Children's Somatization Inventory: Psychometric Properties of the Revised Form (CSI-24)

Lynn S. Walker,[1] PhD, Joy E. Beck,[1,2] PhD, Judy Garber,[3] PhD, and Warren Lambert,[4] PhD

[1]*Department of Pediatrics, Vanderbilt University School of Medicine,* [2]*Department of Psychology, Vanderbilt University,* [3]*Department of Psychology and Human Development, Vanderbilt University, and* [4]*Vanderbilt Kennedy Center for Evaluation & Program Improvement*

Objective   To conduct a multimethod psychometric evaluation to refine the Children's Somatization Inventory (CSI) and to investigate its dimensionality.   **Method**   The CSI was administered to 876 pediatric patients with chronic abdominal pain at their initial visit to a pediatric gastroenterology clinic. Tools from three psychometric models identified items that most effectively measured the construct of somatization and examined its dimensionality.   **Results**   Eleven statistically weak items were identified and removed, creating a 24-item CSI (CSI-24). The CSI-24 showed good psychometrics according to the three measurement models and correlated .99 with the original CSI. The CSI-24 has one dominant general factor but is not strictly unidimensional.   **Conclusions**   The CSI-24 is a reliable and psychometrically sound refinement of the original CSI. Findings are consistent with the view that somatization has a strong general factor that represents a continuum of symptom reporting, as well as minor components that represent specific symptom clusters in youth with chronic abdominal pain.

Key words   Children's Somatization Inventory; chronic abdominal pain; factor analysis; functional disorders; item response theory; medically unexplained symptoms; somatic symptoms.

Introduction

Although somatic symptoms are common in the general population (Pennebaker, 1982), individuals who frequently experience medically unexplained somatic symptoms may be classified from a psychiatric perspective as having a somatoform disorder (e.g., pain disorder, somatization disorder) and from a biomedical perspective as having a functional somatic symptom or syndrome (e.g., recurrent abdominal pain, fibromyalgia, irritable bowel syndrome) (Campo & Fritsch, 1994). We developed the Children's Somatization Inventory (CSI; Walker, Garber, & Greene, 1991) to assess 35 symptoms (e.g., headache, nausea, heart racing) that often, but not necessarily, occur in the absence of identified disease. The CSI has been used in numerous studies of pediatric patients, including patients with chronic abdominal pain (Dorn et al., 2003; Mulvaney, Lambert, Garber, & Walker, 2006; Robins, Smith, & Proujansky, 2002; Walker, Garber, & Greene, 1993; Walker et al., 1991), headache (Smith,

Martin-Herz, Womack, & McMahon, 1999; White, Alday, & Spirito, 2001), chronic fatigue syndrome (Garralda, Rangel, Levin, Roberts, & Ukoumunne, 1999; Smith, Martin-Herz, Womack, & Marsigan, 2003; van de Putte, Engelbert, Kuis, Kimpen, & Uiterwaal, 2006), factitious illness by proxy (Berg & Jones, 1999), and chronic unexplained pain (Konijnenberg et al., 2005, 2006). The instrument also has been used to assess symptom reporting in community samples of children receiving routine medical care (Rocha & Prkachin, 2007; Tsao & Zeltzer, 2003) and children undergoing routine immunization inoculations (Rocha, Prkachin, Beaumont, Hardy, & Zumbo, 2003).

Our initial development and validation of the CSI was based on a sample of pediatric patients with frequent abdominal pain (Walker et al., 1991). Evidence of the instrument's reliability and validity in nonclinical community samples was subsequently demonstrated both in our work in the United States (Garber, Walker, & Zeman,

1991) and in work by investigators in the Ukraine (Bromet et al., 2000; Litcher et al., 2001) and the Netherlands (Meesters, Muris, Ghys, Reumerman, & Rooijmans, 2003; Muris, Vlaeyen, & Meesters, 2001). All these validation studies noted that some items on the CSI are rarely endorsed and have low item-total correlations. This observation is not surprising given that the CSI includes items from the symptom criteria that define somatization disorder in adults [Diagnostic and statistical manual of mental disorders, third edition, revised (*DSM-III-R*); American Psychiatric Association, 1987], and several of these symptoms (e.g., memory loss, pain with urination) are rarely experienced by children. Thus, the first aim of the current study was to identify items that could be dropped from the CSI to make the instrument more appropriate for youth while maintaining its measurement properties.

Different approaches to scoring the CSI are evident in the literature and reflect different perspectives regarding the dimensionality of somatization. The creation of total scores by summing responses to all 35 CSI items is consistent with a view of somatization as a single dimension of somatic concern. In contrast, creation of subscales by summing responses to items that reflect symptom categories [e.g., gastrointestinal (GI), conversion] defined by the criteria for somatization disorder is consistent with a view of somatization as multidimensional. The empirical literature with adults suggests that the construct of somatization includes a large general factor that represents a general tendency to report somatic symptoms (Liu, Clark, & Eaton, 1997). However, little is known about the dimensionality of somatization in children and adolescents. Exploratory factor analysis (EFA) of the CSI (e.g., Garber et al., 1991; Litcher et al., 2001; Meesters et al., 2003) has yielded inconsistent factors, and no studies have applied confirmatory factor analysis (CFA) to evaluate whether CSI data support a general somatization factor rather than, or in addition to, symptom clusters. Such an investigation has important implications both for the conceptualization of somatization and for the diagnostic criteria for somatization disorder in children and adolescents. Thus, the second aim of the current study was to examine the dimensionality of the CSI.

Our analytic plan used tools from the three main approaches to psychometrics, namely classical test theory (CTT; Cronbach, 1951; Cronbach & Shavelson, 2004), item response theory (IRT) using Rasch modeling (Linacre, 2006b; Rasch, 1960/1980), and EFA and CFA. Because some items on the CSI are rarely endorsed, we expected to identify items with low variance that could be dropped without reducing alpha reliability. We also expected to identify items that contributed little to distinguishing among respondents. Finally, because somatization has been conceptualized both as trait-like somatic distress and as clusters of symptoms associated with different organ systems (Robbins, Kirmayer, & Hemami, 1997), we expected that EFA and CFA would yield evidence of a single large dimension, somatization, and several smaller unique factors defined by items comprising various bodily systems.

Methods
Participants

The CSI data for the current study were collected from several cohorts of pediatric patients that were not included in the initial instrument development sample (Walker et al., 1991) and have been described in detail elsewhere (Lipani & Walker, 2006; Walker, Garber, Smith, Van Slyke, & Claar, 2001; Walker et al., 2006). All participants were recruited during the period from 1993 through 2005 as consecutive new patients referred to the Pediatric Gastroenterology Clinic in Vanderbilt University Medical Center for evaluation of unexplained chronic abdominal pain. Additional eligibility requirements for study participants included being between the ages 8 and 18 years, living with parent(s) or parent figure, speaking English, having no preexisting diagnosis of organic disease (e.g., Crohn's disease, pancreatitis, ulcerative colitis) that would explain abdominal pain, having no other chronic illness (e.g., diabetes), and having no developmental disability (e.g., mental retardation or autism). Approximately 6% of patients screened for eligibility were excluded because of previously diagnosed chronic illness (e.g., diabetes) or developmental disability.

The sample used in the current study comprised 876 children (mean age = 11.66, $SD = 2.47$); 59% were female. The majority of the participants were Caucasian (87.8%), 4.1% were African American, and 3.1% were of other ethnicities (5% were missing this information); these percentages reflect the ethnic composition of patients evaluated in the clinic. Duration of abdominal pain ranged from 1 month to 15 years ($M = 19$ months, $SD = 24$ months). The majority of the children's mothers and fathers (80%) had completed a high school education, with the average educational level being equivalent to some college or technical school.

Procedure

Parental consent and child assent were obtained upon the families' arrival for their clinic visit, prior to the child's

medical evaluation. Parents completed a questionnaire pertaining to child and family demographics. The child protocol was administered in a private room by an interviewer who read the items aloud and asked children to indicate their answers on a printed response sheet; this procedure was used to maximize similarity between administration of the questionnaire at the clinic and at a subsequent telephone follow-up (not reported here) in which items were read to children and they selected responses from a response sheet that had been mailed to them. Families were compensated for their participation. The study was approved by the institutional review board.

Measures

Children's Somatization Inventory, Child Report Form

Children completed the CSI (Walker et al., 1991), a questionnaire that assesses the perceived severity of 35 nonspecific somatic symptoms. The CSI includes items from the symptom criteria for somatization disorder as defined by the *DSM-III-R* (American Psychiatric Association, 1987), items from the Somatization factor of the Hopkins Symptom Checklist (HSCL; Derogatis, Lipman, Rickels, Uhlenhuth, & Covi, 1974), and an additional symptom—constipation—that is common in functional GI disorders. Examples of CSI items include headaches, low energy, dizziness, and chest pain. One symptom derived from the *DSM* criteria, "pain in the genitals," was removed from an earlier version of the CSI because it was not endorsed by study participants. The stem for symptom report on the CSI is the same as that in the HSCL, namely, "How much were you bothered by (*symptom*)?" (Derogatis, Lipman, & Covi, 1973; Derogatis et al., 1974). The standard time period for symptom report on the CSI is 2 weeks. The response format is a 5-point scale ranging from "*not at all*" (0) to "*a whole lot*" (4). Total CSI scores, obtained by summing all item ratings, can range from 0 to 140.

Demographics Questionnaire

Parents completed a 10-item demographic questionnaire regarding the child's ethnicity, family constellation, and caregivers' occupations and educational levels.

Results

In this section, we refer to the original CSI with 35 items as CSI-35 and to the presently revised CSI with 24 items as the CSI-24. The CSI-35 data were 99.9% complete, with 97.6% of the cases having no missing items. With <5% of cases having any missing data, almost any reasonable method of dealing with missing data can be used

(Harrell, 2001). We replaced missing data by single imputation using the expectation maximization (EM) algorithm (Rubin, 1991) to fill in the 2.4% missing item scores based on each patient's nonmissing responses (Little & Rubin, 1987). An advantage of EM imputation is that it preserves the variance, which may be reduced by mean imputation. In this case, the average item mean and standard deviation were the same both before and after imputation ($M = 0.68$, $SD = 0.88$).

The 876 participating patients were split into two groups by an SAS-generated random number to create a learning sample and a cross-validation sample. This split allowed us to freely experiment with the learning sample ($n = 417$), revise the CSI-35, and then examine the revision's psychometric properties with the fresh cases in the cross-validation sample ($n = 459$).

First, we examined the psychometric properties of items from the CSI-35 in the learning sample (Table I). According to CTT, good items should be free of ceilings or floors that limit their variance and produce excessive skew and kurtosis and should have high item-total correlations to give the test high alpha reliability. Thus, the CTT criteria for item evaluation were descriptive statistics (i.e., mean, SD, skew, and kurtosis) and item-total correlations. Rasch (1960/1980) modeling is a one-parameter member of the IRT family (Embretson, 1996; Lord, 1986) and focuses on test construction with practical applications (Bond & Fox, 2001; Edelen & Reeve, 2007). While two- or three-parameter IRT models are available, the Rasch IRT model was preferable because it can be scored by a clinician using a simple total score or a measure score from a table without the use of specialized scoring software. The Rasch model evaluates test items and people on a single latent trait interval scale. Good items show a logistic ogive curve in which the probability of endorsing an item increases with the strength of the latent trait (i.e., somatization) in the person, and the probability goes down as the strength of the trait increases. For example, a child low in somatization might report the occasional headache (an "easy" item), but not deafness, seizures, and blindness ("difficult" items). Items that fit the Rasch model have good ability to distinguish people on the strength of their trait, as shown by acceptable infit and outfit mean squares. Therefore, in addition to CTT criteria, we used Rasch infit and outfit indices as criteria for identifying CSI items with atypical response patterns indicative of poor measurement. "Infit" measures unexpected responses to items with a latent trait level close to the person's symptom level. "Outfit" measures unexpected responses to items with a latent trait level

Table I. Psychometric Properties of 35 CSI Items in the Learning Sample (*n* = 417)

Item[a]	Mean <0.3	SD <0.7	Skew >2.5	Kurtosis >0.6	Item-total <0.4	Infit 0.5–1.5	Outfit 0.5–1.5	Drop
1. Headache	1.35	1.18	0.48	−0.69	0.44	0.98	1.02	
2. Faintness or dizziness	0.68	0.98	1.44	1.43	0.49	0.98	0.86	
3. Pain - heart or chest	0.76	1.01	1.26	0.84	0.41	1.07	1.09	
4. Low energy, slowed down	1.32	1.12	0.54	−0.57	0.55	0.75	0.78	
5. Pain - lower back	0.77	1.1	1.28	0.58	0.4	1.25	1.23	
6. Sore muscles	0.86	1.05	1.12	0.51	0.45	1	0.99	
7. Trouble getting breath	0.48	0.9	2	3.41	0.46	1.14	1.04	
8. Hot or cold spells	1.03	1.17	0.9	−0.21	0.58	0.88	0.81	
9. Numbness or tingling	0.54	0.91	1.83	2.95	0.48	1.02	0.97	
10. Lump in throat	0.41	0.81	2.09	3.9	**0.26**	1.33	**1.63**	•
11. Weakness	0.87	0.98	1.02	0.46	0.6	0.67	0.64	
12. Heavy feelings in arms, legs	0.48	0.89	1.87	2.84	0.58	0.92	0.72	
13. Nausea, upset stomach	2.27	1.2	−0.21	−0.86	0.48	0.84	0.9	
14. Constipation	0.81	1.15	1.32	0.68	0.4	1.3	1.21	
15. Loose BM's, diarrhea	0.86	1.11	1.16	0.41	**0.34**	1.26	1.27	
16. Pain - stomach	2.56	1.15	−0.42	−0.68	0.42	0.87	1	
17. Heart beating too fast	0.42	0.9	2.28	4.51	0.46	1.26	1.07	
18. Difficulty swallowing	0.43	0.86	2.3	5.12	0.46	1.13	1.08	
19. Losing voice	0.34	0.73	2.43	5.93	0.44	1.02	1.04	
20. Deafness (cannot hear)	**0.18**	**0.61**	**3.94**	**15.94**	0.44	1.43	0.89	•
21. Double vision	**0.25**	**0.66**	**3.07**	**9.78**	0.44	1.18	0.84	•
22. Blurred vision	0.39	0.76	2.11	4.17	0.48	0.96	0.87	
23. Blindness	**0.07**	**0.41**	**7.27**	**57.28**	**0.31**	**1.97**	0.86	•
24. Fainting, passing out	**0.07**	**0.35**	**5.76**	**35.8**	**0.23**	1.47	0.94	•
25. Memory loss, amnesia	**0.21**	**0.67**	**3.75**	**14.95**	0.49	1.29	0.89	•
26. Seizures, convulsions	**0.15**	**0.56**	**4.55**	**22.66**	**0.34**	**1.55**	1.04	•
27. Trouble walking	**0.29**	**0.66**	**2.69**	**7.98**	0.48	0.93	0.84	•
28. Paralysis, muscle weakness	**0.26**	**0.62**	**2.79**	**8.95**	0.47	0.9	0.75	•
29. Difficulty urinating	**0.13**	**0.44**	**4.26**	**20.54**	**0.31**	1.19	0.85	•
30. Vomiting, throwing up	0.58	1.08	1.85	2.39	**0.38**	1.5	1.48	
31. Feeling bloated, gassy	1.03	1.18	0.98	−0.04	0.48	1.05	0.99	
32. Food makes you sick	1.39	1.27	0.57	−0.74	0.48	1.04	1.04	
33. Pain - knees, elbows, joints	0.68	1.04	1.45	1.23	0.46	1.13	1.19	
34. Pain - arms, legs	0.59	0.92	1.54	1.72	0.49	0.97	0.91	
35. Pain - urination	**0.19**	**0.57**	**3.42**	**12.36**	**0.36**	1.25	0.96	•

[a]Item wording has been abbreviated for conciseness. The complete questionnaire is available from the first author. Bold and underlined values indicate that the item did not meet the criterion for the psychometric property listed in that column.

Notes. CSI responses range from 0 to 4. The criteria in the table headers are ad hoc rules of thumb to distinguish weaker items from stronger ones.

markedly different from the person's symptom level (Linacre & Wright, 1999). Popular criteria favor fit indices that lie between 0.5 and 1.5 (Linacre, 2006a) or 0.7 and 1.3 (Bond & Fox, 2001).

Response alternatives for the CSI items range from 0 to 4, and many items had means close to the floor (Table I). For example, the item "blindness" (23) had 96% zeros and a mean of 0.07, indicating that the item "blindness" contributed very little information about most patients because it is so rarely endorsed. To visually identify items in Table I with flooring, we underlined means below 0.30. Items with the lowest means generally had problems on multiple criteria. For example, the item "blindness" had an SD less than half that of the higher items, a skew over 7, a remarkable kurtosis of 57, a below-average item-total correlation, and unacceptable Rasch infit. Thus, the multiple criteria from CTT and Rasch gave a consistent message that the item "blindness" does not measure effectively. Moreover, the underlined items in Table I reveal a general pattern of measurement problems for items with means near the floor. Many items close to the floor had weaker psychometric properties as shown by multiple underscores. These floored items require time and effort on the patients' part to complete, but they add little

to the CSI's ability to measure. Therefore, we dropped the 10 lowest scoring items (indicated by ● in Table I), namely those with means <0.30 on a 0–4 scale. These items all had warning flags on other criteria as well, such as item-total correlations or infit. An additional item, ''lump in throat'' (10), showed unacceptable outfit to the Rasch measurement model and a lackluster item-total correlation <0.30, so it too was dropped, leaving a revised CSI with 24 items (i.e., CSI-24). Our goal in dropping the statistically weaker items was to reduce the test length without reducing its reliability. Next we compared the reliability of the CSI-35 and the CSI-24.

The internal consistency reliability of the CSI-35 in the learning sample was good with Cronbach's alpha = .90. The Rasch person separation reliability, an index of measure sensitivity that evaluates how effectively items separate respondents, was .84, indicating that the CSI-35 has a good ability to distinguish between individuals who are higher or lower on somatization. The Rasch items separation reliability was .99, suggesting that the CSI-35 contains a wide range of item difficulties to measure patients who are low or high on somatization. In the learning sample, the reliability of the 24-item CSI was very close to that of the 35-item form, Cronbach's alpha = .88 (down from .90). Rasch person separation reliability of the CSI-24 in the learning sample remained .84, and the Rasch items separation reliability remained .99. These results based on the learning sample suggest that the 24-item version was as reliable as the 35-item version in the learning sample.

Dimensionality of the CSI-24 in the Cross-Validation Sample

Having revised the CSI in the learning sample ($n = 417$), we next evaluated its psychometric properties in the cross-validation sample ($n = 459$). According to substantive theory, the CSI should reflect an underlying trait of somatization but should not be exclusively unidimensional because somatization can be expressed in many distinct bodily systems (e.g., nervous, musculoskeletal, circulatory, GI). A child who has functional GI symptoms will not necessarily have musculoskeletal symptoms. For this reason, we expected the symptoms on the CSI to have more uniqueness than a strictly unidimensional test.

To describe the dimensionality of the CSI, we began with a principal component analysis (PCA). The goal of the PCA was to describe the variance of the sample as parsimoniously as possible. While there has been debate about the merits of PCA and EFA (Costello & Osborne, 2005), PCA is common in the medical literature and, according to

a widely used text, is the better choice for a unique mathematical solution that represents an empirical summary of the data set (Tabachnick & Fidell, 2001).

The scree plot of principal components in Fig. 1 suggests that a single large component explains almost 30% of the total variance, and a second independent component explains 8% of the variance. For reference, a parallel factor analysis of random numbers is shown by the dashed line (O'Connor, 2000). The third and later components can be ignored as they are similar to chance eigenvalues. To understand the second eigenvalue, we examined its PCA loadings. These loadings were relatively low: Only one item (loose bowel movements) had a loading >0.40. Five additional items (constipation, food intolerance, nausea, bloating, and stomach pain) had loadings between .25 and .40. These items represent common symptoms in the pediatric gastroenterology clinic population from which our sample was drawn.

Next, we conducted a CFA of a one-factor measurement model. The CFA, run with MPLUS (Muthen & Muthen, 2003), suggested that the CSI-24 does not fit a single-factor model. The comparative fit index (CFI) was low 0.74 (a good CFI is >0.95; Hu & Bentler, 1999). The root mean squared error of approximation was 0.07 (a good fit is <0.05, Steiger, 2000; or <0.06, Tabachnick & Fidell, 2001). These results are congruent with those obtained with the PCA and suggest that the CSI-24 items have uniqueness in addition to the somatization trait that they all share.
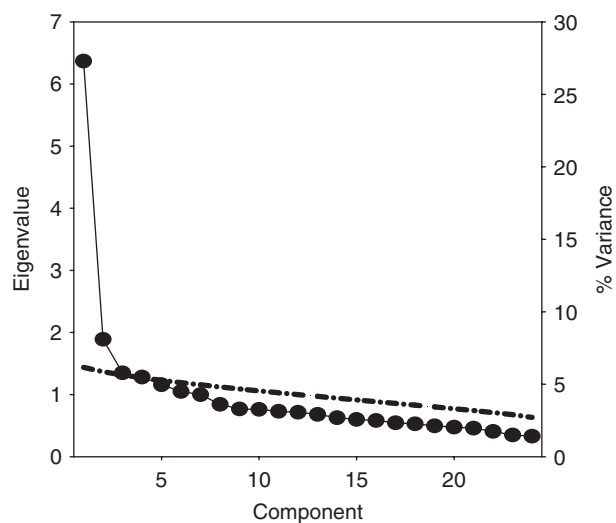

Figure 1. Scree plot of eigenvalues for CSI-24 in the cross-validation sample. The dots show principal component eigenvalues for the CSI-24 in the cross-validation sample. The dashed line is a comparison from a parallel analysis of random numbers in 50,000 simulated data sets with the same number of cases and variables (O'Connor, 2000).

Comprehensive Statistical Properties of the CSI-24 in the Cross-Validation Sample

We evaluated the CSI-24 in the cross-validation sample using 10 criteria from CTT, factor analysis, and the Rasch modeling, as shown in Table II. Columns 1–2 show Cronbach's alpha, columns 3–6 show factor analysis, and columns 7–10 show Rasch modeling results. Below, we interpret the criteria in Table II in order from left to right.

(1) The internal consistency reliability of the CSI-24 was good.
(2) The average correlation between items was moderate. This moderate correlation suggests that the items have uniqueness in addition to their shared construct of somatization.
(3) In a PCA, the second component (GI symptoms) was small, suggesting that it would be difficult to build a robust second factor.
(4) In the CFA, standardized factor loadings were all positive, ranging from medium to high, suggesting that there was a strong first factor that all items measure.
(5–6) In the CFA, the CSI-24 showed a poor fit to a one-factor model, suggesting that it is not strictly unidimensional.
(7–8) Rasch fit criteria were almost always satisfactory, suggesting that the items fit the Guttman–Rasch model.
(9) The CSI-24's sensitivity to differences between people was good.
(10) Items covered a wide range of difficulty values.

A useful feature of IRT models is the ability to assess the extent to which items are targeted to a given sample. Tests can be constructed to assess individuals across a wide range of a trait or can be targeted to assess individuals in a particular range of a trait (e.g., those with clinically significant levels of a trait). Item information curves show that items with low measure scores evaluate healthier

subjects with the least error, whereas items with high measure scores are well targeted toward more severe clinical cases. Most of the CSI items had "high-end" measure scores and would thus yield more accurate information about the most severe cases of somatization. This targeting of items toward the high end fits a measure that is used to identify individuals who may have clinically significant somatization.

Finally, we compared CSI-35 scores with CSI-24 scores in the cross-validation sample. The correlation between total scores was very high ($r = .99$, $p < .001$), indicating that the CSI-24 is a refinement of the original CSI-35 rather than a distinct new instrument.

Relation of the CSI-24 to Child Age and Gender

Scores on the CSI-24 were somewhat higher for older children ($r = .19$); the effect size for age was between "small" and "medium" according to Cohen (1992). Scores also were higher for females, $t(830) = 3.52$, $p = .0005$; the effect size for gender was "medium," $d = $ (mean difference)/$SD_{pooled} = 0.52$.

Scoring of the CSI-24

Administration and scoring instructions for the CSI-24 are presented in Appendix I. Researchers who desire equal-interval scores may prefer to use the Rasch measure scores rather than CSI total sum scores. The Rasch measure scores corresponding to each CSI total score are provided in Appendix II. A figure illustrating the relation between CSI raw total scores and Rasch measure scores is presented in Appendix III. In the midrange, raw total scores and Rasch measure scores are closely related, but at the extremes, the nonlinear Rasch measure scores give a more equal-interval measurement.

Some investigators have modified the CSI by rescoring the 5-point response format into a dichotomous score and calculating a total score as the number of symptoms endorsed (Garber et al., 1991; Garralda et al., 1999; Konijnenberg et al., 2005; Litcher et al., 2001; van de Putte et al., 2006). According to Cohen (1983),

Table II. CSI-24 Psychometric Characteristics in the Cross-validation Sample ($n = 459$)

Classical test theory			Factor analysis			Rasch modeling			
1	2	3	4	5	6	7	8	9	10
Alpha reliability	Ave item r	#2 PCA Eigenvalue	Std CFA loadings 1 factor	Fit CFI (GE 0.95)	Fit RMSEA (LT 0.05)	Infit	Outfit	Person separation reliability	Item separation reliability
.87	.23	1.9	0.34–0.65	0.74	0.07	0.71–1.36	0.68–1.30	.84	.99

Notes. According to Nunnally and Bernstein (1994), reliability of .80 is sufficient for research scales used to characterize groups, but for scales used to make decisions about individuals, "a reliability of .90 is the bare minimum, and a reliability of .95 should be considered the desirable standard" (p. 254). Items that fit the Rasch model have infit and outfit statistics in the range 0.50–1.50 (Linacre, 2006a) or 0.70–1.30 (Bond & Fox, 2001).

dichotomous scoring throws information away, reducing statistical power. We investigated the impact of dichotomous scoring of the CSI by applying three cut points to our data. Cutpoint A dichotomized responses as 0, 1 versus 2, 3, 4; cutpoint B dichotomized responses as 0, 1, 2 versus 3, 4; and cutpoint C dichotomized 0, 1, 2, 3 versus 4. Continuous CSI-24 scores correlated $r = .96$ with cut A scores; $r = .92$ with cut B scores; and $r = .74$ with cut C scores. Cronbach's alpha was reduced from .88 (continuous scores) to .84 (cut A), .80 (cut B), or .68 (cut C). Apparently cutpoints A and B do only slight harm, but cutpoint C seriously reduces reliability. In addition to the loss of reliability, dichotomous scoring makes cross-site comparisons of means problematic and therefore is not recommended.

Discussion

We conducted a multimethod psychometric evaluation to refine the 35-item CSI and to examine the dimensionality of somatization. Using CSI data from 876 pediatric patients with chronic abdominal pain, we identified 11 statistically weak CSI items in a randomly selected learning sample. These items were removed, resulting in the 24-item CSI. An evaluation of the CSI-24 in a cross-validation sample showed that it has good psychometric properties that meet the requirements of CTT, Rasch measurement models, and CFA. The CSI-24 correlates .99 with the CSI-35 and has better psychometric properties, less respondent burden, and items that are more appropriate for youth. Thus, the CSI-24 is preferable to the original 35-item measure unless the specific items deleted from the original measure are of particular interest.

EFA and CFA examined the dimensionality of the CSI-24. A large factor representing the presence of multiple symptoms explained almost 30% of the total variance. All CSI items had positive standardized loadings of medium to high magnitude on this factor, indicating that one dominant general factor underlies symptom reporting associated with multiple organ systems. The second factor (GI symptoms) was quite weak and may be unreliable. These findings are similar to those of Liu and colleagues (1997) who examined the latent structure of somatization symptoms in a population sample of several thousand adults and found that all items loaded strongly on a stable general factor, although smaller factors also were evident. Our findings suggest that somatization in children also has a strong general factor that represents a continuum of symptom reporting and may have minor components that represent specific symptom clusters.

Recent prospective studies provide evidence of the continuity of the general component of somatization from childhood through adolescence and young adulthood (Dhossche, Ferdinand, van der Ende, & Verhulst, 2001; Mulvaney et al., 2006; Steinhausen & Winkler Metzke, 2007). This study contributes further evidence suggesting that a trait-like component of somatization persists across development even as specific symptoms may change.

CSI scores describe the subjective severity of somatic symptoms regardless of etiology. Appropriate medical evaluation is necessary to rule out disease. Similarly, the presence of a somatoform disorder cannot be assumed without appropriate psychiatric evaluation. The CSI should not be used to make individual diagnoses of somatization disorder but must be combined with other medical and psychiatric examinations to identify possible explanations for the symptoms endorsed. However, the CSI may be useful in tracking somatic symptoms over time or monitoring treatment response in patients whose clinical evaluation yields the diagnosis of a functional somatic syndrome or somatoform disorder (cf. Gledhill & Garralda, 2006).

The CSI-24 differs from the 8-item somatic complaints scale of the Child Behavior Checklist (Achenbach, 1991) in that it includes a broad range of symptoms that represent the criteria for various somatoform disorders, as well as symptoms representing the criteria for functional somatic syndromes such as irritable bowel syndrome, fibromyalgia, and chronic fatigue syndrome. Therefore, the CSI-24 will be useful in advancing research on both a trait-like construct of somatization and functional syndromes associated with specific clusters of medically unexplained symptoms.

The present study was based on a predominantly Caucasian clinical sample and may not be representative of somatization in other ethnic groups or in nonclinical community populations. We selected a sample of patients with a primary complaint of chronic abdominal pain because this patient population was used in the original CSI validation study (Walker et al., 1991) and, moreover, abdominal pain is the most common unexplained somatic complaint of childhood and is frequently associated with a variety of other unexplained somatic symptoms. Interestingly, in addition to the major component of somatization, our CSI data showed a minor component characterized by GI symptoms. It is possible that factor analysis of CSI data from patients with other primary complaints (e.g., chest pain, headaches) would yield a minor component consisting of symptoms from the corresponding organ system. Additional research is needed to evaluate whether the factor structure of the CSI in our sample is reproduced in various clinical and community populations.

Recent theoretical and empirical literature highlights an ongoing debate regarding competing approaches to the classification of medically unexplained symptoms (e.g., Hiller, 2006; Kroenke, 2006; Sharpe, Mayou, & Walker, 2006) and whether specific symptom clusters represent one or several syndromes (Aggarwal, McBeth, Zakrzewska, Lunt, & Macfarlane, 2006; Ciccone & Natelson, 2003; Moss-Morris & Spence, 2006). To date, this literature has focused on adult populations. The CSI-24 is a psychometrically sound instrument that can be used to acquire pediatric data relevant to the classification of these symptoms. For example, the CSI-24 can be used to assess whether other pediatric somatic syndromes (e.g., fibromyalgia, chronic fatigue syndrome) provide further evidence of a general symptom reporting factor across multiple organ systems. These data will be important in evaluating the extent to which pediatric somatic syndromes represent discrete entities or are variants of a more general tendency to report somatic symptoms.

Funding

Appendix I
Administration and Scoring Instructions

A. It is recommended that the CSI be administered orally, with the interviewer reading each item aloud and the child selecting a response from a card with the response options. In the case of telephone administration, the response card may be sent to the child prior to the interview, or the interviewer may instruct the child (or his/her parent) to write the options on a piece of paper.

B. If CSI has no missing items, compute the total score by summing the items. Items are scored 0–4. For 24 items, item sum scores range from 0 to 96.

C. If the CSI has at least 20 nonmissing items, compute the total score with two steps:

   a. Calculate the average item score from all non-missing items. Add the ratings of the completed items and divide by the number of items completed.

   b. Multiply the average item score obtained in step 1 by 24, the total number of items (average item score × 24).

D. Excel can do these calculations automatically: $= 24 \times$ AVERAGE (A1:A24), where the item scores are typed into the first column cells A1–A24. The Excel formula gives the right total whether or not there are missing items.

E. If there are fewer than 20 nonmissing items, researchers may either score the CSI or consider it incomplete.

F. Researchers who desire equal-interval scores may prefer to use the Rasch measure scores rather than sum scores. The Rasch measure scores corresponding to each CSI total score are provided in Appendix II. The relation between Rasch measure scores and raw CSI total scores is illustrated in Appendix III.
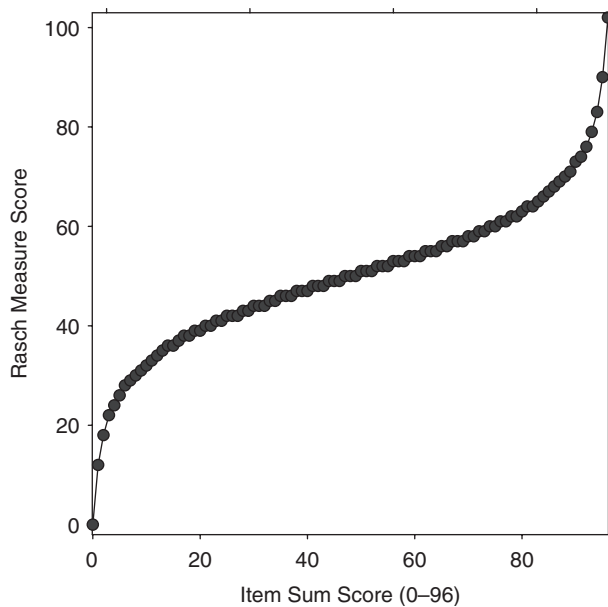
Appendix II. CSI Raw Sum Scores and Corresponding Rasch Measure Scores

Sum score	Rasch measure score	Sum score	Rasch measure score	Sum score	Rasch measure score
0	0	33	45	66	56
1	12	34	45	67	57
2	18	35	46	68	57
3	22	36	46	69	57
4	24	37	46	70	58
5	26	38	47	71	58
6	28	39	47	72	59
7	29	40	47	73	59
8	30	41	48	74	60
9	31	42	48	75	60
10	32	43	48	76	61
11	33	44	49	77	61
12	34	45	49	78	62
13	35	46	49	79	62
14	36	47	50	80	63
15	36	48	50	81	64
16	37	49	50	82	64
17	38	50	51	83	65
18	38	51	51	84	66
19	39	52	51	85	67
20	39	53	52	86	68
21	40	54	52	87	69
22	40	55	52	88	70
23	41	56	53	89	71
24	41	57	53	90	73
25	42	58	53	91	74

(continued)

Appendix II. Continued

Sum score	Rasch measure score	Sum score	Rasch measure score	Sum score	Rasch measure score
26	42	59	54	92	76
27	42	60	54	93	79
28	43	61	54	94	83
29	43	62	55	95	90
30	44	63	55	96	102
31	44	64	55		
32	44	65	56		

Appendix III. The figure shows the relation between traditional sum total scores (X) and Rasch measure scores (Y). In the midrange, traditional total scores and Rasch measure scores are closely related, but at the extremes, the nonlinear measure scores stretch to give a better evaluation of extreme scores, with equal-interval measurement even for extreme scores.

References

Achenbach, T. M. (1991). *Manual for the child behavior checklist 4/18 and 1991 profile*. Burlington: Department of Psychiatry, University of Vermont.

Aggarwal, V. R., McBeth, J., Zakrzewska, J. M., Lunt, M., & Macfarlane, G. J. (2006). The epidemiology of chronic syndromes that are frequently unexplained: Do they have common associated factors? *International Journal of Epidemiology, 35*, 468–476.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.

Berg, B., & Jones, D. P. H (1999). Outcomes of psychiatric intervention in factitious illness by proxy (Munchausen's syndrome by proxy). *Archives of Disease in Childhood, 81*, 465–472.

Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ: Erlbaum.

Bromet, E. J., Goldgaber, D., Carlson, G., Panina, N., Golovakha, E., Gluzman, S., et al. (2000). Children's well-being 11 years after the Chernobyl catastrophe. *Archives of General Psychiatry, 57*, 563–571.

Campo, J. V., & Fritsch, S. L. (1994). Somatization in children and adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry, 33*(9), 1223–1235.

Ciccone, D. S., & Natelson, B. H. (2003). Comorbid illness in women with chronic fatigue syndrome: A test of the single syndrome hypothesis. *Psychosomatic Medicine, 65*, 268–275.

Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement, 7*(3), 249–253.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155–159.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation, 10*(7). Retrieved December 4, 2007, from http://pareonline.net/getvn.asp?v=10&n=7.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297–334.

Cronbach, L. J., & Shavelson, R. J. (2004). My current thoughts on coefficient alpha and successor procedures. *Educational and Psychological Measurement, 64*, 391–418.

Derogatis, L. R., Lipman, R. S., & Covi, L. (1973). SCL-90: An outpatient psychiatric rating scale–preliminary report. *Psychopharmacology Bulletin, 9*, 13–28.

Derogatis, L. R., Lipman, R. S., Rickels, K., Uhlenhuth, E. H., & Covi, L. (1974). The Hopkins Symptom Checklist. *Behavioral Science, 19*, 1–15.

Dhossche, D., Ferdinand, R., van der Ende, J., & Verhulst, F. (2001). Outcome of self-reported functional-somatic symptoms in a community sample of adolescents. *Annals of Clinical Psychiatry, 13*(4), 191–199.

Dorn, L. D., Campo, J. C., Thato, S., Dahl, R. E., Lewin, D., Chandra, R., et al. (2003). Psychological comorbidity and stress reactivity in children and adolescents with recurrent abdominal pain and anxiety disorders. *Journal of American Academy of Child and Adolescent Psychiatry, 42*, 66–75.

Edelen, M. O., & Reeve, B. B. (2007). Applying item response theory (IRT) modeling to questionnaire

development, evaluation, and refinement. *Quality of Life Research, 16*(1), 5–18.

Embretson, S. E. (1996). The new rules of measurement. *Psychological Assessment, 8*, 341–349.

Garber, J., Walker, L. S., & Zeman, J. (1991). Somatization symptoms in a community sample of children and adolescents: Further validation of the Children's Somatization Inventory. *Psychological Assessment, 3*, 588–595.

Garralda, E., Rangel, L., Levin, M., Roberts, H., & Ukoumunne, O. (1999). Psychiatric adjustment in adolescents with a history of chronic fatigue syndrome. *Journal of the American Academy of Child and Adolescent Psychiatry, 38*, 1515–1521.

Gledhill, J., & Garralda, M. E. (2006). Functional symptoms and somatoform disorders in children and adolescents: The role of standardized measures in assessment. *Child & Adolescent Mental Health, 11*, 208–214.

Harrell, F. E. (2001). *Regression modeling strategies*. New York: Springer.

Hiller, W. (2006). Don't change a winning horse. *Journal of Psychosomatic Research, 60*, 345–347.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55.

Konijnenberg, A. Y., de Graeff-Meeder, E. R., van der Hoeven, J., Kimpen, J. L. L., Buitelaar, J. K., Uiterwaal, C., et al. (2006). Psychiatric morbidity in children with medically unexplained chronic pain: Diagnosis from the pediatrician's perspective. *Pediatrics, 117*, 889–897.

Konijnenberg, A. Y., Uiterwaal, C., Kimpen, J., van der Hoeven, J., Buitelaar, J., & de Graeff-Meeder, E. (2005). Children with unexplained chronic pain: Substantial impairment in everyday life. *Archives of Disease in Childhood*, published online 17 May 2005; 10.1136/adc.2004.056820.

Kroenke, K. (2006). Physical symptom disorder: A simpler diagnostic category for somatization spectrum conditions. *Journal of Psychosomatic Research, 60*, 335–339.

Linacre, J. M. (2006a). *Misfit diagnosis: Infit outfit mean-square standardized*. Retrieved June 1, 2006, from http://www.winsteps.com/winman/diagnosingmisfit.htm.

Linacre, J. M. (2006b). Rasch analysis of rank-ordered data. *Journal of Applied Measurement, 7*, 129–139.

Linacre, J. M., & Wright, B. D. (1999). *WINSTEPS computer program* (Version 3.47). Chicago: MESA Press.

Lipani, T., & Walker, L. S. (2006). Children's appraisals and coping with pain: Relation to maternal ratings of worry and restriction in family activities. *Journal of Pediatric Psychology, 31*, 667–673.

Litcher, L., Bromet, E., Carlson, G., Gilbert, T., Panina, N., Golovakha, E., et al. (2001). Ukrainian application of the Children's Somatization Inventory: Psychometric properties and associations with internalizing symptoms. *Journal of Abnormal Child Psychology, 29*, 165–175.

Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: John Wiley & Sons.

Liu, G., Clark, M. R., & Eaton, W. W. (1997). Structural factor analyses for medically unexplained symptoms of somatization disorder in the Epidemiologic Catchment Area Study. *Psychological Medicine, 27*, 617–626.

Lord, F. M. (1986). Maximum likelihood and Bayesian parameter estimation in item response theory. *Journal of Educational Measurement, 23*, 157–162.

Meesters, C., Muris, P., Ghys, A., Reumerman, T., & Rooijmans, M. (2003). The Children's Somatization Inventory: Further evidence for its reliability and validity in a pediatric and community sample of Dutch children and adolescents. *Journal of Pediatric Psychology, 28*, 413–422.

Moss-Morris, R., & Spence, M. (2006). To ''lump'' or to ''split'' the functional somatic syndromes: Can infectious and emotional risk factors differentiate between the onset of chronic fatigue syndrome and irritable bowel syndrome? *Psychosomatic Medicine, 68*, 463–469.

Mulvaney, S., Lambert, W., Garber, J., & Walker, L. S. (2006). Trajectories of symptoms and impairment for pediatric patients with functional abdominal pain: A 5-year longitudinal study. *Journal of American Academy of Child and Adolescent Psychiatry, 45*, 737–744.

Muris, P., Vlaeyen, J., & Meesters, C. (2001). The relationship between anxiety sensitivity and fear of pain in healthy adolescents. *Behavioural Research and Therapy, 39*, 1357–1368.

Muthen, L. K., & Muthen, B. O. (2003). *Traditional latent variable modeling using Mplus*. Los Angeles: www.statmodel.com.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instrumentation, and Computers, 32*, 396–402.

Pennebaker, J. W. (1982). *The psychology of physical symptoms*. New York: Springer.

Rasch, G. (1960/1980). *Probabilistic models for some intelligence and attainment tests*. Chicago: University of Chicago Press.

Robbins, J. M., Kirmayer, L. J., & Hemami, S. (1997). Latent variable models of functional somatic distress. *Journal of Nervous and Mental Disease, 185*, 606–615.

Robins, P. M., Smith, S. M., & Proujansky, R. (2002). Children with recurrent abdominal pain: Comparison of community and tertiary care samples. *Children's Health Care, 31*, 93–106.

Rocha, E. M., & Prkachin, K. M. (2007). Temperament and pain reactivity predict health behavior seven years later. *Journal of Pediatric Psychology, 32*(4), 393–399.

Rocha, E. M., Prkachin, K. M., Beaumont, S. L., Hardy, C. L., & Zumbo, B. D. (2003). Pain reactivity and somatization in kindergarten-age children. *Journal of Pediatric Psychology, 28*, 47–57.

Rubin, D. B. (1991). EM and beyond. *Psychometrika, 56*(2), 241–254.

Sharpe, M., Mayou, R., & Walker, J. (2006). Bodily symptoms: New approaches to classification. *Journal of Psychosomatic Research, 60*, 353–356.

Smith, M. S., Martin-Herz, S. P., Womack, W. M., & Marsigan, J. L. (2003). Comparative study of anxiety, depression, somatization, functional disability, and illness attribution in adolescents with chronic fatigue or migraine. *Pediatrics, 111*, 876–881.

Smith, M. S., Martin-Herz, S. P., Womack, W. M., & McMahon, R. J. (1999). Recurrent headache in adolescents: Nonreferred versus clinic population. *Headache, 39*, 616–624.

Steiger, J. H. (2000). Point estimation, hypothesis testing, and interval estimation using the RMSEA: Some comments and a reply to Hayduck and Glaser. *Structural Equation Modeling, 72*, 149–162.

Steinhausen, H. C., & Winkler Metzke, C. (2007). Continuity of functional-somatic symptoms from late childhood to young adulthood in a community sample. *Journal of Child Psychology & Psychiatry, 48*(5), 508–513.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston: Allyn and Bacon.

Tsao, J. C., & Zeltzer, L. K. (2003). Sex differences in pain-related symptoms and self-initiated school nurse visits among pre-adolescents. *Journal of Pain and Symptom Management, 25*, 427–480.

van de Putte, E. M., Engelbert, R. H., Kuis, W., Kimpen, J. L., & Uiterwaal, C. (2006). How fatigue is related to other symptoms. *Archives of Disease in Childhood*, published online 5 June 2006; 10.1136/adc.2006.094623.

Walker, L. S., Garber, J., & Green, J. W. (1991). Somatization symptoms in pediatric abdominal pain patients: Relation to chronicity of abdominal pain and parent somatization. *Journal of Abnormal Child Psychology, 19*, 379–394.

Walker, L. S., Garber, J., & Greene, J. W. (1993). Psychosocial correlates of recurrent childhood pain: A comparison of pediatric patients with recurrent abdominal pain, organic illness, and psychiatric disorders. *Journal of Abnormal Psychology, 102*, 248–258.

Walker, L. S., Garber, J., Smith, C. A., Van Slyke, D. A., & Claar, R. L. (2001). The relation of daily stressors to somatic and emotional symptoms in children with and without recurrent abdominal pain. *Journal of Consulting and Clinical Psychology, 69*, 85–91.

Walker, L. S., Williams, S. E., Smith, C. A., Garber, J., Van Slyke, D. A., & Lipani, T. A. (2006). Parent attention versus distraction: Impact on symptom complaints by children with and without chronic functional abdominal pain. *Pain, 122*, 43–52.

White, K. S., Alday, C. S., & Spirito, A. (2001). Characteristics of children presenting to a behavioral treatment program for pediatric headache. *Journal of Clinical Psychology in Medical Settings, 8*, 109–117.