21,* 901–908.

Roberts, T. A. (1991). Gender and the influences of evaluation on self-assessments in achievement settings. *Psychological Bulletin, 109,* 297–308.

Robins, R. W., John, O. P., & Caspi, A. (1994). Major dimen-

sions of personality in early adolescence: The Big Five and beyond. In C. F. Halverson, Jr., G. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood.* Hillsdale, NJ: Erlbaum.

Rutter, M. (1989). Isle of Wight revisited: Twenty-five years of child psychiatric epidemiology. *Journal American Academy of Child and Adolescent Psychiatry, 28,* 633–653.

Sameroff, A. J., & Haith, M. M. (1996). Interpreting developmental transitions. In A. J. Sameroff & M. M. Haith (Eds.), *The five to seven year shift: The age of reason and responsibility* (pp. 3–16). Chicago: University of Chicago Press.

Schaefer, C., & O'Conner, K. (1983). *Handbook of play therapy.* New York: Wiley.

Schaefer, E. S., & Hunter, W. M. (1983). *Mother-infant interaction and maternal psychosocial predictors of kindergarten adaptation.* Paper presented at the biennial meeting of the Society for Research in Child Development, Detroit.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Reviews of Educational Research, 46,* 407–441.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87,* 245–251.

Stevenson, H. W., & Newman, R. S. (1986). Long-term predictions of achievement and attitudes in mathematics and reading. *Child Development, 57,* 646–659.

Stipek, D. (1992). The child at school. In M. H. Bornstein & M. E. Lamb (Eds.), *Developmental psychology: An advanced textbook* (3d ed., pp. 579–625). Hillsdale, NJ: Erlbaum.

Stipek, D. J., & Daniels, D. (1988). Declining perceptions of competence: A consequence of changes in the child or in the educational environment? *Journal of Educational Psychology, 80,* 352–356.

Stipek, D., Recchia, S., & McClintic, S. (1992). Self-evaluation in young children. *Monographs of the Society for Research in Child Development, 57*(1, Serial No. 226).

Stanger, C., & Lewis, M. (1993). Agreement among parents, teachers, and children on internalizing and externalizing behavior problems. *Journal of Clinical Child Psychology, 22,* 107–115.

Vaughn, B. E. (1995). *Psychometric and validity concerns regarding the Pictorial Scale of Perceived Competence and Acceptance for children.* Paper symposium presented at the biennial meeting of the Society for Research in Child Development, Indianapolis.

Wigfield, A., Eccles, J. S., Kwang, S. Y., Harold, R. D., Arbreton, A. J. A., Freedman-Doan, C., & Blumenfeld, P. (1997). Changes in children's competence beliefs and subjective task values across the elementary school years: A 3-year study. *Journal of Educational Psychology, 89,* 451–469.

Woltmann, A. G. (1952). Puppet shows as a psychotherapeutic method. In L. Bender (Ed.), *Child psychiatric techniques.* Springfield, IL: C. C. Thomas.

Wylie, R. C. (1989). *Measures of self-concept.* Lincoln: University of Nebraska Press.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99,* 432–442.