Editor's Note

Clinicians and researchers commonly use instruments such as questionnaires and scales to increase diagnostic accuracy and to measure complex constructs. Many of these instruments have been developed and tested on relatively homogeneous, white, middle-class samples. Since many of the constructs being measured (behavioral problems, self-esteem, locus of control, for example) may be interpreted differently in different population groups (for example racial/ethnic and social class), we really can't be certain that the instruments are accurately measuring what we think they are in these groups. It therefore becomes important to assess the psychometric properties (validity, reliability, subscale structure) of these instruments in diverse patient groups to assure conceptual and measurement equivalency (conceptual equivalency means "is it measuring the same underlying concept," and measurement equivalency means "do the scores mean the same thing"). This paper by Kostanecka and colleagues provides an example of one such assessment.

Lee M. Pachter, DO
Associate Editor

Behavioral Health Screening in Urban Primary Care Settings: Construct Validity of the PSC-17

Anna Kostanecka, MD,* Thomas Power, PhD,*‡ Angela Clarke, PhD,† Marley Watkins, PhD,§ Cheryl L. Hausman, MD,* Nathan J. Blum, MD*‡

ABSTRACT: The Pediatric Symptom Checklist–17 (PSC-17) is a brief form of the Pediatric Symptom Checklist that is designed to screen for behavioral health problems in primary care settings. It has been proposed to have three subscales: externalizing, internalizing, and attention problems. In the context of developing a behavioral health screening program in an inner-city primary care practice, we evaluated the construct validity of the PSC-17. A total of 331 families with children between 4 and 12 years of age who were seen for well-child care during the study were invited to complete the PSC-17 and 320 families (96.5%) did so. A confirmatory factor analysis was performed and the Comparative Fit Index and root mean square error of approximation fit statistics were calculated to determine whether the data fit the proposed three-factor model. We found that although the PSC-17 contained three subscales, several items did not load predominantly on the subscale that they were proposed to measure. Specifically, although the five items on the internalizing subscale loaded only on this subscale, only four of the seven externalizing items loaded exclusively on the externalizing subscale, and only two of the five attention items loaded exclusively on the attention problems subscale. Clinicians using the PSC-17 in urban low-income communities should recognize that the externalizing and attention problems subscales of the PSC-17 may not be valid measures of these dimensions of child behavior in this population.

*(J Dev Behav Pediatr 29:*124–128, 2008) **Index terms:** behavioral health screening, pediatric primary care, Pediatric Symptom Checklist–17, construct, validity.

Primary care is a major venue for the delivery of behavioral health services to children in this country.[1–3] Primary care providers (PCPs) have an important role in identifying children with or at risk of behavioral health problems. Therefore, there is a significant need for behavioral health screening measures that are valid and efficient to use in the context of primary care. The Pediatric Symptom Checklist (PSC) is a 35-item questionnaire that was developed to screen for behavior health problems in primary care settings.[4] The total score on this measure has been shown to be reasonably accurate in identifying children at risk of behavioral health problems.[4,5] A limitation of the PSC is that it is relatively long and may not be efficient to use in many practices.

In an effort to create a briefer screening tool, Gardner and colleagues[6] conducted a factor analysis of the PSC and selected items with the most salient loadings from the factors identified through this analysis (i.e., internalizing, attention, and externalizing problems). The 17-item measure, known as the PSC-17, uses both the total score and

From the Departments of *Pediatrics and †Psychology, The Children's Hospital of Philadelphia; ‡Department of Pediatrics, The University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania; and §Division of Psychology in Education, Arizona State University, Tempe, Arizona.

Received June 2007; accepted January 2008.

Address correspondence to: Nathan Blum, MD, The Children's Hospital of Philadelphia, Room 210 Children's Seashore House Building, 3405 Civic Center Boulevard, Philadelphia, PA 19104; e-mail: blum@email.chop.edu.

subscale scores to identify children in need of further evaluation. One study in a primary care setting demonstrated that the PSC-17 had fair predictive validity for detecting children with ADHD, disruptive behavior disorders, or depression, but was less effective in detecting children with anxiety disorders.[7]

Although use of the PSC-17 is increasing,[7,8] there is limited research on the validity of this brief measure. To our knowledge, the factor structure of the PSC-17 has never been analyzed. This is potentially problematic because it has been shown that items in a brief questionnaire may yield results that differ from those produced by the same items embedded in a longer version of the instrument.[9,10] In other words, items in the PSC-17 may be rated somewhat differently than those in the original PSC because of changes in the context within which the items are embedded (e.g., sequence of items and the overall length of the measure).

In the context of developing screening methods for behavioral health problems in a primary care center serving primarily urban, low-income families, we sought to evaluate the construct validity of the PSC-17 to address questions about whether the items on each subscale specifically assess the dimension of child behavior that they have been purported to measure. In addition, previous studies evaluating the predictive validity of the PSC-17 were conducted in a psychiatric clinic[6] or in primary care centers serving predominantly white children (90% white, 6% black, 4% other),[7] which raised questions about the applicability of this measure for use in urban primary care settings serving primarily African American children.

METHODS

Participants

The data were collected in a large, inner-city primary care practice of 15 pediatricians who serve a predominantly low-income population. Approximately 70% of the families seen in this practice are eligible for state-federal–funded medical assistance. The PSC-17 was administered to caregivers of children between the ages of 4 to 12 years while they were waiting for a well child visit. Most of the caregivers were able to complete the questionnaires independently; assistance was offered when caregivers had difficulty understanding items. During the 40 data collection sessions in which one of us (A.K.) was available to distribute the questionnaire, 331 eligible families were identified and 320 (96.5%) completed the questionnaire. Four families were excluded as they reported they had previously completed the questionnaire, three families were excluded because their child's developmental age was below 4 years and the children had an established diagnosis of autism and/or mental retardation. Two families left the office before data were collected, and one refused to complete the questionnaire. Because we did not collect any identifying information, the study was determined to be exempt by the Institutional Review Board of The Children's Hospital of Philadelphia.

Measures

The Pediatric Symptom Checklist–17 (PSC-17) is a 17-item parent-report measure that is designed to screen for behavioral health problems in children presenting to a primary care practice.[6] It has been proposed to have three subscales: internalizing problems (five items), externalizing problems (seven items), and attention problems (five items). Children are identified as being at risk on this measure by having a score above a specified cutoff on any subscale or on the total measure.[6,7]

Statistical Analysis

Given that the PSC-17 has a proposed factor structure,[6] confirmatory factor analysis (CFA) was employed to empirically test the validity of this factor structure in the study sample.[11] Based on the number of factors, the variable-to-factor ratio, the estimated communalities, and the expected nonnormal distribution of the data, it was estimated that 250 or more participants were required for this analysis.[12] The actual sample of 320 exceeded this estimate and was sufficient to conduct a CFA of this measure.[13]

CFAs using maximum likelihood estimation methods were conducted on covariance matrices with EQS 6.1.[14] The Comparative Fit Index (CFI) and root mean square error of approximation (RMSEA) fit statistics were selected a priori as methods for assessing whether our data support the proposed factor structure. High values of CFI and low values of RMSEA are indicative of good model fit. As recommended by Hu and Bentler,[15] we used a combination rule that required a CFI cutoff value >0.95 and an RMSEA value <0.06 to determine good model fit. Three CFA models were tested: (1) a restricted one-factor model whereby all 17 items loaded onto a single factor, (2) a restricted two-factor model whereby the five original internalizing items formed the first factor and the remaining 12 items loaded onto the second factor, and (3) a restricted three-factor model mirroring the structure reported by Gardner and colleagues.[6]

RESULTS

Parents completed the measure for 320 children, 186 (52.5%) of whom were male, with a mean age (SD) of 7.3 (2.5) years. Eighty-eight percent of caregivers described the child's race as African American, 7% as biracial, and 5% as other or not specified. Questionnaires were completed by mothers (76%), fathers (10%), and other caregivers (14%). Sixty-three percent of the families rated their annual income equal to or less than $40,000/year and 80% of participants qualified for Head Start and/or free lunch at school. English was used as the primary language in 94% of households. Two percent of families reported more than one primary language including English; 2% reported a language other than English as the primary language, but were able to complete the Pediatric Symptom Checklist–17 (PSC-17) without an interpreter, and 2% of families did not indicate a primary language.

For the 320 completed PSC-17 questionnaires, less than 0.3% of the data points were missing and no ques-

Table 1. Robust Maximum Likelihood (ML) χ^2 and Fit Statistics for Pediatric Symptom Checklist–17 Confirmatory Factor Analysis Models (N = 320)

Model	ML χ^2	*df*	CFI	RMSEA
Restricted one factor	583.53	119	0.762	0.097
Restricted two factors	385.16	117	0.869	0.073
Restricted three factors	321.93	116	0.900	0.064
Unrestricted three factors	182.37	88	**0.959**	**0.047**

CFI, Comparative Fit Index; RMSEA, root mean square error of approximation. CFIs >0.95 and RMSEA indices <0.06 indicate good fit.

tionnaire was missing more than three items. Total scores for 22.5% of the questionnaires were at or above 15 points, the recommended cutoff to indicate that children are at risk of behavioral health problems.[6] Using the proposed subscale cutoff scores recommended by Gardner et al,[6] an additional 12.4% screened positive. Of those without an elevated total score, 1.6% had an elevated score on the internalizing subscale, 5.3% had an elevated score on attention subscale, and 6.9% had an elevated score on the externalizing subscale (adds up to more than 12.4% because 1.4% had more than one elevated subscale).

Results of the confirmatory factor analysis (CFA) testing the proposed restricted factor structures are shown in Table 1. The three-factor model fit the data better than the one- or two-factor solutions. However, this restricted three-factor model did not meet the Comparative Fit Index (CFI) and root mean square error of approximation (RMSEA) criteria for a good fit with the data, and it was marred by high interfactor correlations (0.60 – 0.80). Subsequently, we tested a three-factor, unrestricted model.[16] We selected an anchor item for each proposed factor by identifying the item with the highest loading on each factor based on an exploratory factor analysis. This unrestricted approach differs from the restricted factor analysis in that all of the nonanchor items were free to load on any factor (i.e., subscale) and on more than one factor (i.e., subscale). The unrestricted three-factor model clearly met criteria for good fit. Further, the difference in CFI values between the unrestricted model and the restricted three-factor model was 0.059, suggesting that the unrestricted model was meaningfully better.[17]

Table 2 shows the standardized loadings of PSC-17 items from the unrestricted three- factor model. Overall, the findings demonstrated that the PSC-17 contains three factors, but many of the items loaded significantly ($p < .05$) on more than one factor or subscale. The analyses indicated that the internalizing subscale of the PSC-17 is composed of five items, each of which load on this factor and not the other factors. However, only four of the seven externalizing items load exclusively on the externalizing subscale (Table 2, item numbers 4, 8, 14, 16) and only two of the five attention items load exclusively on the attention problems subscale (Table 2 item numbers 1, 17).

DISCUSSION

The results of a series of confirmatory factor analyses (CFAs) indicated that the three-factor model proposed by Gardner and colleagues[6] provides a better description of the data than alternative one- and two-factor models;

Table 2. Standardized Loadings (and 95% Confidence Interval) of Pediatric Symptom Checklist–17 Items from the Unrestricted Three-Factor Model

	Factor			
Item	Internalizing	Externalizing	Attention	R^2
2. Feels sad, unhappy	**0.528** (0.398–0.658)	0.121 (−0.045 to 0.286)	0.031 (−0.135 to 0.196)	.577
6. Feels hopeless	**0.690** (0.568–0.823)	0.054 (−0.111 to 0.222)	0.007 (−0.160 to 0.173)	.373
9. Is down on him- or herself	**0.818** (0.691–0.950)	−0.099 (−0.274 to 0.074)	0.083 (−0.084 to 0.250)	.284
15. Worries a lot	**0.747** (0.643–0.843)	[a]	[a]	.342
11. Seems to be having less fun	**0.543** (0.406–0.675)	0.187 (0.015–0.357)	−0.095 (−0.262 to 0.073)	.321
8. Fights with other children	[a]	**0.739** (0.635–0.845)	[a]	.519
12. Does not listen to rules	0.098 (−0.034 to 0.227)	**0.329** (0.170–0.483)	**0.264** (0.105–0.418)	.497
5. Does not understand other people's feelings	**0.187** (0.049–0.321)	**0.432** (0.266–0.592)	0.034 (−0.132 to 0.200)	.547
14. Teases others	−0.095 (−0.248 to 0.059)	**0.759** (0.563–0.952)	−0.155 (−0.353 to 0.042)	.661
10. Blame others for his or her troubles	**0.376** (0.253–0.496)	**0.324** (0.173–0.473)	0.148 (−0.003 to 0.298)	.489
4. Refuses to share	−0.014 (−0.161 to 0.132)	**0.615** (0.437–0.793)	−0.040 (−0.222 to 0.141)	.366
16. Takes things that do not belong to him/her	0.138 (0.004–0.269)	**0.464** (0.304–0.620)	0.114 (−0.048 to 0.275)	.344
1. Fidgety, unable to sit still	−0.096 (−0.223 to 0.032)	0.149 (−0.011 to 0.310)	**0.703** (0.537–0.869)	.457
3. Daydreams too much	**0.391** (0.257–0.523)	−0.030 (−0.200 to 0.139)	**0.250** (0.084–0.416)	.414
17. Distracted easily	[a]	[a]	**0.824** (0.718–0.930)	.558
7. Has trouble concentrating	**0.242** (0.123–0.363)	0.089 (−0.065 to 0.244)	**0.499** (0.350–0.648)	.385
13. Acts as if driven by a motor	0.064 (−0.060 to 0.190)	**0.254** (0.102–0.409)	**0.456** (0.308–0.609)	.678

Statistically significant parameters are shown in bold type. [a]An anchor item, which was the item with the highest factor loading on an exploratory factor analysis.

however, it fails to meet criteria for good model fit. Sources of misfit were identified by unrestricted factor analyses. Although all the proposed externalizing items loaded significantly on the proposed externalizing problems subscale, three of the items significantly cross-loaded, suggesting that they did not specifically refer to the proposed factor. In fact, "blames others" had a higher loading on the internalizing subscale than the externalizing subscale. The attention problems subscale was particularly problematic. All five of the proposed attention items demonstrated significant loadings, but three of the items cross-loaded. The item "daydreams too much" had a much higher loading on the internalizing subscale than the attention problems subscale. In contrast, the internalizing scale represented a robust factor consisting of five items with high loadings that did not cross-load significantly with the other factors. Overall, this study does not confirm the construct validity of the Pediatric Symptom Checklist–17 (PSC-17), at least for the low-income, urban sample in which the study was conducted.

One of the possible explanations for our failure to confirm the factor structure of the PSC-17 is that these factors were determined based on a factor analysis of the original PSC, which had 35 items. This is potentially problematic in that changing the context in which items are embedded (i.e., extracting 17 items from the original PSC) may change the factor structure and item content of factors.[18] Responses to items may be affected by such issues as their location in the questionnaire and expectations elicited by items preceding or following the item in the questionnaire.[19,20]

A second possible explanation for our failure to confirm the factor structure of the PSC-17 is that our study was conducted with a population that differed substantially from that in which the PSC-17 was developed.[5] It is possible that parents in the present study who were primarily low income and African American, interpreted items in the PSC-17 differently than parents in the predominantly white samples from which the PSC-17 was developed.[6] It is also possible that high level of comorbidities in this population, especially externalizing behaviors overlapping with attention problems, caused those subscales to be less clearly defined. Finally, we investigated families attending a single primary care center, and it is possible that factors unique to this center or this community affected the results of the factor analysis. Evaluation of the factor structure of the PSC-17 in a more diverse sample may help to distinguish these possible explanations for our findings.

This study focused on examining the construct validity of the PSC-17 and did not address issues of predictive validity. However, it is noteworthy that 22.5% of participants in this study had a total score above the cutoff score of 15 points, indicative of the risk of behavioral health problems. Moreover, if the subscales are used, >34% of the cases screened positive. This rate is much higher than that reported in middle- or mixed-income population samples, which have found positive PSC scores in the 12%–14% range among school-age children.[4] However, it is also substantially higher than the rate identified by Murphy et al,[21] who found that 22% of individuals screened positive using the 35-item PSC in an inner-city, low-income population.

The findings of this study strongly indicate the need for additional research to develop a valid and efficient behavioral health screening tool that can be used in inner-city primary care settings. Although this study affirmed the meaningfulness of the three-factor structure proposed by Gardner and colleagues,[6] it also suggests the need to modify many of the items of the PSC-17 so that they provide a purer estimate of the construct that they are designed to assess. This study does not address whether use of the PSC-17 total score alone would be a valid predictor of children in need of behavioral health evaluation, although it has been reported to have predictive validity similar to that of some longer rating scales in one study.[7] Until more research is completed, clinicians using the PSC-17 in urban low-income communities should recognize that the externalizing and attention problems subscales of the PSC-17 may not be valid measures of these dimensions of child behavior in this population.

ACKNOWLEDGMENTS

This study was supported in part by project #33463 the Maternal and Child Health Bureau (Title V, Social Security Act), Health Resources and Services Administration, Department of Health and Human Services. Also, this study was supported by an award from the Chairman of the Department of Pediatrics at The Children's Hospital of Philadelphia. We thank the primary care physicians, their patients, and families for their contribution to clinical research through the Pediatric Research Consortium (PeRC) at The Children's Hospital of Philadelphia.

REFERENCES

1. Bernal P. Hidden morbidity in pediatric primary care. *Pediatr Ann.* 2003;32:413–418.
2. Cooper S, Valleley RJ, Polaha J, et al. Running out of time: physician management of behavioral health concerns in rural pediatric primary care. *Pediatrics.* 2006;118:132–138.
3. Gardner W, Kelleher KJ, Wasserman R, et al. Primary care treatment of pediatric psychosocial problems: a study from pediatric research in office settings and ambulatory sentinel practice network. *Pediatrics.* 2000;106;e44.
4. Jellinek MS, Murphy JM, Little M, et al. Use of the pediatric symptom checklist to screen for psychosocial problems in pediatric primary care. *Arch Pediatr Adolesc Med.* 1999;153:254–260.
5. Jutte DP, Burgos A, Mendoza F, et al. Use of the Pediatrics Symptom Checklist in a low-income, Mexican American population. *Arch Pediatr Adolesc Med.* 2003;157:1169–1176.
6. Gardner W, Murphy M, Childs G, et al. The PSC-17: A brief pediatric symptom checklist with psychosocial problem subscales. *Ambul Child Health.* 1999;5:225–236.
7. Gardner W, Lucas A, Kolko D, et al. Comparison of the PSC-17 and alternative mental health screens in an at-risk primary care sample. *J Am Acad Child Adolesc Psychiatry.* 2007;46:611–618.
8. Borowski IW, Mozayeny S, Ireland M. Brief psychosocial screening at health supervision and acute care visits. *Pediatrics.* 2003;112:129–133.
9. Smith GT, McCarthy DM, Anderson KG. On the sins of short-form development. *Psychol Assess.* 2000;12:102–111.
10. Podsakoff PM, MacKenzie SB, Podsakoff NP, et al. Common method biases in behavioral research: a critical review of the literature and recommended remedies. *J Appl Psychol.* 2003;88:879–903.
11. Byrne BM. *Structural Equation Modeling with EQS: Basic Concepts, Applications, and Programming*. 2nd ed. Mahwah, NJ: Erlbaum; 2006.

12. Mundfrom DJ, Shaw DG, Ke TL. Minimum sample size recommendations for conducting factor analyses. *Int J Test.* 2005; 5:159–168.
13. Tomarken AJ, Waller NG. Structural equation modeling: strengths, limitations, and misconceptions. *Ann Rev Clin Psychol.* 2005;1: 31–65.
14. Bentler PM. *EQS 6 Structural Equations Program Manual.* Encino, CA: Multivariate Software; 2006.
15. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equation Model.* 1999;6:1–55.
16. Hoyle RH, Duvall JL. Determining the number of factors in exploratory and confirmatory factor analysis. In: Kaplan D, ed. *The Sage Handbook of Quantitative Methodology for the Social Sciences.* Thousand Oaks, CA: Sage; 2004:301–315.
17. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct Equation Model.* 2002;9: 233–255.
18. Alterman AI, McDermott PA, Mulvaney FD, et al. Comparison of embedded and isolated administration of the California Psychological Inventory's socialization subscale. *Psychol Assess.* 2003;15:64–70.
19. Knowles ES. Item context effects on personality scales: measuring changes the measure. *J Pers Soc Psychol.* 1988;55:312–320.
20. Schumann H, Ludwig J. The norm of even-handedness in surveys as in life. *Am Sociol Rev.* 1983;48:112–120.
21. Murphy JM, Reede J, Jellinek MS, et al. Screening for psychosocial dysfunction in inner city children: further validation of the Pediatric Symptom Checklist. *J Am Acad Child Adolesc Psychiatry.* 1992;31:221–233.

Literary Quotes

Development of Self-Awareness- III- Richard Hughes

Two of the submissions in this space in 2006 had to do with the psychological phenomenon of children's development of self-awareness. By this term is meant the recognition of ourselves as separate from our surroundings. This is not an aspect of personality usually assessed in pediatrics or psychiatry, although developmental psychologists have given it some attention. Particularly noteworthy is the recent publication in English of the insightful observations of Dr. Dolph Kohnstamm, retired professor of developmental psychology in Leiden, entitled "I am I. Sudden Flashes of Self-Awareness in Childhood" (U.K. Athena Press, 2007).

An acquaintance with this phenomenon makes one more aware of illustrations of its appearance in novels and other non-scientific literature. A recently noted example can be found in Richard Hughes's "A High Wind in Jamaica." Ten-year-old Emily is one of the English children who are captives on a pirate ship.

"And then an event did occur, to Emily, of considerable importance. She suddenly realized who she was . . .."

"She stopped dead, and began looking all over her person which came within the range of eyes. She could not see much, except a fore-shortened view of the front of her frock, and her hands when she lifted them for inspection: but it was enough for her to form a rough idea of the little body she suddenly realized to be hers . . .."

"What agency had so ordered it that of all the people in the world who she might have been, she was this particular one, this Emily: born in such-and-such a year out of all the years in Time, and incased in this particular rather pleasing little casket of flesh?. . .

"Why had all this not occurred to her before?. . ."

I can recall such a sudden flash myself but have not noticed much discussion of the phenomenon in the professional literature. Should this be a milestone in social or cognitive development of concern to child health experts? How would it be assessed?

William B. Carey, MD

REFERENCE

1. Hughes R. *A High Wind in Jamaica.* Falmouth, Cornwall, Triad/Grenada, 1976; page 85–86.