FACTOR STRUCTURE OF CHILD BEHAVIOR SCALE SCORES IN PERUVIAN PRESCHOOLERS

ERIN L. MEYER AND BARBARA A. SCHAEFER

The Pennsylvania State University

CESAR MERINO SOTO

San Martín de Porres University and the Scientific University of the South

CRYSTAL S. SIMMONS, REBECCA ANGUIANO, JEREMY BRETT, ALEA HOLMAN,
JUSTIN F. MARTIN, AND HEIDI K. HATA

University of California, Berkeley

KIMBERLY J. ROBERTS

Columbia Missouri Public School District, University of California, Berkeley

ZENA R. MELLO

University of Colorado, Colorado Springs

FRANK C. WORRELL

University of California, Berkeley

Behavior rating scales aid in the identification of problem behaviors, as well as the development of interventions to reduce such behavior. Although scores on many behavior rating scales have been validated in the United States, there have been few such studies in other cultural contexts. In this study, the structural validity of scores on a Spanish translation of the six-factor Child Behavior Scale (CBS) was assessed in a sample of 265 Peruvian preschool children who ranged from 2 to 6 years in age. Exploratory factor analysis yielded a four-factor structure, and reliability estimates for scores on the four factors were adequate. The authors suggest replicating the study and examining the utility of CBS scores in predicting future problem behaviors in this population. © 2011 Wiley Periodicals, Inc.

Researchers have argued that behavior problems in preschool children are among the best predictors of adjustment problems later on in life (e.g., Crick, Casas, & Mosher, 1997; Feil & Walker, 1995; Lahey, Loeber, Burke, & Applegate, 2005; Lerner, Inui, Trupin, & Douglas, 1985; Visser, Van Der Ende, Koot, & Verhulst, 2003). For example, Lerner et al. (1985) explored the relationship between problem behaviors in preschool and psychiatric disorders later in life using follow-up data on 88 children who were originally assessed when they were between the ages of 3 and 5 years. These researchers found that children who had higher ratings of behavior problems had an increased risk of developing psychiatric disorders within the subsequent 11 years. These findings support the need for early identification of children with problem behaviors so that interventions can be implemented to prevent later maladjustment (Kellam et al., 2008; Lahey et al., 2005; Walker et al., 1998).

Teachers can be valuable resources in assessing problem behaviors in children because they work with the same children for many hours each day. Consequently, teachers observe a variety of student behaviors in the school setting. Given teachers' experience, teacher rating scales of children's problem behaviors are useful in identifying children at risk for subsequent maladjustment. Many behavior rating instruments have been developed and their scores normed for use in the United States, but only a few have been adapted for use in other cultural contexts. Problem behaviors among

Correspondence to: Erin Meyer, Department of Educational and School Psychology and Special Education, The Pennsylvania State University, 125 CEDAR Building, University Park, PA 16802. E-mail: elb206@psu.edu

children occur in many cultures, and behavior rating instruments with psychometrically sound scores can be beneficial in identifying these behaviors. The purpose of this study was to examine scores on a Spanish translation of the Child Behavior Scale (CBS; Ladd & Profilet, 1996) completed by preschool teachers on a group of Peruvian preschoolers.

USING U.S. INSTRUMENTS IN OTHER CULTURAL CONTEXTS

Adapting assessments for use in other cultures includes considering differences in both values and language, and adapted scales must be examined to verify whether they measure the constructs they were initially developed to measure (Geisinger, 1994). The Child Behavior Checklist (CBCL; Achenbach, 1992) is a behavior rating measure that was initially developed in the United States, but has been adapted for use in several other cultures. Although similar behavior constructs of the CBCL were evident in some cultures, different behavioral constructs emerged in other countries. For example, although the eight-factor structure of the CBCL for ages 6 through 18 (Achenbach & Rescorla, 2001) found in the United States was replicated in a Turkish sample using a Turkish translation of the instrument (Dumenci, Erol, Achenbach, & Simsek, 2004), only six factors were supported in an Australian sample using the original CBCL (Heubeck, 2000). In another cross-cultural examination of the CBCL, Weisz, Weiss, Suwanlert, and Chaiyasit (2003) obtained an eight-factor solution in a Thai sample, but only two of the eight factors demonstrated substantial agreement (kappa $\geq$ .61) with the United States sample. Differences in constructs across countries can be due to a variety of factors, including differences in culture, language, interpretation of behavior, and age in which behavior problems are salient in the cultural context.

Similar concerns have emerged with Spanish adaptations of behavior rating instruments. In 2002, Méndez, Hidalgo and Inglés assessed the psychometric properties of scores on the Matson Evaluation of Social Skills with Youngsters (MESSY) using a sample of 634 adolescents from Spain. Previous studies of MESSY scores (e.g., Matson, Rotatori, & Helsel, 1983; Spence & Liddle, 1990) had yielded six- and seven-factor solutions. However, Méndez et al.'s analysis yielded only four factors in the Spanish sample.

Also in 2002, Carney and Merrell translated the Preschool and Kindergarten Behavior Scales (PKBS; Merrell, 1994) into Spanish. Previous research (Merrell, 1994) had provided support for Social Skills and Problem Behavior factors, with internal consistency coefficients at or above .90. Carney and Merrell had parents and teachers rate the problem behaviors and social skills of 45 children from 3 to 6 years of age using the English and Spanish versions of the PKBS. The correlation between the English and Spanish versions was .93 for the Social Skills subscale scores and .94 for the Problem Behavior subscale scores. The internal consistency coefficients, measured by Cronbach's alpha, were .93 and .96 for the Spanish version's Social Skills and Problem Behaviors scores, respectively. Carney and Merrell concluded that the validity evidence in support of scores in the Spanish version of the PKBS was strong; however, the sample was too small to allow for structural validity analyses.

THE CHILD BEHAVIOR SCALE

There is a particularly great need for instruments in countries with a primary language other than English and limited resources for research and instrument development. Although there are many rating scales with robust psychometric properties that can be translated, most of these are commercially available instruments and there is a cost for each protocol. Thus, translations of *public domain* behavior rating scales that yield psychometrically sound scores in the United States potentially can allow other cultures to benefit from the information they provide. Public domain instruments do not require paying for individual protocols when conducting assessments. Currently,

there are no rating scales with representative norms for children of school age in Peru and no scales assessing problem behaviors in preschoolers (Merino, Livia, & Díaz, 2006). There is one rating scale assessing learning behaviors in Peruvian preschoolers, with limited validity evidence (Hahn, Schaefer, Merino Soto, & Worrell, 2009). However, there are no instruments available for assessing problem behaviors in Peru.

The behavior rating measure used in this study is the CBS (Ladd & Profilet, 1996). Ladd and Profilet (1996) were interested in developing a public domain instrument "for assessing young children's behaviors with peers at school, using teachers as informants" (p. 1010). The original CBS, consisting of 59 items, was initially intended to tap five constructs (Ladd & Profilet, 1996, p. 1010): "aggressive behavior with peers, prosocial behavior with peers, and three types of withdrawn behaviors (i.e., asocial with peers, excluded by peers, and anxious-fearful behavior)." The scale development study included two independent cohorts of kindergarten children with assessments in both fall and spring. Principal components analyses (Cohort 1 Fall) provided support for a six-factor structure (70.3% variance accounted for) consisting of 35 items, and this structure was confirmed with the fall data from Cohort 2 and the spring data from both cohorts. Items retained on each of the factors had coefficients greater than or equal to .55, minimal to no complex loadings, and content that related to one of the hypothesized behavioral constructs.

The six factors were labeled as follows: Aggressive with Peers (7 items, $.89 \leq \alpha \leq .92$), Prosocial with Peers (7 items, $.91 \leq \alpha \leq .92$), Asocial with Peers (6 items, $.87 \leq \alpha \leq .89$), Excluded by Peers (7 items, $.93 \leq \alpha \leq .96$), Anxious-Fearful (4 items, $.77 \leq \alpha \leq .79$), and Hyperactive-Distractible (4 items, $.88 \leq \alpha \leq .93$). Scores on all but one of the subscales were above the .80 value recommended for screening measures (Salvia & Ysseldyke, 2004). Ladd and Profilet (1996) also reported other evidence in support of the CBS subscale scores, including stability coefficients from fall to spring (.54 to .71), subscale intercorrelations consistent with theory, and convergent validity coefficients with classroom observation data, peer ratings, and subscale scores on the Child Behavior Profile-Teacher Report Form (Achenbach, 1991).

A longitudinal study was also completed on the CBS with 396 youth aged 6 to 13 years in Grades 1 through 8 (Ladd, Herald-Brown, & Andrews, 2009). Confirmatory factor analyses were conducted for each age cohort using the 35 items accepted by Ladd and Profilet (1996). In addition to the hypothesized six-factor model, two five-factor models were also examined. The six-factor model provided the best fit to the data each year. In addition, each subscale formed a single factor that was invariant over time. Estimates of internal consistency and stability suggested that the externalizing behavior scales (e.g., Aggressive with Peers, Hyperactive-Distractible) were more consistent and stable across grades than the scales assessing internalizing behaviors (e.g., Anxious-Fearful). The use of additional measures (i.e., peer nomination, number of mutual friendships, CBCL, and Teacher's Report Form), as well as additional informants (i.e., parents and peers), provided further validity support for CBS subscale scores.

THE CURRENT STUDY

The purpose of this study was to assess the reliability and structural validity of scores on a Spanish translation of the CBS (Ladd & Profilet, 1996) in a sample of preschoolers in Peru. If the validity of CBS scores were to be supported in the Peruvian sample, the instrument might be useful in identifying Peruvian children who can benefit from interventions aimed at the reduction of problem behaviors. Three questions were addressed in this study: (1) What factors will CBS scores yield in a sample of Peruvian preschoolers? (2) Will the scores on the Peruvian factors be reliable? and (3) Will the factors in Peru be similar to the factors obtained in the United States?

Setting

The study was conducted in the Republic of Peru, which is located on the western coast of South America between Chile and Ecuador. Approximately 29.5 million people live in Peru, and 29% of the population is 14 years of age or younger (Central Intelligence Agency [CIA], 2009). The ratio of males to females in the population is about one to one. Forty-five percent of the population is Amerindian, 37% is Mestizo (Amerindian and White), 15% is White, and 3% is Black, Japanese, Chinese or another ethnicity. Approximately 44.5% of the population of Peru lives below the poverty level and there is a 9% unemployment rate. Peru has two official languages, Spanish and Quechua, but Aymara and several other Amazonian languages are also spoken (CIA, 2009). In Peru, the ability to speak and write Spanish is essential to receive services from the government; therefore, the Ministry of Education requires that the instruction in the schools be given in Spanish (Library of Congress Federal Research Division, 2009).

Peruvian education is divided into four systems: preschool education (ages 3 to 6), elementary education (ages 6 to 11), secondary education (ages 11 to 16), and postsecondary education (Merino Soto, Díaz Casapía, Ponce, & Morales, 2007). Although the first 12 years of schooling are compulsory (Minsterio de Educación, 2005), the government provides different levels of support across the schools. About 20% of preschoolers, mostly from low-income communities, are served in preschools that receive minimal support from the government (Cueto & Diaz, 1999), and it is from these schools that this study's sample was drawn.

Participants

Participants consisted of a convenience sample of 265 students, ranging in age from 2 to 6 years ($M = 3.8$, $SD = .85$) attending 10 nongovernmental preschools in Lima, Peru. Fifty-two percent of the students were female. The students were from low-income households, and the parents had completed elementary or some secondary education. Many of the parents were immigrants to Lima from rural areas of Peru and were typically engaged in housework (mothers) or were small-scale merchants (fathers).

Students were rated on the CBS by 23 female teachers, who were between 20 and 30 years old ($M = 25$) and had 1 to 8 years of teaching experience. Each teacher rated approximately 12 students, on average. Two teachers had degrees, several had 1-year diplomas from a technical institute, and the rest had completed secondary school. The majority of teachers lived in the same communities as the students. Each preschool had a coordinator, and a school psychologist was assigned to the preschools.

Measure

The CBS (Ladd & Profilet, 1996) is a teacher-completed behavior rating scale. Types of behavior assessed by the CBS fall into six main categories: Aggressive with Peers, Prosocial with Peers, Asocial with Peers, Excluded by Peers, Anxious-Fearful, and Hyperactive-Distractible. Teachers rate student behaviors on a 3-point Likert scale (1 = *Doesn't apply*, 2 = *Applies sometimes*, 3 = *Certainly applies*). As previously indicated, CBS scores are structurally valid and yield moderate to high internal consistency estimates in preschool-age and older samples in the United States (Ladd et al., 2009; Ladd & Profilet, 1996). Although the final structure accepted in the United States consists of 35 items, given the cultural differences between Peru and the United States and the exploratory nature of this study, the original 59 items were included in the analyses.

The CBS was translated in two phases. First, items were translated into Spanish by a school psychologist based in Lima. Then, a second translator who was not a psychologist looked at the

translation based on the semantics in relation to the original meaning in English. Inconsistencies in translation were reconciled by the two translators in favor of the semantics of the item in the local Spanish idiom rather than a literal translation of the English words. The Spanish adaptation of the CBS is available on request.

Procedure

The 10 preschools used in this study had Coordinators who were willing to allow their school psychologist to collect the data as part of a training workshop on child development for the teachers. Teachers who completed the CBS participated in a $2^1/_2$ hour workshop conducted by one of the authors, a school psychologist in Peru. In the workshop, teachers were briefed about the utility of rating scales and the advantages of assessments based on teacher observations. The teachers were introduced to items on learning and social behavior that were similar to those on the CBS, but they were only given the CBS itself when asked to rate the participants. At that time, the CBS was described, and teachers were told how to complete the form. Teachers were given a form for each student in their classroom and were allowed 4 weeks to complete the forms. They were asked to rate the behavior of their students as accurately as possible and told that they would be able to use the information from the ratings to speak to their students' parents about their children's development and needs. Over the course of the 4 weeks, the school psychologist was sometimes present at each site and available to teachers to answer questions. Teachers were not compensated for completing the rating scales, but after they had turned in the ratings on their students, they received training in interpreting the rating scale and communicating this knowledge to parents.

RESULTS

Several criteria were used to determine whether factor analysis would be appropriate. First, descriptive statistics were computed on the data obtained for each CBS item. Univariate descriptive statistics revealed that seven items had skewed distributions (> 2), and one item had a kurtotic distribution (>7). Multivariate normality was investigated through the Mahalanobis distance test. Eighteen of the 265 cases had significant Mahalanobis distance values ($d = 98.34$, $df = 59$, $p < .001$). However, given the nature of the response format, the expectation of non-normal distributions for problem behaviors, and the assumption that these cases were part of the intended population, these items were retained for further analysis. Additionally, an investigation of scatterplots indicated that the assumption of linearity was not violated. Correlations greater than .30 (Tabachnick & Fidell, 2001), a significant value on Bartlett's (Bartlett, 1954) test of sphericity ($\chi^2 = 9193.21$, $df = 1711$, $p < .001$), and a Kaiser-Meyer Olkin measure of sampling adequacy (Kaiser, 1974) value of .91 indicated that the correlation matrix was factorable.

An exploratory factor analysis using principal axis factor extraction was used to investigate the latent structure of CBS scores. Several decision rules were used to determine the number of factors to retain, including Cattell's (1966) scree test, minimal average partials (Velicer, 1976), and parallel analysis (Horn, 1965). Parallel analysis and the scree test suggested that four factors should be retained. However, minimal average partials and results from previous studies (Ladd et al., 2009; Ladd & Profilet, 1996) suggested a six-factor solution. Thus four-, five-, and six-factor solutions were examined.

Ladd and Profilet (1996) used a correlated factor model in the initial study of CBS scores, so this model was also used in this study to provide a basis of comparison. However, both oblique (promax, oblimin) and orthogonal (varimax) rotations were examined. A value of .40 was used as the floor for salient coefficients (Snook & Gorsuch, 1989), and a minimum of three salient coefficients without cross-loadings was required for a factor to be considered meaningful.

MacCallum, Widaman, Zhang, and Hong (1999) suggested using a combination of criteria to determine whether the size of a sample was adequate for factor analysis, including examination of communality estimates (high, wide, or low), overdetermination of factors (a combination of at least three or four variables per factor along with simple structure), and overall sample size. The communalities for the data in this study were wide (ranging from .16 to .68), each factor in the final solution had a sufficient number of items (> 6 items), and the overall sample size was 265. Mundfrom, Shaw and Ke (2005) offered more specific criteria based on a Monte Carlo simulation. Based on Mundfrom et al.'s criteria, a four-factor solution with wide communalities and a 7-to-1 variable-to-factor ratio requires a sample size of at least 160. Thus, based on the criteria examined, the current sample of 265 was adequate.

The six-factor structure using varimax rotation had eight complex items (i.e., salient coefficients on more than one factor) and five items with no salient coefficients on any of the six factors. Four items in the five-factor solution were complex, and three items had no salient coefficients on any of the factors. Finally, in the four-factor solution, four items were complex and seven items did not have salient coefficients on any of the four factors. Thus, none of the orthogonal solutions provided an adequate fit to the data.

A promax rotation ($\kappa = 4$) was used to extract six-, five-, and four-factor oblique solutions. The six-factor promax solution had six complex items. Additionally, five items had no salient coefficients on any of the five factors. Thus, the six-factor structure did not provide an adequate fit. Factor intercorrelations for the five-factor structure ranged from -.07 to .46, and there were only two salient coefficients on the fifth factor. Additionally, the items with salient coefficients on the fifth factor were complex, also having salient coefficients on the first factor. The four-factor solution had no complex items, five items had no salient coefficients on any of the four factors, and factor intercorrelations ranged from .23 to .44. An examination of the factor structures using oblimin rotation ($\delta = 0$) yielded similar results. Thus, the four-factor structure provided the most parsimonious fit to the data, and this structure also had the closest match to the subscales identified by Ladd and Profilet (1996).

Consequently, the 54-item four-factor structure using promax rotation was selected for interpretation. Pattern and structure coefficients for this solution are provided in Table 1. All items on the first factor, labeled Externalizing Behavior, related to overactive, disruptive, and negative interpersonal behaviors. Items on the second factor, labeled Prosocial Behavior, related to positive interpersonal interactions. All salient items on the third factor, labeled Internalizing Behavior, related to negative intrapersonal behaviors. Items on the fourth factor, labeled Socially Isolated, related to either peer avoidance of a child or a child's avoidance of peers. Coefficient alphas were computed using the items with salient coefficients. They ranged from .81 to .95 (see Table 2) and were adequate for screening decisions (Salvia & Ysseldyke, 2004). Factor correlations were in the low to moderate range (see Table 2). Table 3 contains the CBS items with their original assignments by Ladd and Profilet (1996), as well as their factor assignment in the current study.

DISCUSSION

We examined the structural validity and internal consistency of scores on a Spanish translation of the CBS (Ladd & Profilet, 1996) in a sample of Peruvian preschoolers. A theoretically feasible four-factor structure was obtained, with reliability estimates for scores that are acceptable for screening purposes. The four constructs found in the Peruvian sample differed somewhat from the six that have been found with American samples. However, the factors identified in this study are consistent with factors from other behavior rating measures that have been adapted for use in other countries, which suggests that similar constructs may be present across scales and cultural contexts. For example, the Externalizing Behavior factor on the CBS in Peru coincides with the Aggressiveness/Antisocial Behavior factor on the MESSY (Matson et al., 1983), the Externalizing

Table 1

Structure of the Child Behavior Scale (Principal Axis Extraction, Promax Rotation (N = 265)

Item	Pattern (Structure) Coefficients				h^2
	I	II	III	IV	
1. Restless. Runs about or jumps up and down. Doesn't keep still.	**.93** (.71)	.30 (−.06)	.09 (.26)	−.33 (−.04)	0.68
4. Fights with other children	**.87** (.79)	.03 (−.29)	−.12 (.22)	−.08 (.19)	0.63
2. Squirmy, fidgety child	**.87** (.71)	.12 (−.19)	.04 (.25)	−.36 (−.06)	0.62
23. Kicks, bites, or hits other children	**.80** (.77)	−.06 (−.36)	−.18 (.21)	.06 (.30)	0.61
50. Annoys or irritates other children	**.77** (.77)	.06 (−.29)	−.05 (.30)	.12 (.37)	0.61
35. Aggressive child	**.73** (.73)	.04 (−.28)	−.10 (.26)	.17 (.39)	0.56
36. Taunts and teases other children	**.71** (.73)	−.05 (−.35)	−.11 (.26)	.15 (.38)	0.56
52. Disrupts peers' activities	**.69** (.73)	−.05 (−.37)	.08 (.38)	−.04 (.25)	0.54
48. Argues with peers	**.69** (.69)	−.06 (−.35)	.00 (.30)	−.07 (.20)	0.48
10. Is disobedient	**.65** (.73)	−.17 (−.48)	.16 (.45)	−.16 (.16)	0.6
58. Will continue to bother or hurt other children even when they are clearly upset	**.64** (.67)	.16 (−.19)	−.02 (.30)	.31 (.49)	0.54
16. Bullies other children	**.63** (.62)	.09 (−.20)	−.03 (.25)	.13 (.33)	0.41
22. Inconsiderate of others	**.62** (.74)	−.12 (−.43)	−.03 (.36)	.22 (.46)	0.6
14. Tells lies	**.61** (.69)	−.03 (−.35)	.07 (.38)	.10 (.35)	0.5
20. Blames others	**.59** (.69)	−.07 (−.38)	.11 (.42)	.07 (.33)	0.5
59. Bossy toward peers	**.59** (.65)	.03 (−.28)	.03 (.33)	.16 (.37)	0.45
47. Loses temper easily in conflicts with peers	**.58** (.63)	.05 (−.26)	.05 (.33)	.15 (.36)	0.43
38. Threatens other children	**.57** (.68)	.06 (−.28)	.01 (.36)	.37 (.56)	0.58
3. Destroys own or others' belongings	**.56** (.63)	−.06 (−.33)	−.08 (.26)	.20 (.39)	0.43
7. Irritable; quick to "fly off the handle"	**.52** (.64)	−.04 (−.34)	.08 (.38)	.17 (.39)	0.44
5. Not much liked by other children	**.49** (.59)	−.04 (−.29)	−.13 (.23)	.38 (.52)	0.47
29. Tends to react to other children's distress by teasing them or making things worse	**.48** (.61)	.06 (−.25)	.03 (.34)	.39 (.56)	0.5
13. Fussy or over-particular	**.46** (.63)	−.15 (−.44)	.26 (.51)	−.02 (.27)	0.48
40. Kind toward peers	−.08 (−.38)	**.80** (.79)	.10 (−.24)	.01 (−.17)	0.64
46. Cooperative with peers	−.02 (−.39)	**.77** (.81)	−.06 (−.37)	−.02 (−.23)	0.65
56. Offers to help or comfort when other children are upset	.02 (−.35)	**.74** (.76)	−.01 (−.32)	−.11 (−.28)	0.59
26. Helps other children.	.11 (−.23)	**.69** (.67)	−.08 (−.30)	−.00 (−.15)	0.46
42. Listens to classmates	.04 (−.24)	**.68** (.64)	.07 (−.18)	−.04 (−.16)	0.42
49. Friendly toward other children	.12 (−.22)	**.67** (.65)	−.04 (−.27)	−.07 (−.20)	0.43
41. Can be trusted, is dependable	−.17 (−.46)	**.63** (.71)	.02 (−.31)	−.06 (−.26)	0.54
34. Seems concerned when other children are distressed	.02 (−.22)	**.63** (.59)	.05 (−.17)	.05 (−.08)	0.35
53. Shows concern for moral issues (e.g., fairness, welfare of others)	−.07 (−.32)	**.62** (.64)	−.06 (−.29)	.13 (−.06)	0.43
39. Takes turns with play materials	−.22 (−.39)	**.55** (.59)	.10 (−.18)	.06 (−.11)	0.38
6. Is worried. Worries about many things.	.10 (−.01)	**.46** (.34)	.16 (.05)	.05 (.04)	0.16

(*Continued*)

Table 1
Continued

| Item | Pattern (Structure) Coefficients | | | | h^2 |
	I	II	III	IV	
28. Shows a recognition of the feelings of others; is empathic	.08 (−.06)	**.45** (.37)	.04 (−.06)	.10 (.04)	0.16
44. Compromises in conflict with peers	.06 (−.14)	**.40** (.40)	−.09 (−.21)	.03 (−.07)	0.17
12. Tends to be fearful or afraid of new things or new situations	−.04 (.24)	.08 (−.18)	**.71** (.67)	.02 (.22)	0.46
25. Prefers to play alone	−.05 (.22)	.14 (−.12)	**.64** (.61)	.14 (.31)	0.41
31. Likes to be alone	−.24 (.10)	.10 (−.11)	**.64** (.60)	.30 (.40)	0.47
21. Gives up easily	.16 (.42)	−.12 (−.39)	**.59** (.66)	−.15 (.14)	0.49
24. Stares into space	.05 (.36)	.07 (−.23)	**.57** (.65)	.26 (.45)	0.49
37. Often unoccupied	−.17 (.18)	−.14 (−.31)	**.55** (.58)	.15 (.30)	0.38
8. Appears miserable, unhappy, tearful, or distressed	.04 (.31)	.06 (−.20)	**.50** (.57)	.22 (.39)	0.37
17. Inattentive	.30 (.53)	−.18 (−.46)	**.48** (.63)	−.16 (.15)	0.52
57. Withdraws from peer activities	.19 (.34)	.10 (−.16)	**.47** (.51)	−.03 (.18)	0.28
11. Has poor concentration or short attention span	.24 (.45)	−.18 (−.42)	**.45** (.57)	−.16 (.12)	0.42
15. Has speech difficulty	.05 (.28)	−.02 (−.23)	**.43** (.50)	.11 (.27)	0.26
45. Is ignored by peers	.08 (.32)	−.01 (−.21)	.11 (.32)	**.54** (.60)	0.39
33. Peers avoid this child	.23 (.45)	.05 (−.21)	.10 (.36)	**.53** (.64)	0.47
27. Peers refuse to let this child play with them	.29 (.47)	−.07 (−.28)	−.11 (.22)	**.53** (.62)	0.46
32. Keeps peers at a distance	−.17 (.16)	.00 (−.17)	.33 (.43)	**.52** (.57)	0.41
54. Ridiculed by peers	.13 (.36)	.06 (−.19)	.18 (.38)	**.49** (.58)	0.39
30. Not chosen as a playmate by peers	.16 (.34)	.09 (−.14)	.12 (.31)	**.47** (.55)	0.34
55. Avoids peers	−.07 (.20)	.06 (−.14)	.30 (.41)	**.47** (.54)	0.35
9. Has twitches, mannerisms, or tics of the face and body	.10 (.31)	−.10 (−.28)	.23 (.38)	.21 (.34)	.21
18. Doesn't share toys	.37 (.48)	−.18 (−.36)	−.01 (.25)	.09 (.26)	.26
19. Cries easily	.12 (.30)	.04 (−.17)	.34 (.42)	.12 (.27)	.21
43. Excluded from peers' activities	.26 (.39)	−.05 (−.23)	−.01 (.23)	.34 (.44)	.26
51. Solitary child	−.29 (.02)	−.06 (−.15)	.39 (.39)	.32 (.35)	.27

Note. Salient coefficients (>.40) are indicated in bold. Factor I = Externalizing Behavior; Factor II = Prosocial Behavior; Factor III = Internalizing Behavior; Factor IV = Socially Isolated; h^2 = communality.

Table 2
Factor Correlation Matrix for the Four-Factor Solution of the CBS With Promax Rotation (N = 265)

Factor	I	II	III	IV	α
I. Externalizing Behavior	–				.95
II. Prosocial Behavior	−0.44	–			.85
III. Internalizing Behavior	0.43	−0.38	–		.86
IV. Socially Isolated	0.36	−0.23	0.34	–	.81

Note. Factor I = Externalizing Behavior; Factor II = Prosocial Behavior; Factor III = Internalizing Behavior; Factor IV = Socially Isolated.

Table 3
Items From Ladd and Profilet's (1996) Six-Factor Structure and the Final Four-Factor Structure (N = 265)

Ladd and Profilet's Initial Factor Assignments	Four Factors (Current Study)			
	Externalizing Behavior	Prosocial Behavior	Internalizing Behavior	Socially Isolated
Aggressive with Peers	4, 16, 23, 35, 36, 38, 48			
Prosocial with Peers		26, 28, 34, 40, 46, 53, 56		
Asocial with Peers			25, 31, 57	32, 35
Excluded by Peers	5			27, 30, 33, 45, 54
Anxious-Fearful		6	8, 12	
Hyperactive-Distractible	1, 2		11, 17	
Not Assigned to a factor	3, 7, 10, 13, 14, 20, 22, 29, 47, 50, 52, 58, 59	39, 41, 42, 44, 49	15, 21, 24, 37,	

Problem subscale on the PKBS (Merrell, 1994), and the Rule-Breaking Behavior and Aggressive Behavior factor on the CBCL (Achenbach, 1991). Similarly, the Prosocial Behavior factor on the CBS parallels the Social Skills/Assertiveness factor on the MESSY (Matson et al., 1983) and the Social Skills factor on the PKBS (Merrell, 1994).

Factor Differences Across Cultural Contexts

Although these results are promising, several differences from the previous examinations of this scale should be considered. First, although items related to aggression and prosocial behavior loaded on hypothesized factors in the Peruvian sample, items for Asocial with Peers, Excluded by Peers, Anxious-Fearful, and Hyperactive-Distractible behaviors were split across factors. This finding raises a question about whether these latter constructs are manifested or perceived differently in Peru than in the United States. Are the behaviors investigated by the CBS emic (culture-specific) or etic (similar across cultures), or both (Hart et al., 2000)? An emic perspective explains the differing factor structures across U.S. and Peruvian cultures, whereas an etic perspective explains the similarities.

Another possible explanation is that the psychometric properties of the subscales that split apart in this study are not as strong as the other scales. For example, Anxious-Fearful subscale scores were not as reliable or valid as the other subscales in the Ladd et al. (2009) study. A third possibility has to do with familiarity with rating scales. In an examination of two behavior rating scales in Trinidad and Tobago, Worrell, Watkins, and Hall (2006) found fewer factors in that country than in the United States. They postulated that teachers who are not used to completing rating scales might make fewer fine-tuned distinctions about children's behavior than teachers in the United States, where there is a long history of using these scales.

Translation Concerns

The translation of the items in this study was completed by two individuals based in Lima. Geisinger (1994) summarized several of the issues related to the adaptation of an instrument to a different culture. Adaptation refers to changes in reference to wording, content, and culture so that an assessment can be made as equivalent as possible across cultures. To adapt a test for use in a different culture while maintaining the meaning of the original test as much as possible, several steps should be followed. The measure should first be translated into the new language and the content of the items adapted as needed. Then, experts in both languages and cultures need to review the measure to make sure the adaptation is appropriate. Third, changes should be made to the adaptation

of the test as recommended by the reviewers. Next, the assessment should undergo pilot testing, followed by field testing. The scores on the assessment should also be re-standardized, considering that the results may mean different things in different cultures. Finally, validation research should then be performed on the adapted assessment, and it should be re-normed for the new population.

Some of these guidelines were not followed in the current study (e.g., both translators were from Peru instead of using experts from both countries). However, the CBS does not have national norms in the United States either. The scale is a screening evaluation whose scores have been validated for use on the basis of empirical findings, and this is the method being used in Peru. The fact that the items factored into meaningful constructs in Peru with the same conceptual meaning as the U.S. factors provides some assurance that the translation was successful in capturing the constructs. However, this is an initial study of the instrument in this context, and more research should be conducted before the results are put into widespread use.

Measurement Concerns

One concern about the CBS involves the response format of the CBS items. Several items have more than one prompt (e.g., "Can be trusted, is dependable"), which may make it difficult for the respondent to complete the scale (Comrey, 1988). When several descriptors are grouped together and a child exhibits the identified behaviors to a differing extent, the respondent is forced to give a response that is not entirely true to reality. To avoid this problem, Comrey suggested making items as simple as possible so that the respondent can provide answers more easily. Additionally, a 3-point Likert scale may not provide an accurate representation of the continuous nature of the underlying constructs. A 5-point (Dawis, 1987) or 7-point (Comrey, 1988; Gorsuch, 1997) scale has been recommended to address problems with skew and to ensure that item means are different.

Another measurement concern involves the valence of the items. The Prosocial Behavior factor is composed almost entirely of items that are phrased in a positive manner, whereas all other items in the scale are phrased negatively. Inspection of the item means for the Prosocial Behavior factor also reveals that 10 of 13 items have mean values greater than 2, although no other factors have items with mean values greater than 2. Consequently, the relationship between the items of the Prosocial Behavior factor may be influenced by having a similar distribution as well as similar content (Bernstein & Teng, 1989; Gorsuch, 1997). Prosocial Behavior could be described as a difficulty factor because the factor is produced by items with different distributions from the rest of the items in the scale, although this problem did not manifest itself in the structural analysis of the items.

Limitations

This study had several limitations. The sample was limited to Peru's capital and to individuals from low socioeconomic backgrounds. Thus, it is not clear whether the factors found in this study will generalize to a more diverse population. Additionally, the ideal procedures for translation were not followed. Nonetheless, the results of the study are promising in a context where no other instruments are available.

CONCLUSION

These findings suggest that further research should be conducted on CBS scores in Peru. In particular, research needs to be conducted to determine whether the factor structure obtained in this study can be replicated using other Peruvian samples and whether factor structures differ across age, sex, ethnicity, and other demographic variables. Concurrent validity should also be investigated through comparisons with other information in the system (e.g., parent reports, disciplinary infractions),

Research also needs to be conducted to determine whether elevated problem behavior scores on the CBS are predictive of later behavior difficulties and psychopathology, beginning with a longer term follow-up of participants in this study, if possible. Further evidence for the validity of CBS scores in Peru could open the door for school personnel to engage in early intervention for students with behavioral difficulties. To the extent that the findings are replicated and demonstrate evidence of predictive validity, the development of local norms is a useful next step in the long term. Studies such as this one can help to determine whether factor structures truly differ between the United States and Peru, and can also provide valuable opportunities for cross-cultural collaboration among researchers interested in working with children.

References

Achenbach, T. M. (1991). Manual for the Child Behavior Checklist/4-18 and 1991 profile. Burlington: University of Vermont, Department of Psychiatry.

Achenbach, T. M. (1992). Manual for the Child Behavior Checklist/2-3 and 1992 profile. Burlington: University of Vermont, Department of Psychiatry.

Achenbach, T. M., & Rescorla, L. A. (2001). Manual for the ASEBA school-age forms & profiles. Burlington: University of Vermont, Research Center for Children, Youth and Families.

Bartlett, M. S. (1954). A further note on the multiplying factors for various χ^2 approximations in factor analysis. Journal of the Royal Statistical Society, 16, 296–298.

Bernstein, I. H., & Teng, G. (1989). Factoring items and factoring scales are different: Spurious evidence for multidimensionality due to item categorization. Psychological Bulletin, 105, 467–477. doi:10.1037/0033-2909.105.3.467

Carney, A. G., & Merrell, K. W. (2002). Reliability and comparability of a Spanish-language form of the Preschool and Kindergarten Behavior Scales. Psychology in the Schools, 39, 367–373. doi:10.1002/pits.10033

Cattell, R. B. (1966). The scree test for the number of factors. Multivariate Behavioral Research, 1, 245–276. doi:10.1207/s15327906mbr0102_10

Central Intelligence Agency. (2009). The world factbook: Peru. Retrieved February 21, 2010, from https://www.cia.gov/library/publications/the-world-factbook/geos/pe.html

Comrey, A. L. (1988). Factor-analytic methods of scale development in personality and clinical psychology. Journal of Consulting and Clinical Psychology, 56, 754–761. doi:10.1037/0022-006X.56.5.754

Crick, N. R., Casas, J. F., & Mosher, M. (1997). Relational and overt aggression in preschool. Developmental Psychology, 33, 579–588. doi:10.1037/0012-1649.33.4.579

Cueto, S., & Diaz, J. (1999) Impacto de la educacion preescolar en el rendimiento en primer grado de primaria en escuelas públicas urbanas de Lima [Impact of preschool education on achievement in first grade in urban public schools in Lima]. Revista de Psicologia, 18, 75–91.

Dawis, R. V. (1987). Scale construction. Journal of Counseling Psychology, 34, 481–489. doi:10.1037/0022-0167.34.4.481

Dumenci, L., Erol, N., Achenbach, T. M., & Simsek, Z. (2004). Measurement structure of the Turkish translation of the Child Behavior Checklist using confirmatory factor analytic approaches to validation of syndromal constructs. Journal of Abnormal Child Psychology, 32, 335–340. doi:10.1023/B:JACP.0000026146.67290.07

Feil, E. G., & Walker, H. M. (1995). The early screening project for young children with behavior problems. Journal of Emotional and Behavioral Disorders, 3, 194–202. doi:10.1177/106342669500300401

Geisinger, K. F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. Psychological Assessment, 6, 304–312. doi:10.1037/1040-3590.6.4.304

Gorsuch, R. L. (1997). Exploratory factor analysis: Its role in item analysis. Journal of Personality Assessment, 68, 532–560. doi:10.1207/s15327752jpa6803_5

Hahn, K. R., Schaefer, B. A., Merino Soto, C., & Worrell, F. C. (2009). The factor structure of Preschool Learning Behaviors Scale scores in Peruvian children. Canadian Journal of School Psychology, 24, 318–331. doi:10.1177/0829573509349015

Hart, C. H., Yang, C., Nelson, L. J., Robinson, C. C., Olsen, J. A., Nelson, D. A., et al., (2000). Peer acceptance in early childhood and subtypes of socially withdrawn behaviour in China, Russia, and the United States. International Journal of Behavioral Development, 24, 73–81. doi:10.1080/016502500383494

Heubeck, B. G. (2000). Cross-cultural generalizability of CBCL syndromes across three continents: From the USA and Holland to Australia. Journal of Abnormal Child Psychology, 28, 439–450. doi:10.1023/A:1005131605891

Horn, J. (1965). A rationale and test for the number of factors in factor analysis. Psychometrika, 30, 179–185. doi:10.1007/BF02289447

Kaiser, H. F. (1974). An index of factorial simplicity. Psychometrika, 39, 31–36. doi:10.1007/BF02291575

Kellam, S. G., Hendricks, B.C., Poduska, J. M., Ialongo, N. S., Wang, W., Toyinbo, et al. (2008). Effects of a universal classroom behavior management program in first and second grades on young adult behavioral, psychiatric, and social outcomes. Drug and Alcohol Dependence, 95, S5–S28. doi: 10.1016/j.drugalcdep.2008.01.004

Ladd, G. W., Herald-Brown, S. L., & Andrews, R. K. (2009). The Child Behavior Scale (CBS) revisited: A longitudinal evaluation of CBS subscales with children, preadolescents, and adolescents. Psychological Assessment, 21, 325–339. doi:10.1037/a0016205

Ladd, G. W., & Profilet, S. M. (1996). The Child Behavior Scale: A teacher-report measure of young children's aggressive, withdrawn, and prosocial behaviors. Developmental Psychology, 32, 1008–1024. doi:10.1037/0012-1649.32.6.1008

Lahey, B. B., Loeber, R., Burke, J. D., & Applegate, B. (2005). Predicting future antisocial personality disorder in males from a clinical assessment in childhood. Journal of Consulting and Clinical Psychology, 73, 389–399. doi:10.1037/0022-006X.73.3.389

Lerner, J. A., Inui, T. S., Trupin, E. W., & Douglas, E. (1985). Preschool behavior can predict future psychiatric disorders. Journal of the American Academy of Child Psychiatry, 24, 42–48. doi:10.1016/S0002-7138(09)60408-8

Library of Congress Federal Research Division. (2009, May 7). A country study: Peru. Retrieved February 21, 2010, from http://memory.loc.gov/frd/cs/petoc.html

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. Psychological Methods, 4, 84–99. doi:10.1037/1082-989X.4.1.84

Matson, J. L., Rotatori, A. F., & Helsel, W. J. (1983). Development of a rating scale to measure social skills in children: The Matson Evaluation of Social Skills with Youngsters (MESSY). Behaviour Research and Therapy, 21, 335–340. doi:10.1016/0005-7967(83)90001-3

Méndez, F. X., Hidalgo, M. D., & Inglés, C. J. (2002). The Matson Evaluation of Social Skills with Youngsters: Psychometric properties of the Spanish translation in the adolescent population. European Journal of Psychological Assessment, 18, 30–42.

Merino, C., Livia, J., & Díaz, M. (2006). Experiences in the search for effective services for children with behavioral and emotional problems in Peru. Preventing School Failure, 50(2), 43–47.

Merino Soto, C., Díaz Casapía, M., Ponce, L. Z., & Morales, L. B. (2007). School psychology in Peru. In S. R. Jimerson, T. D. Oakland, & P. T. Farrell (Eds.), The handbook of international school psychology (pp. 299–307). Thousand Oaks, CA: Sage Publications.

Merrell, K. W. (1994). Preschool and Kindergarten Behavior Scales. Austin, TX: PRO-ED.

Ministerio de Educación. (2005). Reglamento de Educación Básica Regular [Regulation of Basic Education](D.S. N° 013-2004-ED). Lima: Ministerio de Educacion.

Mundfrom, D. J., Shaw, D. G., & Ke, T. L. (2005). Minimum sample size recommendations for conducting factor analyses. International Journal of Testing, 5, 159–168. doi:10.1207/s15327574ijt0502_4

Salvia, J. & Ysseldyke, J. E. (2004). Assessment: In special and inclusive education (9th ed.). Boston: Houghton Mifflin.

Snook, S. C., & Gorsuch, R. L. (1989). Component analysis versus common factor analysis: A Monte Carlo study. Psychological Bulletin, 106, 148–154. doi:10.1037/0033-2909.106.1.148

Spence, S. H., & Liddle, B. (1990). Self-report measures of social competence for children: An evaluation of Social Skills for Youngsters and the List of Social Situation Problems. Behavioral Assessment, 12, 317–336.

Tabachnick, B. G., & Fidell, L. S. (2001). Using multivariate statistics (4th ed.). New York: Harper Collins College.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. Psychometrika, 41, 321–327. doi:10.1007/BF02293557

Visser, J. H., Van Der Ende, J., Koot, H. M., & Verhulst, F. C. (2003). Predicting change in psychopathology in youth referred to mental health services in childhood or adolescence. Journal of Child Psychology and Psychiatry, 44, 509–519. doi:10.1111/1469-7610.00140

Walker, H. M., Kavanagh, K., Stiller, B., Golly, A., Severson, H., & Feil, E. G. (1998). First step to success: An early intervention approach for preventing school antisocial behavior. Journal of Emotional and Behavioral Disorders, 6, 66–80. doi: 10.1177/106342669800600201

Weisz, J. R., Weiss, B., Suwanlert, S., & Chaiyasit, W. (2003). Syndromal structure of psychopathology in children of Thailand and the United States. Journal of Consulting and Clinical Psychology, 71, 375–385. doi:10.1037/0022-006X.71.2.375

Worrell. F. C., Watkins, M. W., & Hall, T. E. (2006). Behavioural rating scales for elementary schools. Berkeley, CA: Author.