

A Validation Study of the Penn Interactive Peer Play Scale with Urban Hispanic and African American Preschool Children

By: Marcelo Castro, Julia L. Mendez, John Fantuzzo

Castro, M., [Mendez, J.L.](#), & Fantuzzo, J. (2002). A validation study of the Penn Interactive Peer Play Scale with urban Hispanic and African American preschool children. *School Psychology Quarterly*, 17(2), 109-127. DOI.1521/scpq.17.2.109.20856

Made available courtesy of American Psychological Association:

<http://www.apa.org/journals/spq/>

This article may not exactly replicate the final version published in the APA journal. It is not the copy of record

*****Note: Figures may be missing from this format of the document**

Promoting young children's readiness to learn is a national priority. With the arrival of the new millennium, it is appropriate to reassess our National Education Goal that all American children will start school ready to learn by the year 2000 (U.S. Department of Education, 1992). While the 1990s have yielded improvements in the status of young children's well-being, too many young children still face "socially-toxic" environments that threaten the development of competencies necessary for early school success (Garbarino, 1995). Approximately one in five American children currently lives in poverty (Children's Defense Fund, 1998), and is exposed to multiple risk factors—including poor health care, lack of appropriate housing, family stress, and community violence—that severely threaten their development (Huston, McLoyd, & Garcia Coll, 1994). In addition, risk associated with membership in low-income families disproportionately affects children from ethnic minority communities (McLoyd, 1990, 1998; Swanson & Spencer, 1991).

Poor children are more likely than nonpoor children to have difficulties in school, ranging from behavioral and emotional difficulties to retention and poor school performance (Duncan, Brooks-Gunn, & Klebanov, 1994; Huston, McLoyd, & Garcia Coll, 1994). These risks are especially pronounced for children living in densely populated urban areas, who are exposed disproportionately to risk factors associated with less desirable educational outcomes (BrooksGunn, Duncan, & Aber, 1997). As a result, school psychologists in urban communities are faced with the increasing need to engage in greater outreach to preschool children and their families in order to facilitate elementary school adjustment for incoming students (Carlton & Winsler, 1999; Rimm-Kaufman & Pianta, 1999).

Within the early childhood arena, compensatory educational programs have been developed to promote school readiness and social competence among young children by providing a buffer against the ill effects of poverty and urban living. Early childhood education is often guided by a developmental-ecological perspective that asserts the importance of supporting children's acquisition of age-appropriate competencies within relevant social contexts (Bredekamp & Copple, 1997). For preschool children, relating successfully with classmates is a critical,

developmentally salient task that is a primary indicator of healthy adjustment (Cicchetti, 1990; Vaughn et al., 2000). The development of positive peer relationships during the preschool years has been associated with positive adjustment in kindergarten as well as academic success in the elementary grades and high school (Ladd, Kochenderfer, & Coleman, 1996; Ladd & Price, 1987). In contrast, longitudinal research has linked poor peer relations with detrimental consequences during later developmental periods, including emotional maladjustment, delinquent behavior, and school failure (Denham & Holt, 1993; Parker & Asher, 1987; Parker, Rubin, Price, & DeRosier, 1995).

Classroom peer play represents a primary context in which preschool children acquire and express social competencies (Fantuzzo, Coolahan, Mendez, McDermott, & Sutton-Smith, 1998; Raver & Zigler, 1997). The repeated interpersonal interactions that occur during peer play, especially those involving prosocial behavior or aggressive encounters, are important experiences that impact children's social development (Fisher, 1992; Ladd & Price, 1987). Exposure during play to the ideas and feelings of peers enables children to move beyond egocentric thought to consider perspectives of others. Research on peer play with African American children has established relations among successful play interaction and children's adaptable temperament, emotion regulation, receptive language, and expressive language (Mendez, Fantuzzo, & Cicchetti, in press). Peer acceptance is also associated with higher levels of student motivation and engagement in school, which are key components of school success (Coolahan, Fantuzzo, Mendez, & McDermott, 2000; Wentzel, 1999).

Quality assessment of children's peer play is critical in order to create appropriate programming and curricula designed to enrich children's social relationships (NAEYC, 1991; Raver & Zigler, 1997). Despite the existence of a large body of work on children's interpersonal competence (see Hart, Olsen, Robinson, & Mandelco, 1997 for a review), the assessment of interactive peer play among diverse groups is confined to a few studies. Critiques of existing measurement tools show that most scales were developed to assess functioning for groups of white middle-class children (Kennedy, 1988; McLoyd, 1998). In general, the lack of standardized assessment instruments developed for use with minority preschool children limits the quantitative research focus to labor-intensive, observational work. Other weaknesses of prior studies include lack of differentiation between influences of culture, social class, or even features of culture such as linguistic or ethnic differences (Phinney & Landin, 1998).

A recommended research strategy for obtaining more information about ethnic minority children's behavioral repertoire is to develop and validate culturally relevant assessment instruments (Guerra & Jagers, 1998; Mendez, McDermott, & Fantuzzo, 2002). An efficient approach to obtaining information about large numbers of at-risk children is to utilize behavior ratings scales. Specifically, teachers have been shown to be reliable raters of children's social behavior (Connolly & Doyle, 1984; Hoge & Colardaci, 1989). Studies have documented the validity of teacher judgments in distinguishing students with social and behavioral difficulties from those without problems (Gresham, Macmillan, & Bocian, 1996). Other studies supported the accuracy of teacher judgment in the identification of students with internalizing and externalizing behavioral disorders (Fantuzzo et al., 1995; McConaughy, 1993).

In response to the need for culturally sensitive and developmentally appropriate assessments of social competence, Fantuzzo and colleagues (1995) developed the Penn Interactive Peer Play Scale (PIPPS). This instrument is aimed at identification of children's behavioral strengths and needs within the context of peer play in the classroom or home environments. One of the advantages of this instrument is that the scale items were developed with the assistance of teachers and parents of minority preschool children from low-income families to ensure appropriate and meaningful item content. A second advantage is that this measure can be used as an assessment tool to inform school-based intervention programs (see Fantuzzo et al., 1997 for an illustration). The psychometric properties of the scale are also strong, with studies establishing reliability among teachers, parents, and teacher assistants from different educational and ethnic back-grounds, and validity using multiple sources of social competence data (Fantuzzo et al., 1995; Fantuzzo & Coolahan et al., 1998). A limitation of the current PIPPS instrument is that it has not been tested with preschool children outside of the Northeast region of the country, nor has it been translated for use with Spanish-speaking adult raters.

Studies of specific populations that differ along major demographic or social position variables (i.e., income, education, immigrant status) are necessary in order to identify relevant developmental competencies promoted and supported by a particular culture (Garcia Coll et al., 1996). For example, Gaskins's (1999) observational data regarding Mayan infant exploratory play were not interpreted using a deficit, comparative perspective with American infants. Instead, researchers considered a prevailing Mayan cultural belief that child development is "primarily internally generated," with children's play being a "self-motivated" activity that is "primarily independent from adult supervision and participation" (p. 49), to explain cross-cultural differences in play behavior. Other work that investigates and informs play constructs that are relevant for diverse cultural groups is critical in order to produce high-quality assessment of social behavior (Weinberger & Starkey, 1994). Moreover, this need is most pressing for young children from Hispanic backgrounds, who represent the fastest growing population of children in the United States. Projections indicate that Hispanics will likely become the largest U.S. ethnic minority population as early as 2010. By the year 2020, more than one-fifth of all U.S. children are projected to come from Hispanic backgrounds, a figure that is almost double the proportion found in 1990 (McLoyd, 1998; U.S. Census Bureau, 1999; U.S. Department of Health and Human Services, 1997).

Given the need for high-quality assessment for Hispanic preschool children, the present study was designed to investigate the generalizability of the validated dimensions of the PIPPS (Rogers, 1998). Establishing generalizability of findings requires empirical tests to determine if findings obtained through the study of one population are comparable or equivalent for another distinct population, also referred to as "ecological representativeness" (Kerlinger, 1996, p. 300; Messick, 1995). This study examined the generalizability of play construct research initially conducted with low-income African American children in the Northeastern region of the United States to a sample of low-income Hispanic and African American children in the Southeastern United States. To accomplish this task, a Spanish version of the original PIPPS was developed using recommended translation procedures (Guerra & Jagers, 1998). The primary objectives of the study were to (a) produce a Spanish translation of the PIPPS, and (b) collect evidence of construct and concurrent validity for this rating scale with a new population. It was hypothesized that the ratings obtained from Spanish-speaking early childhood educators would yield a latent

structure of three play constructs. This research also allowed for a replication of the PIPPS dimensions with African American children in a different geographic region of the U.S. Finally, differences in interactive peer play dimensions across ethnicity and gender were examined in the Southeast sample.

METHOD

Participants

Eight preschools were selected as research sites to represent the cultural and ethnic diversity of the children receiving subsidized childcare services in MiamiBade County, Florida. Within the target area, significant numbers of Hispanic and African American children attended the local preschools. Recruitment efforts focused on centers with higher participation rates by Spanish-speaking families in order to secure participation of a significant number of Hispanic children in the research project. For inclusion in the study, participating children needed to be between 3 and 5 years of age and receive subsidized preschool childcare under Child Development Services. Only one child per family was eligible for the study.

Ethnicity and race were operationalized for purposes of this research to provide demographic information that helps characterize the members of this community. African American children were defined as English-speaking, American children who were of African descent. Other cultural groups that are sometimes encompassed under a term of African American, such as Haitian children, were not included in the sample due to variation that arises from a dialect of Haitian Creole used by some families. In sum, families in the African American category were English-speaking, while families in the Hispanic category (of any racial background) spoke Spanish in the home. Close to two-thirds of Hispanic participants within this study were of Cuban background. Other descriptive data on Hispanic families, including degree of acculturation or bilingualism, were not collected due to limited access to family composition information. Reported family income did not exceed \$20,000 per year.

A total of 400 parents were offered participation in the study. Of these, 64 declined. Across families, there was participation by male and female parents, although this information was not recorded. Of the remaining 336 participants, 12 families were dropped from the sample due to missing information. Parents in four cases also asked the researchers to have their child removed from the study for individual reasons. The final sample consisted of 320 preschool children enrolled in 24 classrooms. Children were between the ages of 3 and 5 years ($M = 3$ years, 10 months, $SD = 8$ months). Females comprised 45.6% of the sample. Reported ethnicity indicated that 231 Hispanic children (72.2%) and 89 African American children (27.8%) participated in the study.

Any child who was classified or receiving services under any disability category was excluded from the sample, excepting children who were receiving speech services. This decision controlled for the amount of time that children spent outside of the classroom receiving services related to their particular disability. Speech and language disabilities are diagnosed at a much higher rate in subsidized child care and the associated services generally do not prevent students from participating in regular classroom activities, as compared with students with physical disabilities or other classifications.

A total of 24 female early childhood educators provided ratings of children's social behavior. Teacher age ranged from 21 to 48 years with a mean age of 35. Eighteen teachers were fluent in Spanish and from Hispanic backgrounds. Six teachers were English-only speakers and 1 of these teachers was of Hispanic ethnicity. The remaining 5 English-only speakers were African American females. Teaching experience varied from 10 months to 5.5 years with the mean years of experience approximately 3 years. Only 4 participating teachers possessed a bachelor's degree. The remainder of the teachers had obtained a high school diploma and had completed 30 hours of childcare training as mandated by the state department with jurisdiction over early childhood education.

Procedures

Parental consent was obtained by conducting group meetings at the eight preschool centers. In addition, the research objectives were also explained to teachers during staff meetings. All participating teachers received training on performing observations prior to completion of measures. The primary author directly addressed questions and concerns that parents had regarding the study. Data were collected from teachers within a 2-month time period. Eighteen Spanish-speaking early childhood educators of Hispanic background were asked to complete the Spanish-translated PIPPS measure, while six English-speaking teachers completed the English version of the PIPPS. Five of these English-only speakers were of African American background while one teacher was of Hispanic background. The data collection process yielded 203 Spanish version PIPPS and 117 English version PIPPS teacher forms. The majority of Spanish PIPPS were ratings of Hispanic children ($n = 191$), while the majority of English PIPPS were ratings of African American children ($n = 77$). Forty Hispanic children were rated with an English PIPPS and 12 African American children were rated with a Spanish PIPPS.

Instrumentation

For research purposes, the PIPPS instrument was translated into Spanish. Establishing equivalence of meaning of items across PIPPS versions was accomplished using the preferred method of back-translation (Guerra & Jagers, 1998). First, a bilingual Head Start Curriculum Specialist with expertise in children's behavior terminology completed translation of the PIPPS from English to Spanish. Next, the Spanish-translated version was reviewed by three parents and three teachers representing at least three different Latin-American countries that comprised a majority of the target sample's background (specifically, Cuba, Nicaragua, and Honduras). Reviewers pointed to specific ambiguities possible due to culture-specific idiom differences. For example, Item 27 involved translation of the English word "disrupts" as "interrumpir" for some and "desbaratar" for others.

Following discussion, a final Spanish version was developed and translated back into English by a second bilingual translator who was not a participant in the initial translation. Both translators who were used in this process had experience within early childhood education settings and were completing master's level studies in special education. The English translation of the instrument was compared to the original PIPPS scale in order to determine if the essential meaning of the items in each scale was preserved throughout the translation process. This step yielded a comparable Spanish translation of the PIPPS scale, with the exception of Item 12, "Tattles." This item presented some difficulty in the translation process because the translated item, "rumorea, dice secretos" was construed by some of the scale reviewers as "a positive and normal thing."

After discussion, it was determined that the item would remain on the measure, and the fitness of the “Tattles” item would be examined empirically through factor analysis. The interpretation of all items by members of the Hispanic culture would be assessed within the selected data analyses.

Measures

Demographic questionnaire

Information regarding each child’s birthday, gender, and ethnicity was obtained from a questionnaire completed by parents.

Peer play interaction

The Penn Interactive Peer Play Scale (PIPPS; Fantuzzo et al., 1995) was developed in collaboration with Head Start teachers and parents to describe peer play interactions of African American children. The teacher version consists of 32 items identifying common play behaviors that facilitate or interfere with peer interactions in the classroom. Multiple factor analytic studies (Coolahan et al., 2000; Fantuzzo et al., 1995; Fantuzzo & Coolahan et al., 1998) yielded three underlying dimensions of children’s play behaviors: Play Interaction, Play Disruption, and Play Disconnection. The Play Interaction factor consists of items reflecting creative, cooperative, and helpful behaviors that facilitate successful peer play interactions. Play Disruption captures children’s aggressive and antisocial play behaviors, while Play Disconnection accounts for withdrawn and avoidant behaviors that impede active participation in play. Internal consistency using Cronbach’s alpha shows strong reliability for these factors ($r = .92, .91, \text{ and } .89$, respectively). Interrater reliability between teachers and assistant teachers was .85, indicating strong agreement between adults in the classroom, despite variation in levels of educational training and also racial background.

Concurrent validity has been established using teacher rating scales, peer nominations, and observations of classroom free play (Fantuzzo & Coolahan et al., 1998). Children who showed high ratings for Play Interaction also received high ratings from teachers on the Social Skills Rating Scale (SSRS; Gresham & Elliot, 1990). Interactive children were accepted by peers and observed as engaged in peer play. Play Interaction also correlated with positive learning styles, as measured by the Preschool Learning Behaviors Scale (PLBS; McDermott, Green, Francis, & Stott, 1996), teacher ratings of adaptable temperament, and observations of emotion regulation and autonomous behavior (Mendez et al., in press). Children rated high on the Play Disconnection factor received low peer ratings as assessed by sociometric procedures, and received teacher ratings of passivity, inattentiveness, and lack of motivation. Children receiving high ratings for Play Disruption received high scores for classroom conduct problems on the Conners’ Teacher Rating Scale (CTRS; Conners, 1990).

Problem behaviors

The Problem Behavior Scale of the preschool version of the Social Skills Rating System for Teachers (SSRS-T; Gresham & Elliot, 1990) was used to assess concurrent validity for the Spanish translation of the PIPPS. This rating scale is a 10-item checklist that yields two subscales measuring externalizing and internalizing behavior problems. The teacher rates each question on a 3-point scale (never, sometimes, very often). Research investigating the use of the SSRS-T with urban, African American Head Start children provided the empirical basis for using

the Problem Behavior scales of the SSRS in the present investigation (Fantuzzo, Manz, & McDermott, 1998). The research indicates that only the Problem Behavior Scales were replicated with this population, that is, the construct validity was identical to the original published scale structure, whereas the Social Skills Scales revealed different constructs from the original. Problem behavior scales demonstrated adequate internal consistency with alpha coefficients of .88 and .77 for externalizing and internalizing behavior problems (Fantuzzo & Manz et al., 1998). Higher order factor analytic analyses conducted on the empirically derived Social Skills and Problem Behavior factors with this population of preschool children revealed only one dimension with social skills and problem behaviors producing positive and negative loadings, respectively. In the absence of validation data on the Social Skills Rating Scale with a Hispanic population, the two Problem Behavior subscales were used to provide an initial test of concurrent validity for the Spanish translation of the PIPPS.

Data Analysis

Factor analysis is a common technique used to investigate whether constructs are generalizable across cultural groups (Guerra & Jagers, 1998; Kerlinger, 1996; Knight & Hill, 1998). Construct validity was first assessed independently with the total sample of 320 children. Data were subject to exploratory common factor analysis, with squared multiple correlations as the initial communality estimates. This procedure allows a factor structure to emerge that is unique to the sample without imposing an existing structure onto the new population of interest (Knight & Hill, 1998). Analyses included orthogonal varimax and oblique promax (where $k = 2-6$) rotational procedures. Multiple criteria were applied to determine the optimal factor structure. Specifically, the ideal structure should (a) satisfy Cartell's (1966) scree test; (b) retain five or more salient items per factor, with loadings $\geq .40$ regarded as salient; (c) yield reasonable internal consistency ($\geq .70$; Kline, 1993); (d) remain invariant across random and important demographic subgroups; and (e) demonstrate psychological sense through parsimonious coverage of data and compatibility with prior research (Wood, Tataryn, & Gorsuch, 1996).

Upon establishing an overall factor structure, the Wrigley-Neuhaus factorial matching method (Guadagnoli & Velicer, 1991) was used to compare the Southeast structure with the original PIPPS factor structure obtained with 523 African American children from the Northeast United States. The procedure yields coefficients of congruence reflecting the comparability of a dimension across factor solutions and is an appropriate analysis for examining newly derived structures for similarities with existing data sets (Guadagnoli & Velicer, 1991). A coefficient of 1.0 would indicate perfect agreement of item variance across structures. Next, intercorrelations were calculated for the Northeast and Southeast samples and compared using a chi-square test for equality of independent correlations as another indicator of construct equivalence across ethnic groups (see Knight & Hill [1998] discussion). Concurrent validity for the constructs of the Spanish PIPPS was then evaluated by determining bivariate correlations between the three PIPPS factors and the two SSRS problem behavior factors. It was hypothesized that the problem behaviors would positively correlate with Play Disruption and Play Disconnection, and that the magnitude of correlation would be higher between Externalizing and Disruption and Internalizing and Disconnection, respectively.

Lastly, a repeated measures analysis of variance (ANOVA) was used to examine gender and ethnic differences within the Southeast sample across the PIPPS dimensions. The specific goal of

this step was to identify interaction effects among gender and ethnicity for the PIPPS factors Play Interaction, Play Disruption, and Play Disconnection. The repeated measures procedure is a robust data analytic strategy with sufficient statistical power and parsimony due to the simultaneous consideration of children's three PIPPS factor scores as the repeated measure.

RESULTS

Table 1 displays the factor solution for the Southeast sample collected within this study. A three-factor promax structure ($k = 2$) met all of the stated criteria for retention. The solution yielded the most parsimonious simple structure as reflected by its maximum hyperplane count. The first factor accounted for 49.7% of common variance, contained items associated with children's play strengths and prosocial behaviors, and was named Play Interaction. The second factor accounted for 25.9% of common variance, comprised aggressive or acting out behaviors, and was called Play Disruption. The third factor accounted for 12.8% of common variance, conveyed passive or withdrawn play behaviors, and was named Play Disconnection. Examination of internal consistency showed strong reliability for each subscale (.87, .88, and .83, respectively). The findings demonstrate that Spanish-translated items yield remarkably similar dimensions of peer play as compared with the English version of the PIPPS teacher rating scale.

TABLE 1. Promax Factor Structure for the Total Southeast Sample with Varimax Prerotation

	Factor 1	Factor 2	Factor 3
Play Interaction (Alpha = .87)			
Comforts others who are hurt or sad	.83	-.04	.04
Encourages others to join play	.74	.02	-.10
Shows creativity in making up play stories and activities	.69	.16	-.11
Helps other children	.65	.02	-.15
Helps settle peer conflicts	.65	.04	.03
Shows positive emotions during play (e.g., smiles, laughs)	.64	-.06	.02
Directs others' actions politely	.63	-.11	-.01
Verbalizes stories during play	.64	.08	-.11
Flexible—can go with the flow	.48	-.19	-.05
Tattles	.43	.34	-.05
Disagrees without fighting	.42	-.17	.16
Play Disruption (Alpha = .88)			
Is physically aggressive	-.09	.80	-.01
Grabs others' things	-.04	.78	.04
Starts fights and arguments	-.01	.77	-.00
Disrupts the play of others	-.12	.74	.07
Destroys others' things	-.03	.65	.09
Verbally assaults others	.03	.65	.03
Is rejected by others	.07	.43	.33
Rejects the play ideas of others	.05	.42	.37
Cries, whines, and shows temper	-.07	.41	.36
Play Disconnection (Alpha = .83)			
Hovers outside play group	.07	.11	.70
Withdraws	.03	.08	.70
Needs help to start playing	-.14	-.13	.68
Confused in play	-.10	.03	.58
Needs teacher's direction	-.27	.03	.50
Refuses to play when invited	-.04	.25	.47
Is ignored by others	-.04	.25	.46
Seems unhappy	-.09	.18	.41
Wanders aimlessly	-.08	.38	.40
Remaining Non-Loading Items			
Does not take turns	.11	.29	.24
Shares toys with other children	-.28	.05	.04
Has difficulty moving from one activity to another	-.15	.02	.29

Note. $N = 320$.

Promax Factor Structure for the Total Southeast Sample with Varimax Prerotation

Indicators of Construct Congruence

To assess similarity empirically, structural congruence coefficients were computed between the Southeast sample and the original validated scale structure obtained for 523 African American children. Table 2 depicts the strong congruence between similar factors (i.e. Play Interaction for each sample), as evidenced by high Wrigley-Neuhaus coefficients (Guadagnoli & Velicer, 1991). Moreover, the differentiation between unlike factors across the two geographically distinct

samples is sufficient. This procedure was also repeated following the deletion of the item “Tattles,” a question that elicited disparate interpretations by Spanish-speaking adults during the translation discussion. Following removal of “Tattles,” the Wrigley-Neuhaus coefficient for the Play Interaction factor increased from .71 to .74, which suggests that this behavior may differ across regional or cultural dimensions. Specifically, this item demonstrated a positive loading on the Play Interaction factor within the Southeast, predominantly Hispanic sample, while “Tattles” contributed to the Play Disruption factor in the Northeast, African American sample.

TABLE 2. Structural Congruence between Two Geographically Distinct Samples

Factor	W-N Coefficient	W-N*
Play Interaction	.71 (.13)	.74(.12)
Play Disruption	.82 (.22)	.81(.20)
Play Disconnection	.83 (.25)	.83(.25)

Note. W-N Coefficient = Wrigley Neuhaus Coefficient of Congruence. Nonparenthetical values indicate coefficients of congruence for each factor and its counterpart. Values in parentheses indicate coefficients for each factor compared with noncounterpart factors.

*Denotes Wrigley Neuhaus calculated after removing the item “Tattles.”

N = 320 for total Southeast sample. *N* = 523 for the African American Northeast sample.

Structural Congruence between Two Geographically Distinct Samples

The patterns of interfactor correlations for the two geographically different samples were also examined for significant differences. Similar structures should demonstrate comparable relations between the constructs produced by this measure. Chi-square tests yielded no significant differences ($p < .20$), which indicates that the correlations do not differ among the regional samples. The patterns of relationships among the PIPPS scales across samples showed negative correlations for Play Interaction as compared with Play Disconnection ($-.55, -.37$, respectively for Northeast and Southeast samples) and for Play Interaction compared with Play Disruption ($-.14, -.18$, respectively). The correlation between Play Disconnection and Play Disruption across samples was also comparable and in the expected direction ($.37, .52$, respectively).

Concurrent Validity

Initial support for concurrent validity was found by examining the relations between the PIPPS factors and the teacher SSRS problem behavior scales. Positive correlations were found between the SSRS Externalizing Scale and Play Disruption ($r = .52$; $p < .01$) and Play Disconnection ($r = .28$; $p < .01$). Similarly, positive correlations were found between the SSRS Internalizing Scale and Play Disconnection ($r = .39$; $p < .01$) and Play Disruption ($r = .31$; $p < .01$). The Play Interaction scale, an indicator of children’s behavioral strengths in play, was unrelated to teacher ratings of problem behaviors in the classroom. The pattern of correlations confirms our hypotheses regarding convergent and divergent validity.

Gender and Ethnic Differences

Table 3 presents the group means and standard deviations for each play construct derived with the Southeast sample. A repeated measures ANOVA procedure revealed a significant between-group effect for ethnicity [$F(1, 316) = 17.14, p < .0001$], with African American children receiving higher group mean scores than Hispanic children for Play Disruption. In addition, a

significant Play \times Gender interaction effect was obtained [$F(2, 632) = 13.75, p < .0001$]. Tukey's post hoc analysis revealed that females received higher group mean scores for Play Interaction ($p < .05$). Conversely, males received higher group mean scores for Play Disruption ($p < .05$). No significant mean differences were found for Play Disconnection. Additionally, no other significant interaction effects were detected.

TABLE 3. Group Means and Standard Deviations for PIPPS Constructs by Child Ethnicity and Gender

Construct	Males	Females	African American	Hispanic
Play Interaction	47.75 (9.55)	51.23 (10.21)	50.42 (10.77)	48.80 (9.64)
Play Disruption	50.90 (9.77)	47.31 (8.68)	52.50 (10.21)	48.18 (8.92)
Play Disconnection	49.33 (9.60)	49.36 (9.64)	50.78 (11.96)	48.80 (8.51)

Note. $N = 320$. Standard scores with $M = 50$ and $SD = 10$ were used in the study. Standard deviations are presented in parentheses.

Group Means and Standard Deviations for PIPPS Constructs by Child Ethnicity and Gender

DISCUSSION

The primary objective of this study was to address the pressing need for culturally and linguistically appropriate measurement for assessing preschool children's social behavior. Factor analysis and factor congruence procedures using the Penn Interactive Peer Play Scale (PIPPS) were employed to determine if validated dimensions of preschool peer play behavior were generalizable to a new sample. Results confirmed that constructs found with low-income African American children in the Northeast also emerged with a sample of low-income, predominantly Hispanic, preschool children in the Southeast region of the United States. Psychometrically sound structures were obtained with an ethnically and linguistically diverse sample in support of the original three PIPPS dimensions—Play Interaction, Play Disruption, and Play Disconnection. The comparability of the structure of the Spanish and English versions of the PIPPS holds promise for the use of this measure in research with Spanish-speaking early childhood educators.

In considering construct validity across ethnic groups, Knight and Hill (1998) discuss evaluation procedures for considering both item and scale equivalence. Within the present study, some information exists regarding the value of the set of PIPPS items with a Spanish-speaking population. For example, coverage of the items across all three PIPPS dimensions was comparable to the original factor structure loadings obtained with African American children. Also, three items did not yield a salient loading within either the African American or the Hispanic factor structures, which indicates that these items are poorly phrased in either language. Lastly, the item "Tattles," which presented some difficulty during the back-translation process, revealed its uniqueness in an expected way within the factor analysis. Specifically, Spanish-speaking teachers indicated that "tattles" or "rumorea, dice secretos," was a behavior accompanied by a positive rather than a negative connotation. In this study, Spanish-speaking teachers viewed this item as an age-appropriate social behavior, which explains the positive loading of this item on the Play Interaction factor. This departure from the English PIPPS factor structure likely signifies a culturally specific belief that influences interpretation of children's social behavior.

While these “culture-specific” interpretations represent a rich area for additional study, an important general finding is that the set of PIPPS items conveys their meaning after translation into Spanish. Furthermore, evidence from the series of factor analysis procedures and factor-matching techniques provides support for the three subscales of the PIPPS. Specifically, Play Interaction captures prosocial play behavior, Play Disconnection captures disengagement from play, and Play Disruption captures problematic or aggressive classroom play. An important question, which cannot be addressed within the present study, is the meaning of individual scores on each subscale across ethnic groups. For example, as discussed by Knight and Hill (1998), do teacher ratings on the English PIPPS and the Spanish PIPPS reflect the same intensity or magnitude of the construct across ethnic groups? Future research that combines observational techniques with PIPPS teacher ratings using equal sample sizes of two or more ethnic groups would inform these questions.

The consideration of assessment procedures conducted with non-English-speaking teachers, parents, and children is becoming an increasingly important area for researchers and practitioners. Unfortunately, prior research has frequently failed to report percentage of participants tested in English versus Spanish or examine a language-of-test-administration effect (Busch-Rossnagel, 1993; Guerra & Jagers, 1998). An important initial step that must be conducted before examining relationships between variables of interest is to ensure that available measures for non-English speakers are valid for use with these populations. Other researchers have demonstrated success with utilizing factor analytical procedures to develop comparable assessment instruments measuring intended constructs with Spanish-speaking populations (Diaz-Guerrero & Diaz-Loving, 1990). Yet, an important caution, in particular for work with linguistic minority populations, where the meaning of language may vary according to country of origin or region of the country, is to assess for these subtle linguistic differences (Guerra & Jagers, 1998).

Evidence for concurrent validity was also found in the patterns of correlation between the PIPPS and the SSRS. Specifically, relations between the SSRS problem behavior scales and the disruptive and disconnected dimensions of the PIPPS were supportive of convergent validity. Further, divergent validity was established as the interactive peer play dimension of the PIPPS did not correlate with teacher reports of problem behaviors. This initial evidence is limited, however, as problem behavior ratings are insufficient to validate complex dimensions of peer play interaction. Additional rating scales that focus on children’s interpersonal skill and social competence would enhance our capacity to conduct research and intervention with preschool children (Fantuzzo et al., 1997). An important next step in the validation process of the Spanish PIPPS is to employ observational measures of children’s social competence and play behavior to investigate convergent and divergent validity. Other research with Hispanic and African American preschool children in shared classroom environments has shown concordance between observer-rated and teacher-rated peer play (Milfort & Greenfield, 2001). Replication of this study utilizing the Spanish translation of the PIPPS would yield additional validity data and also extend our knowledge regarding play styles of diverse groups of minority children attending preschool.

The present study enhances the ability of school psychologists to assess preschool social competence by developing a Spanish translation that demonstrates reliability and validity with a

low-income population from a major Southeast region. Yet, this population is not representative of all Spanish-speaking communities in this country and further testing and modification of this Spanish translation is warranted. However, the replication of the construct and concurrent validity of the PIPPS with a Spanish-speaking sample does extend the generalizability of the measure beyond the African American population addressed by Fantuzzo et al. (1995) and subsequent authors (Coolahan et al., 2000; Fantuzzo & Coolahan et al., 1998). As this initial study sets forth reliable play constructs, subsequent work can investigate precisely the relations between the PIPPS and other published rating scales, classroom observations of children's play, and cultural and contextual influences on Hispanic children's play behavior. The role of language, in particular variation in the dominant language of the classroom, may be an important variable in explaining Spanish-speaking children's social competence with peers.

Other limitations of this study will need to be remedied in order to build a comprehensive literature regarding early education and social development of Hispanic children. First, the design of this study did not permit us to disentangle language issues among the children and the teacher raters. The majority of children who spoke Spanish were also rated with the Spanish version of the PIPPS. As stated earlier, an important next step for the validation of this translation will be to obtain and compare ratings collected from Spanish-speaking educators for both English- and Spanish-speaking children. As the English version was validated with observational data, we recommend the use of play observations of Spanish-speaking children as an essential criterion for this translated measure.

The existence of rating scales for bilingual adult raters can lead to interesting studies examining bilingual children from different racial or ethnic groups (i.e., African American bilingual children and Cuban American bilingual children). Within large urban centers, children reside within a diverse constellation of cultural and familial backgrounds often united by both Spanish and English language exposure. These naturally occurring situations, oftentimes frequent in urban Head Start programs, provide a unique opportunity for researchers to utilize more sophisticated designs to answer complicated research questions regarding the interplay between culture, development, and experience (Takanishi & DeLeon, 1994). Adding family's level of acculturation as a contextual variable will also inform the development of minority children's peer play and social competence. We expect issues of cultural identification and attitudes to ascend to the forefront as we seek to improve our research capacity by developing a parent PIPPS for use with Spanish-speaking families.

A secondary goal of the present study was to investigate variability within an urban, Spanish-speaking community according to ethnicity and gender. The factor structure developed in this study for urban preschoolers allowed for calculation of standardized scores. Results showed that teachers rated African American children as more disruptive than Hispanic children. The finding is consistent with prior adjustment data obtained with nationally representative samples showing that African American children evidence greater rates of externalizing behavioral difficulties than other ethnic or racial groups (McDermott & Spencer, 1997). Teachers and observers in Head Start classrooms also rated African American children as more disruptive than Hispanic children (Milfort & Greenfield, 2001). No differences were found for Play Disconnection, regardless of ethnicity or gender. Additional data are needed to determine if the PIPPS is sensitive to magnitude of withdrawn behavior within play across ethnic groups. Some national

data have suggested that children from Hispanic backgrounds manifest greater prevalence of diffident or internalizing behavior disorders (McDermott & Spencer, 1997).

Gender analyses revealed that boys evidenced significantly more difficulties with disruption of play while girls demonstrated greater prosocial tendencies during classroom play. These findings are consistent with prior investigations of gender differences among African American preschool children using the PIPPS (Coolahan et al., 2000; Mendez et al., 2002). Moreover, research conducted with other scales has found relationships between male gender and problem behaviors in preschool children (Hart, DeWolf, Wosniak, & Burts, 1992; Pianta & Caldwell, 1990). Results suggest that greater emphasis should be placed on the assessment of specific difficulties that young males exhibit during efforts to connect and sustain peer play in the classroom (Mendez et al., 2002).

The contribution of this research is perhaps one step toward sustained efforts in the development of appropriate measures for use with Spanish-speaking families and Spanish-speaking teachers. It is critical that we stimulate progress in this area of school-based assessment given the projected population increase in the numbers of children from Hispanic families over the next few decades. Additionally, Hispanic children currently comprise the second largest group of ethnic minority children enrolled in Head Start, the largest federally funded program promoting school readiness for low-income children. The importance of early identification of children's social capacities and special needs is apparent in light of research documenting relations between emotional difficulties in preschool children and exposure to multiple risk factors associated with socially toxic urban environments (Campbell, 1997; Garbarino, 1995; Lavigne et al., 1996).

In sum, the major goal of this research was to add to the number of culturally appropriate assessment tools available for preschool children. This instrument can assist school psychologists in the process of identification of children at risk for developing emotional and behavioral problems. The PIPPS can also help promote mental health by helping psychologists and teachers identify socially competent children who can model appropriate peer interactions for classmates (see Play Buddy Project as described in Fantuzzo et al., 1997). The parent version of the PIPPS (Fantuzzo, Mendez, & Tighe, 1998) should also be translated into Spanish and tested using comparable methods. Finally, the predictive validity of the PIPPS should be established for Spanish-speaking children, as research conducted with African American children has linked peer play competence during preschool with academic success in first grade (Hampton, 2000). The overarching objective is to continue to investigate the validity of instruments and the generalizability of research findings across geographically and ethnically diverse populations of children.

REFERENCES

- Bredenkamp, S., & Copple, C. (Eds.). (1997). *Developmentally appropriate practice in early childhood programs*. Washington, DC: NAEYC.
- Brooks-Gunn, J., Duncan, G. J., & Aber, J. L. (Eds.). (1997). *Neighborhood poverty: Context and consequences for children*. New York: Sage.
- Busch-Rossnagel, N. A. (1993). Commonalities between test validity and external validity in basic research on Hispanics. In A. Anastasi (Ed.), *Psychological testing* (pp. 195–214). New York: Macmillan.

- Campbell, S. B. (1997). Behavior problems in preschool children: Developmental and family issues. *Advances in Clinical Child Psychology*, 19, 1–26.
- Carlton, M. P., & Winsler, A. (1999). School readiness: The need for a paradigm shift. *School Psychology Review*, 28(3), 338–352.
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, I, 245–276.
- Children's Defense Fund. (1998). *The state of America's children yearbook 1998*. Washington, DC: Author.
- Cicchetti, D. (1990). The organization and coherence of socioemotional, cognitive, and representational development: Illustrations through a developmental psychopathology perspective on Down syndrome and child maltreatment. In R. Thompson (Ed.), *Nebraska symposium on motivation*, vol. 36: Socioemotional development (pp. 259–366). Lincoln, NE: University of Nebraska Press.
- Conners, C. K. (1990). *Couriers' rating scales, manual*. Toronto: Multi-Health Systems.
- Connolly, J., & Doyle, A. (1984). Assessment of social competence in preschoolers: Teachers versus peers. *Developmental Psychology*, 17(4), 454–462.
- Coolahan, K. C., Fantuzzo, J., Mendez, J. L., & McDermott, P. A. (2000). Preschool peer interactions and readiness to learn: Relationships between classroom peer play and learning behaviors and conduct. *Journal of Educational Psychology*, 92(3), 458–465.
- Denham, S. A., & Holt, R. W. (1993). Preschoolers' likability as cause or consequence of their social behavior. *Developmental Psychology*, 29, 271–275.
- Diaz-Guerrero, R., & Diaz-Loving, R. (1990). Interpretation in cross-cultural personality assessment. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children: Personality, behavior, and context* (pp. 491–523). New York: Guilford.
- Duncan, G. J., Brooks-Gunn, J., & Klebanov, P. K. (1994). Economic deprivation and early childhood development. *Special Issue: Children and poverty. Child Development*, 65(2), 296–318.
- Fantuzzo, J., Coolahan, K. C., Mendez, J. L., McDermott, P. A., & Sutton-Smith, B. (1998). Contextually-relevant validation of constructs of peer play with African American Head Start children: Penn Interactive Peer Play Scale. *Early Childhood Research Quarterly*, 13(3), 411–431.
- Fantuzzo, J. W., Manz, P. H., & McDermott, P. A. (1998). The teacher social skills rating system: An analysis of its use with African American urban children. *Journal of School Psychology*, 36(2), 199–214.
- Fantuzzo, J., Mendez, J., & Tighe, E. (1998). Parental assessment of peer play: Development and validation of the parent version of the Penn Interactive Peer Play Scale. *Early Childhood Research Quarterly*, 13(4), 655–672.
- Fantuzzo, J., Sutton-Smith, B., Atkins, M., Meyers, R., Stevenson, H., Coolahan, K., Weiss, A., & Manz, P. (1997). Community-based resilient peer treatment of withdrawn maltreated preschool children. *Journal of Consulting and Clinical Psychology*, 64, 1377–1386.
- Fantuzzo, J. W., Sutton-Smith, B., Coolahan, K. C., Manz, P., Canning, S., & Debnam, D. (1995). Assessment of play interaction behaviors in young low-income children: Penn Interactive Peer Play Scale. *Early Childhood Research Quarterly*, 10, 105–120.
- Fisher, E. P. (1992). The impact of play on development: A meta-analysis. *Play & Culture*, 5, 159–181.

- Garbarino, J. (1995). *Raising children in a socially toxic environment*. San Francisco: Jossey-Bass.
- Garcia Coll, C., Lamberty, G., Jenkins, R., McAdoo, H., Crnic, K., Wasik, B., & Vazquez Garcia, H. (1996). An integrative model for the study of developmental competencies in minority children. *Child Development*, 67, 1891–1914.
- Gaskins, S. (1999). Children's daily lives in a Mayan village: A case study of culturally constructed roles and activities. In A. Goncu et al. (Eds.), *Children's engagement in the world: Sociocultural perspectives* (pp. 25–60). New York: Cambridge University Press.
- Gresham, F. M., & Elliott, S. N. (1990). *The social skills rating system*. Circle Pines, MN: American Guidance Services.
- Gresham, F., MacMillan, D., & Bocian, K. (1996). Behavioral earthquakes: Low frequency, salient behavioral events that differentiate students at-risk for behavioral disorders. *Behavioral Disorders*, 21, 277–292.
- Guadagnoli, E., & Velicer, W. (1991). A comparison of pattern matching indices. *Multivariate Behavioral Research*, 23(2), 323–343.
- Guerra, N. G., & Jagers, R. (1998). The importance of culture in the assessment of children and youth. In V. C. McLoyd & L. Steinberg (Eds.), *Studying minority adolescents: Conceptual, methodological, and theoretical issues* (pp. 167–181). Mahwah, NJ: Lawrence Erlbaum.
- Hampton, G. (2000). *The Penn Interactive Peer Play Scale for kindergarten: Building essential linkages in early childhood assessment*. Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia.
- Hart, C. H., DeWolf, M. D., Wosniak, P., & Burts, D. C. (1992). Maternal and paternal disciplinary styles: Relations with preschoolers' playground behavioral orientations and peer status. *Child Development*, 63(4), 879–892.
- Hart, C. H., Olsen, S. F., Robinson, C. C., & Mandlco, B. L. (1997). The development of social and communicative competence in childhood: Review and a model of personal, familial, and extrafamilial processes. *Communication Yearbook*, 20, 305–373.
- Hoge, R., & Colardaci, T. (1989). Teacher-based judgments of academic achievement: A review of the literature. *Review of Educational Research*, 59, 297–313.
- Huston, A. C., McLoyd, V. C., & Garcia Coll, C. (1994). Children and poverty: Issues in contemporary research. *Child Development*, 65, 275–282.
- Kennedy, J. H. (1988). Issues in the identification of socially incompetent children. *School Psychology Review*, 17(2), 276–288.
- Kerlinger, F. N. (1996). *Foundations of behavioral research* (3rd ed.). New York: Holt, Rinehart, & Winston.
- Kline, P. (1993). *The handbook of psychological testing*. New York: Routledge.
- Knight, G. P., & Hill, N. E. (1998). Measurement equivalence in research involving minority adolescents. In V. C. McLoyd & L. Steinberg (Eds.), *Studying minority adolescents: Conceptual, methodological, and theoretical issues* (pp. 183–210). Mahwah, NJ: Lawrence Erlbaum.
- Ladd, G. W., Kochenderfer, B. J., & Coleman, C. C. (1996). Friendship quality as a predictor of young children's early school adjustment. *Child Development*, 67, 1103–1118.
- Ladd, G. W., & Price, J. M. (1987). Predicting children's social and school adjustment following the transition from preschool to kindergarten. *Child Development*, 58, 1168–1189.
- Lavigne, J. V., Gibbons, R. D., Christoffel, K. K., Arend, R., Rosenbaum, D., Binns, H., Dawson, N., Sobel, H., & Isaacs, C. (1996). Prevalence rates and correlates of psychiatric

- disorders among preschool children. *Journal of the American Academy of Child and Adolescent Psychiatry*, 35, 204–214.
- McConaughy, S. (1993). Advances in empirically based assessment of children's behavioral and emotional problems. *School Psychology Review*, 22, 285–307.
- McDermott, P. A., & Spencer, M. B. (1997). Racial and social class prevalence of psychopathology among school-age youth in the United States. *Youth & Society*, 28(4), 387–414.
- McDermott, P. A., Green, L. F., Francis, J. M., & Stott, D. H. (1996). *Preschool Learning Behaviors Scale*. Philadelphia: Edumetric and Clinical Science.
- McLoyd, V. C. (1990). The impact of economic hardship on black families and children: Psychological distress, parenting, and socioemotional development. *Child Development*, 61, 311–346.
- McLoyd, V. C. (1998). Changing demographics in the American population: Implications for research on minority children and adolescents. In V. C. McLoyd & L. Steinberg (Eds.), *Studying minority adolescents: Conceptual, methodological, and theoretical issues* (pp. 3–28). Mahwah, NJ: Lawrence Erlbaum.
- Mendez, J. L., Fantuzzo, J., & Cicchetti, D. (in press). Profiles of social competence in low-income African American preschool children. *Child Development*.
- Mendez, J. L., McDermott, P., & Fantuzzo, J. (2002). Identifying and promoting social competence with African American preschool children: Developmental and contextual considerations. *Psychology in the Schools*, 39(1), 111–123.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50, 741–749.
- Milfort, R., & Greenfield, D. B. (2001). Teacher and observer ratings of Head Start children's social skills. Manuscript submitted for publication.
- National Association for the Education of Young Children; NAEYC. (1991). Guidelines for appropriate curriculum content and assessment in programs serving children ages 3 through 8. *Young Children*, 41(6), 21–38.
- Parker, J. G., & Asher, S. R. (1987). Peer relations and later personal adjustment: Are low-accepted children at risk? *Psychological Bulletin*, 102, 357–389.
- Parker, J. G., Rubin, K. H., Price, J. M., & DeRosier, M. E. (1995). Peer relationships, child development, and adjustment: A developmental psychopathology perspective. In D. Cicchetti & D. J. Cohen et al. (Eds.), *Developmental psychopathology, 2: Risk, disorder, and adaptation* (pp. 96–161). New York: Wiley.
- Phinney, J. S., & Landin, J. (1998). Research paradigms for studying ethnic minority families within and across groups. In V. C. McLoyd & L. Steinberg (Eds.), *Studying minority adolescents: Conceptual, methodological, and theoretical issues* (pp. 3–28). Mahwah, NJ: Lawrence Erlbaum.
- Pianta, R. C., & Caldwell, C. B. (1990). Stability of externalizing symptoms from kindergarten to first grade and factors related to instability. *Development & Psychopathology*, 2(3) 247–258.
- Raver, C. C., & Zigler, E. F. (1997). Social competence: An untapped dimension in evaluating Head Start's success. *Early Childhood Research Quarterly*, 12, 363–385.
- Rimm-Kaufman, S. E., & Pianta, R. C. (1999). Patterns of family-school contact in preschool and kindergarten. *School Psychology Review*, 28(3), 426–438.

- Rogers, M. (1998). Psychoeducational assessment of culturally and linguistically diverse children and youth. In H. B. Vance et al. (Eds.), *Psychological assessment of children: Best practices for school and clinical settings* (2nd ed.). (pp. 355–384). New York: Wiley.
- Swanson, D. P., & Spencer, M. B. (1991). Youth policy, poverty, and African Americans: Implications for resilience. *Education and Urban Society*, 24(1), 148–161.
- Takanishi, R., & DeLeon, P. H. (1994). A Head Start for the 21st century. *American Psychologist*, 49(2), 120–122.
- U.S. Census Bureau. (1999). *School enrollment—social and economic characteristics of students*. Washington, DC: Author.
- U.S. Department of Education. (1992). *Starting school ready to learn. Questions and answers on reaching National Education Goal I: By the year 2000, all children in America will start school ready to learn* (Resource Document). Washington, DC: Author. (ERIC Document Reproduction Service No. ED 355013)
- U.S. Department of Health and Human Services. (1997). *Trends in the well-being of America's children and youth*. Washington, DC: Author.
- Vaughn, B. E., Azria, M. R., Caya, L. R., Newell, W., Krzysik, L., Bost, K. K., & Kazura, K. L. (2000). Friendship and social competence in a sample of preschool children attending Head Start. *Developmental Psychology*, 36(3), 326–338.
- Weinberger, L. A., & Starkey, P. (1994). Pretend play by African American children in Head Start. *Early Childhood Research Quarterly*, 9, 327–343.
- Wentzel, K. R. (1999). Social-motivational processes and interpersonal relationships: Implications for understanding motivation at school. *Journal of Educational Psychology*, 91(1), 76–97.
- Wood, J. M., Tataryn, D. J., & Gorsuch, R. L. (1996). Effects of under- and overextraction on principal axis factor analysis with varimax rotation. *Psychological Methods*, (1), 354–365.