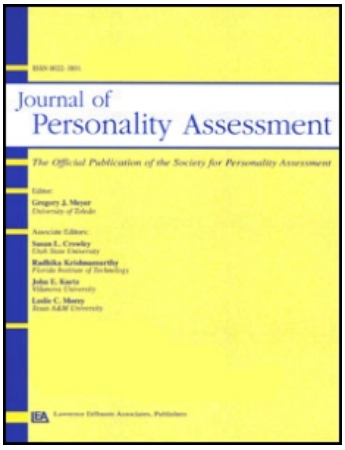

Journal of Personality Assessment
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t775653663

Assessing the Factorial Invariance of Harter's Self-Concept Measures: Comparing Preadolescents With and Without Spina Bifida Using Child, Parent, and Teacher Report

Azure Dee Welborn Thill [a]; Grayson N. Holmbeck [b]; Fred B. Bryant [b]; Cheryl Nelson [b]; Amra Skocic [b]; Nicole Uli [b]
[a] Department of Psychology, North Central College. [b] Department of Psychology, Loyola University Chicago.

To cite this Article Thill, Azure Dee Welborn, Holmbeck, Grayson N., Bryant, Fred B., Nelson, Cheryl, Skocic, Amra and Uli, Nicole(2003) 'Assessing the Factorial Invariance of Harter's Self-Concept Measures: Comparing Preadolescents With and Without Spina Bifida Using Child, Parent, and Teacher Report', Journal of Personality Assessment, 81: 2, 111 — 122
To link to this Article: DOI: 10.1207/S15327752JPA8102_03
URL: http://dx.doi.org/10.1207/S15327752JPA8102_03

JOURNAL OF PERSONALITY ASSESSMENT, *81*(1), 111–122

Assessing the Factorial Invariance of Harter's Self-Concept Measures: Comparing Preadolescents With and Without Spina Bifida Using Child, Parent, and Teacher Report

Azure Dee Welborn Thill
Department of Psychology
North Central College

Grayson N. Holmbeck, Fred B. Bryant, Cheryl Nelson, Amra Skocic, and Nicole Uli
Department of Psychology
Loyola University Chicago

The purpose of this study was to determine whether the structure of children's self-concept, as rated by children themselves on the Self-Perception Profile for Children (SPPC; Harter, 1985), and the structure of children's competence, as rated by their parents on the Rating Scale of Child's Actual Behavior (PRS) and teachers on the Rating Scale of Child's Actual Behavior (TRS; Harter, 1985), are similar across samples of children with and without spina bifida (children were 8 and 9 years of age; $n = 68$ in each sample). Using confirmatory factor analyses, results revealed that a multidimensional model for the SPPC and the TRS fit the data well for both samples. On the other hand, the model for both mothers' and fathers' PRS ratings did not fit the data for the spina bifida sample. Further tests of factorial invariance conducted on the SPPC and TRS revealed that the spina bifida and able-bodied samples tended to have the same pattern and magnitude of factor loadings for both instruments. There was some invariance, however, in the amount of unique error variance accounted for across samples. These results bolster our confidence in using the SPPC and TRS to compare children with and without spina bifida.

Self-concept is a construct widely studied in the field of social development (e.g., Demo, 1992; Harter, 1998) and refers to ways in which children perceive their own attributes or traits. Within the developmental literature, this construct has been approached in two ways. First, self-concept has been conceived as a single, overarching global construct (e.g., Piers & Harris, 1964). Other researchers, however, believe that self-concept is a multifaceted construct such that there are several, unique subcomponents that together form this construct (e.g., Harter, 1985; Marsh & Hocevar, 1985).

One widely used, multidimensional self-concept measure is the 36-item Self-Perception Profile for Children (SPPC; Harter, 1985). Specifically, the instrument consists of one Global Self-Worth subscale and five domain-specific subscales (6 items each): Scholastic Competence, Social Acceptance, Athletic Competence, Physical Appearance, and Behavioral Conduct. Research has demonstrated high internal consistency (e.g., Schumann et al., 1999) and test–retest reliability (e.g., Granleese & Joseph, 1994) for each of the subscales on the SPPC. Furthermore, a teacher version (Teacher's Rating Scale of Child's Actual Behavior [TRS]; Harter, 1985) of the instrument has been adapted from the SPPC, which may also be modified to create a parent version (Parent's Rating Scale of Child's Actual Behavior [PRS]; Harter, 1985). Both of these adult versions of the measure have demonstrated adequate psychometric properties, including convergent and discriminant validity (Cole, Gondoli, & Peeke, 1998). Similar to the child version of the instrument, the TRS and the PRS contain five domain-specific subscales (Scholastic Competence, Social Acceptance, Athletic Competence, Physical Appearance, and Behavioral Conduct). Unlike the SPPC, however, the TRS and the PRS

do not include a Global Self-Worth subscale and instead of 6 items per subscale, the TRS and PRS include 3 items per subscale (for a total of 15 items).

Another important distinction between the child and adult versions of Harter's (1985) measure is terminological. Investigators who use Harter's (1985) measures often refer to the child version as a measure of *self-concept* (e.g., Veerman, ten Brink, Straathof, & Treffers, 1996). On the other hand, investigators who use the parent and teacher versions usually refer to the instrument as a measure of children's *competence* (e.g., Cole et al., 1998). Indeed, it is likely that the best reporter of one's self-concept is the target person (e.g., the child), but it seems reasonable to assume that multiple informants can report on an individual's competence (e.g., parents and teachers). Although some subscales of the measure (e.g., Physical Appearance) do not necessarily assess competence, we use both of these terms so as to be consistent with past literature and to distinguish between child and adult perceptions.

Within the self-concept and competence literatures, research has replicated the factor structure of the SPPC originally obtained by Harter (1985; e.g., Granleese & Joseph, 1993; Schumann et al., 1999; Trent, Russell, & Cooney, 1994; Veerman et al., 1996) as well as the factor structure for the TRS and PRS (Cole et al., 1998). In other words, the finding that this self-concept measure is composed of five unique subscales is quite robust, thus validating the instrument. Moreover, factorial validity and factorial invariance (i.e., the equivalence of factor structures across different groups) has been examined in recent studies. For example, there is evidence that the structure of children's self-concept, as measured by the SPPC, is different in White versus African American populations (Schumann et al., 1999). These differences, though, are primarily the result of factorial complexity within the African American population (i.e., several of the 36 individual items contained in the instrument loaded on more than one factor when a principal components analysis was conducted). Other researchers, however, have found equivalence of factor structures across diverse populations. For example, Veerman et al. (1996) found evidence for the factorial invariance of the SPPC in nonclinic versus clinic samples.

Although previous research has demonstrated factorial invariance for the SPPC across certain populations (e.g., clinic vs. nonclinic groups), the equivalence of factor structure has not been assessed across pediatric versus nonpediatric samples. Because this instrument is often used to compare these two groups (e.g., Appleton et al., 1997), it is important to establish factorial invariance across these two types of samples. There is reason to believe that these two samples of children (and their parents and teachers) may not interpret the instrument in the same manner. With respect to this study, the physical limitations associated with *spina bifida* (a congenital birth defect caused by a failed closure of one or more vertebrae during the early weeks of pregnancy, which typically produces urinary, bowel, orthopedic, and neurological difficulties) likely have important implications for one's self-concept, and these effects may produce differences in the structure of self-concept across samples. Specifically, the Athletic Competence and Physical Appearance subscales may be more important determinants of self-concept among children with spina bifida than they are among able-bodied children. Accordingly, the purpose of this study was to determine whether the structure of self-concept, as measured by Harter's (1985) SPPC, is comparable across samples of children with and without spina bifida. Parent and teacher versions of the instrument, which tap adult ratings of children's competence in comparison to other children, were also examined for factorial invariance. We turn now to a brief description of issues related to factorial invariance and the assessment of self-concept in pediatric versus nonpediatric samples.

ISSUES OF FACTORIAL INVARIANCE

When examining factorial invariance across samples, three basic questions must be addressed, each of which focuses on a different aspect of factorial structure. Each of the Harter (1985) instruments (i.e., the SPPC, TRS, and PRS) were examined separately; three questions were posed for each instrument.

The first structural question concerns whether the same underlying dimensions of self-concept are in fact relevant for both populations. In other words, is the SPPC a valid multidimensional measure of self-concept for both the spina bifida and able-bodied samples? Specifically, this question addresses whether the domains of Scholastic Competence, Social Acceptance, Athletic Competence, Physical Appearance, and Behavioral Conduct underlie self-concept in both samples as assessed with the SPPC. We addressed this question by imposing the factor structure obtained by Harter and others (e.g., Harter, 1985; Trent et al., 1994) on the SPPC (and PRS and TRS) data for each group separately and evaluated whether the model provided a reasonable fit for each group.

Assuming that the same five domains are relevant for both groups, a second structural question examines whether the relationship between domain-specific subscales and total self-concept is equivalent for children with and without spina bifida. In other words, do the five domains have the same loadings on the general self-concept factor for the two groups of children? Or do children with spina bifida use domain-specific aspects of self-concept differently to define their total self-concept compared to children without spina bifida? We addressed this question of factor loading invariance by evaluating whether the factor loadings were equivalent across groups (with a multigroup confirmatory factor analysis [CFA]).

Assuming that the domains of self-concept have similar factor loadings for both groups of children, a third structural

issue concerns the amount of variability in each specific domain of self-concept that is unrelated to self-concept. In other words, is the amount of variance in each domain-specific subscale that reflects other influences besides self-concept the same for children with and without spina bifida? Or is there more unique, unexplained variance in the specific domains of self-concept for one group relative to the other? Thus, we investigated (a) invariance in the pattern and magnitude of factor loadings across groups and (b) invariance in the unique error variance of domain-specific self-concepts.

It is important to note that answers to the first two of these three structural questions (i.e., whether the pattern and magnitude of factor loadings are invariant across groups) provide sufficient evidence regarding factorial invariance, regardless of the findings pertaining to the third question. This is because the two groups may well differ in the variance of measurement errors for certain domains (i.e., the unique variance in domain-specific subscales), even if there is factorial invariance across groups (Kline, 1998; MacCallum & Tucker, 1991).

STRUCTURAL HYPOTHESES

If our data support the factorial invariance hypothesis for children with and without spina bifida, then we can be more confident in using the Harter (1985) instruments (i.e., the SPPC, TRS, and PRS) to compare the development of self-concept and adult-rated child competence in children with and without this pediatric condition. However, differences across samples may be found. As noted earlier, the Athletic Competence and Physical Appearance subscales of the SPPC may play a more important role in defining overall self-concept in the spina bifida sample compared to the able-bodied sample. Thus, we hypothesized that these two subscales would load more highly on total self-concept in the spina bifida sample when compared with the able-bodied sample. A similar argument could be made for adult ratings of children's competence; teachers and parents may be particularly attuned to Athletic Competence and Physical Appearance when rating a child's level of competence because of the visible physical disabilities associated with spina bifida. Thus, we expected these subscales to load higher on children's overall competence in the spina bifida sample than in the able-bodied sample.

METHOD

Participants

As part of a larger longitudinal study investigating the transition into adolescence (e.g., Holmbeck et al., 1998; Holmbeck et al., 1997; Holmbeck, Johnson, et al., 2002; Holmbeck, Shapera, & Hommeyer, 2002; Holmbeck et al., in press; McKernon et al., 2001; findings of this study do not overlap with any of the findings from these earlier publications), families of children with and without spina bifida participated when each child was 8 to 9 years of age. Specifically, the sample included 68 families of children with spina bifida (37 boys, 31 girls; *M* age = 8.34) and a matched comparison sample of 68 families of able-bodied children (37 boys, 31 girls; *M* age = 8.49).

The two samples were successfully matched on 10 demographic variables (see Table 1). A range of family incomes was represented in both samples. The majority of participants were

TABLE 1
Demographics: Comparisons Across Samples

Demographic Characteristics	*Spina Bifida*				*Able Bodied*				*Statistical Tests*	*Effect Size (r)*
	M	*SD*	*%*	*n*	*M*	*SD*	*%*	*n*		
Child age	8.34	0.48			8.49	0.50			$t(134) = -1.75$	.149
Maternal age	37.74	5.19			37.74	4.84			$t(134) = 0.00$	.000
Paternal age	41.02	5.45			40.63	6.50			$t(105) = 0.33$	.032
Child gender										
Male			54.41	37			54.41	37	$\chi^2(1) = 0.00$	.000
Female			45.59	31			45.59	31		
Child ethnicity										
White			82.35	56			91.18	62	$\chi^2(1) = 2.30$	.130
Other			17.65	12			8.82	6		
Child birth order	2.12	1.38			2.06	1.29			$t(129) = 0.27$	.024
Marital status										
Two-parent intact			80.88	55			69.12	47	$\chi^2(1) = 2.51$	.136
Nonintact			19.12	13			30.88	21		
Maternal income	5.75	2.57			5.73	2.45			$t(130) = 0.05$	.005
Paternal income	6.24	2.50			6.35	2.22			$t(105) = -0.24$	.023
Hollingshead SES	43.12	10.57			46.46	10.89			$t(131) = -1.80$	.155

Note. $N = 68$ for each sample. Family income is rated on a scale ranging from 1 to 11 with 1 < \$10,000, 5 = \$40,000 to 49,999, 10 = \$90,000 to 99,999, and 11 > \$100,000. The Hollingshead (1975) four-factor index of socioeconomic status (SES) is based on a composite of maternal education, paternal education, maternal occupational status, and paternal occupational status. All statistics were nonsignificant ($p > .05$).

White (82% in the spina bifida sample, 91% in the able-bodied sample). Although biological mothers from all families from both groups participated in the study, only 55 (81%) fathers/stepfathers from the spina bifida sample and 52 (76%) of the fathers in the able-bodied sample participated.

Information on a number of physical status variables for the spina bifida group were obtained based on maternal report and/or from information gleaned from the child's medical chart: (a) spina lesion level was 32% sacral, 54% lumbosacral or lumbar, 13% thoracic; (b) spina bifida type was 82% myelomeningocele, 12% lipomeningocele, 6% other; (c) shunt status was 71% shunted, 28% not shunted; and (d) ambulation was 18% no assistance, 63% assistance with braces, 19% assistance with a wheelchair. The average number of shunt surgeries among those with shunts was 2.50 (SD = 2.91).

Participant Recruitment

Participating families in the spina bifida group were recruited from lists provided by four sources: (a) a children's hospital, (b) a children's hospital that cares exclusively for youngsters with physical disabilities, (c) a university-based medical center, and (d) a statewide spina bifida association. A recruitment letter was sent to all parents of children within the 8- to 9-year-old age range (and those who would reach this age within the following year). Letters were followed up with phone calls. Out of 310 nonoverlapping child names from the four sources, 72 families lived too far away (greater than 120 miles from the laboratory); 64 declined to participate; 56 could not be reached (due to invalid addresses and phone numbers); 16 did not speak English; 14 children turned 10 years of age before a visit could be scheduled; 11 children did not have spina bifida; and 9 were excluded for miscellaneous reasons. A comparison of participating children (n = 68) with children from families that declined to participate (n = 64) revealed no differences with respect to lesion level, $\chi^2(2, N = 116) = 0.62$, $p > .059$ (effect size, $r = .073$), or type of spina bifida (myelomeningocele vs. lipomeningocele), $\chi^2(1, N = 119) = 1.63$, $p > .05$ (effect size, $r = .117$).

Participating families from the able-bodied comparison group were recruited by contacting schools where the children with spina bifida were enrolled. To obtain a comparison group the same size as the spina bifida group, it proved unnecessary to contact all possible schools. Instead, a representative listing of schools was chosen based on the following factors: location, the average family income of the surrounding community, and the ethnic distribution in the school. The initial list of schools was based on school enrollment information for the first 42 children with spina bifida who agreed to participate in our study. This list provided us with the necessary number of potential able-bodied participants to yield a satisfactory matching of groups. Of these 42 schools, 24 were ruled out for one of several reasons (e.g., the community was too far away to run multiple families in that community given limited funding resources; there was a racial or income distribution in the school that could have produced matching difficulties). Of the remaining 18 schools, 12 agreed to participate and 6 declined. At the participating schools, recruitment letters (as well as self-addressed, stamped envelopes) were sent home with the comparison children in our age range; parents could then return a slip indicating their consent to participate. To obtain the sample used in this study, roughly 1,700 letters were sent. The low recruitment rate is attributable, at least in part, to the longitudinal nature of the study that was described in detail in the recruitment letter.

Procedure

Trained graduate and undergraduate research assistants visited families and children in their homes. Home visits lasted approximately 3 hr. Parents and children completed questionnaires as well as audiotaped and videotaped interaction tasks (e.g., families were presented with an unfamiliar board game and given 10 min to come up with their own rules and play the game). The order of all questionnaires and family interaction tasks were counterbalanced across families. Release of information for teacher measures was obtained from all families (teacher response rates were 96% for the spina bifida sample and 84% for the comparison sample). At the end of each home visit, families were paid $50 and children were given their choice of a small gift (e.g., a colored pencil).

Measures

Although several measures of child and family adjustment were obtained through questionnaires and observed interactions, this study focused only on self-report measures of self-concept/competence. Three versions of Harter's (1985) Self-Perception Profile were used to assess self-concept/competence for this study: the SPPC, the TRS, and the PRS.

SPPC. As mentioned previously, the SPPC is a 36-item self-report measure of self-concept. The instrument includes one Global Self-Worth subscale and five domain-specific subscales (Scholastic Competence, Social Acceptance, Athletic Competence, Physical Appearance, and Behavioral Conduct). Six items are used to assess each of the six subscales. For each item, children read two statements (e.g., "'Some kids feel that they are very good at their school work.' BUT 'Other kids worry about whether they can do the school work assigned to them.'") and decided which statement was more like themselves. After deciding which statement was more like themselves, children then decided if that statement was "sort of true for me" or "really true for me." For each item, a score of 4 represents the most satisfactory self-assessment, and a score of 1 represents the least satisfactory self-assessment, after recoding reverse-coded items.

TABLE 2
Means, Standard Deviations, Intercorrelations, and Alpha Coefficients for the Spina Bifida and Able-Bodied Samples on the SPPC, TRS, and PRS

Subscale	*1*	*2*	*3*	*4*	*5*	*M*	*SD*	
SPPC								
1. Scholastic Competence		0.32	0.32	0.39	0.51	3.17	0.58	.70
2. Social Acceptance	0.56		0.31	0.16	0.33	3.01	0.62	.71
3. Athletic Competence	0.38	0.27		0.39	0.46	2.83	0.65	.72
4. Physical Appearance	0.29	0.48	0.26		0.41	3.18	0.64	.76
5. Behavioral Conduct	0.47	0.43	0.17	0.13		3.09	0.54	.60
M	2.75	2.86	2.57	2.91	3.02			
SD	0.78	0.77	0.85	0.79	0.62			
α	0.71	0.67	0.74	0.74	0.53			
TRS								
1. Scholastic Competence		0.19	–0.02	0.27	0.30	3.24	0.73	.90
2. Social Acceptance	0.38		0.38	0.51	0.42	3.17	0.75	.93
3. Athletic Competence	0.50	0.38		0.23	0.06	2.81	0.57	.87
4. Physical Appearance	0.09	0.24	0.18		0.42	3.39	0.52	.89
5. Behavioral Conduct	0.38	0.12	0.09	0.14		3.31	0.71	.90
M	2.35	2.93	1.84	3.25	3.35			
SD	0.74	0.87	0.71	0.67	0.85			
α	0.82	0.92	0.88	0.88	0.93			
PRS (mothers)								
1. Scholastic Competence		0.23	–0.01	0.06	0.26	3.39	0.70	.80
2. Social Acceptance	0.24		0.38	0.40	0.09	3.38	0.54	.78
3. Athletic Competence	0.25	0.41		0.33	–0.04	2.86	0.62	.79
4. Physical Appearance	0.08	0.12	0.22		0.20	3.78	0.36	.67
5. Behavioral Conduct	0.14	–0.07	–0.20	0.25		3.28	0.56	.81
M	2.69	3.18	2.03	3.52	3.28			
SD	0.79	0.59	0.62	0.48	0.57			
α	0.84	0.79	0.76	0.76	0.75			
PRS (fathers)								
1. Scholastic Competence		0.20	0.31	0.18	0.38	3.29	0.63	.72
2. Social Acceptance	0.19		0.26	0.23	0.16	3.31	0.58	.75
3. Athletic Competence	0.23	0.09		0.29	0.12	2.85	0.60	.75
4. Physical Appearance	–0.03	0.42	–0.04		0.27	3.68	0.42	.67
5. Behavioral Conduct	0.29	0.01	0.03	0.24		3.41	0.61	.78
M	2.93	3.24	2.08	3.57	3.49			
SD	0.64	0.66	0.68	0.45	0.47			
α	0.80	0.73	0.81	0.76	0.69			

Note. α = Cronbach's alpha (a reliability index of internal consistency). Covariance matrices can be constructed by multiplying the correlation coefficients with the standard deviation of each subscale. For each instrument (SPPC, TRS, and PRS), statistics for the spina bifida sample are tabled below the diagonal; statistics for the able-bodied sample are tabled above the diagonal. Samples sizes for the SPPC data were: spina bifida group, $n = 62$; able-bodied group, $n = 68$. Sample sizes for the TRS data were: spina bifida group, $n = 65$; able-bodied group, $n = 56$. Sample sizes for the mothers' PRS data were: spina bifida group, $n = 66$; able-bodied group, $n = 68$. Sample sizes for the fathers' PRS data were: spina bifida group, $n = 53$; able-bodied group, $n = 52$. SPPC = Self-Perception Profile for Children; TRS = Teacher's Rating Scale of Child's Actual Behavior; PRS = Parent's Rating Scale of Child's Actual Behavior.

Within each subscale, the directions of the statements were counterbalanced (i.e., for 3 of the items, the satisfactory statements are on the right, and for 3 of the items, the satisfactory statements are on the left). Domain scores were obtained by calculating the mean of the 6 items within each subscale. Table 2 presents descriptive statistics, intercorrelations, and reliabilities for each SPPC subscale for each sample. Note that the statistics in Table 2 are based on only 62 participants in the spina bifida group (compared to 68 in the able-bodied group). This reduced sample size is the result of some children's failure to complete and/or follow the instructions on the SPPC.

TRS. The TRS was developed by Harter (1985) to parallel the child's self-report measure of self-concept. As noted earlier, the TRS assesses teacher's perceptions of children's competence across the same domains included in the SPPC. Similar to the SPPC, teachers read a set of two statements (e.g., "'This pupil is really good at his/her schoolwork.' OR 'This pupil can't do the school work assigned.'") and decided which statement was, in their opinion, more like the student. They then decided whether they thought each statement was "sort of true" or "really true" for the student. Teachers were asked to rate the child's actual behavior for each item as opposed to rating how they thought the child would respond to each question. Furthermore, in contrast to the SPPC, there are only 15 items on the teacher version of the instrument (3 items, instead of 6, are used to assess the five domain-specific subscales; global self-worth is not assessed with the teacher version). The directions of the statements

within each subscale were counterbalanced. Calculating the mean of the 3 items within each subscale provides a domain score that can be directly compared to the child's scores. Table 2 presents descriptive statistics, intercorrelations, and reliabilities for each TRS subscale for each sample. Note that the statistics in Table 2 are based on only 65 participants in the spina bifida group and 56 in the able-bodied group. This reduced sample size is the result of teachers' participation rate or failure to complete and/or follow the instructions on the TRS.

PRS. Mothers and fathers independently completed a parent version of the instrument (PRS), which was adapted from Harter's (1985) TRS. Like the TRS, the PRS parallels the child-report version and assesses parent ratings of child competence. Similar to both the child and teacher versions of the instrument, each parent read two statements (e.g., "'My child is really good at his/her school work.' OR 'My child can't do the school work assigned.') and decided which statement was, in their opinion, more like their child. Parents then decided if that statement was "sort of true for my child" or "really true for my child." Similar to the teacher version of the instrument, there are only 15 items on the parent version (3 items per subscale; global self-worth is not assessed). The directions of statements within each subscale were counterbalanced. Domain scores were achieved by calculating the mean of the 3 items within each subscale. The domain scores from the PRS can be directly compared to the domain scores of both the SPPC and the TRS. Table 2 presents descriptive statistics, intercorrelations, and reliabilities for each PRS subscale for each sample. Note that the statistics in Table 2 for the mothers are based on 66 participants in the spina bifida group compared to 68 in the able-bodied group. The statistics in Table 2 for the fathers are based on only 53 participants in the spina bifida group and 52 in the able-bodied group. This reduced sample size is a result of fathers' participation rate and because some families were mother-only, single-parent families.

Analytic Strategy

Because Harter (1985) argued that global self-worth is a construct independent of the other subscales tapped by her instrument (Harter, 1985) and because teachers and parents did not complete the Global Self-Worth items, these analyses focused only on the five domain-specific subscales for each of the instruments. It is important to note that the relatively small sizes of these samples precluded the use of CFA to examine the appropriateness of item placement for each subscale of the SPPC. That is, there were too few observations to compare the factor loadings of each individual item within each subscale across the spina bifida and able-bodied samples groups. This, however, was not our purpose. Instead, our goal was to construct each subscale in the same way for both groups and then to use CFA to determine whether the higher order construct of self-concept influenced responses to each subscale equivalently across the two groups. In other words, we sought to answer the question of whether self-concept has the same meaning for children with and without spina bifida when self-concept was measured in terms of these five subscales of the SPPC.

Thus, we did not perform structural analyses at the item level. Instead, we used the five subscale scores (i.e., the mean of the subscale items) as measured variables in our model. Each subscale was then predicted by a single, higher order latent variable (i.e., self-concept/competence) in a one-factor CFA. This type of structural analysis has been termed a *partial aggregation* model (Bagozzi & Edwards, 1998; Bagozzi & Heatherton, 1994). In a partially aggregated measurement model, responses to multiple items measuring a specific subscale are combined (i.e., averaged) to form a composite measure. The summary scores for multiple composite measures are then used as indicators in the measurement model, and these indicators are analyzed together to examine their underlying hierarchical structure.

The primary advantages of the partial aggregation subscale score (molecular) approach, relative to a total disaggregation, individual-item (atomistic) approach are (a) it reduces the number of parameters to be estimated so that analyses can be run with smaller sample sizes and (b) it tends to decrease the amount of measurement error in the indicators by aggregating items to form more reliable, composite indexes; thus, models generally fit the data better (Bagozzi & Edwards, 1998; Bagozzi & Heatherton, 1994; Hull, Lehn, & Tedlie, 1991). It should be noted, however, that this modeling approach has one important disadvantage. Because the measured variables consist of a mean score for each subscale, any distinctiveness among the items of the subscales is obscured (Bagozzi & Edwards, 1998). That is, the individual contributions of the constituent items of each subscale remain unknown when using the partial aggregation approach.

Using the five composite subscales as measured variables in a partially aggregated, one-factor CFA model provided a potentially useful form of hierarchical factor analysis. In particular, the standardized factor loadings in the partial aggregation model represented correlations between the higher order self-concept construct and each subscale; squaring these standardized loadings provided estimates of the proportion of variance in each subscale that was related to global self-concept (see Bagozzi & Heatherton, 1994). In addition, the standardized unique variance for each subscale represented the proportion of its variance that was specific to the particular domain and that shared nothing in common with global self-concept. Applying this modeling approach in multigroup CFA enabled us to determine whether children with and without spina bifida used the first-order SPPC subscales in the same ways to define the broader construct of self-concept.

It is important to clarify that this one-factor model does not assume all five subscales contribute equally to the total

score. Rather, the model assumes that the constituent items for each subscale contribute equally to that subscale but allows the five subscales to have different factor loadings in defining a single, overarching global factor within the spina bifida and able-bodied groups. Multigroup CFA enabled us to test, for each subscale, whether its factor loading differed across the two groups.

Concerning our use of CFA, Jöreskog (1993) distinguished among three forms of CFA application: (a) *strictly confirmatory,* in which the researcher seeks to accept or reject a single a priori model; (b) *alternative models,* in which the researcher specifies several competing models and seeks to determine which one should be selected; and (c) *model generating,* in which the researcher specifies a tentative initial model or models, which may be modified and tested again using the same data. Our use of CFA can be considered strictly confirmatory in the sense that we estimated a single, well-established model and tested its factorial invariance across two samples.

We used LISREL 8 (Jöreskog & Sörbom, 1996; also see Byrne, 1998) to evaluate the goodness of fit of separate one-factor CFA models for the child, teacher, mother, and father data. Matrices of covariances among the subscales of domain-specific self-concept were analyzed (see Table 2 footnote). Multigroup CFA was used to test factorial invariance across the spina bifida and the able-bodied samples. Figure 1 represents the measurement model tested for both samples on all three versions of the Harter (1985) instruments. As seen in this figure, there were five measured variables (the mean of each subscale's items—Scholastic Competence, Social Acceptance, Athletic Competence, Physical Appearance, and Behavioral Control) and one higher order latent construct (self-concept/competence). The path between self-concept and Behavioral Conduct within the model was fixed at 1.0 to define the scale for the variance of the latent factor (see Jöreskog & Sörbom, 1996). Behavioral Conduct was chosen as the fixed variable because it shared the least amount of variance with self-concept compared to the other four subscales when a CFA was computed without fixing any paths.

An important statistical concern in CFA is sample size. At issue is the question of whether the samples at hand provide sufficient power to detect a lack of model fit or to identify differences in parameter estimates across groups. As a rough guideline in establishing target sample size, researchers typically strive to have at least 5 to 10 participants for each estimated parameter in a CFA model (see Kline, 1998). By this criterion, these samples are small but adequate—the one-factor CFA model includes only 10 estimated parameters (i.e., five factor loadings and five unique errors) and yields a 6:1 ratio of participants to estimated parameters. To gauge the adequacy of our sample sizes further, we conducted a retrospective power analysis using procedures outlined by Satorra and Saris (1985; see also Jöreskog & Sörbom, 1996). Assuming that each subscale was 50% reliable, we calculated the smallest between-group difference in factor loadings for which these samples provide 0.80 statistical power (cf. Cohen, 1988). Results indicate that our sample sizes provide 0.80 power to detect absolute differences in standardized loadings that are at least 0.37 in magnitude.

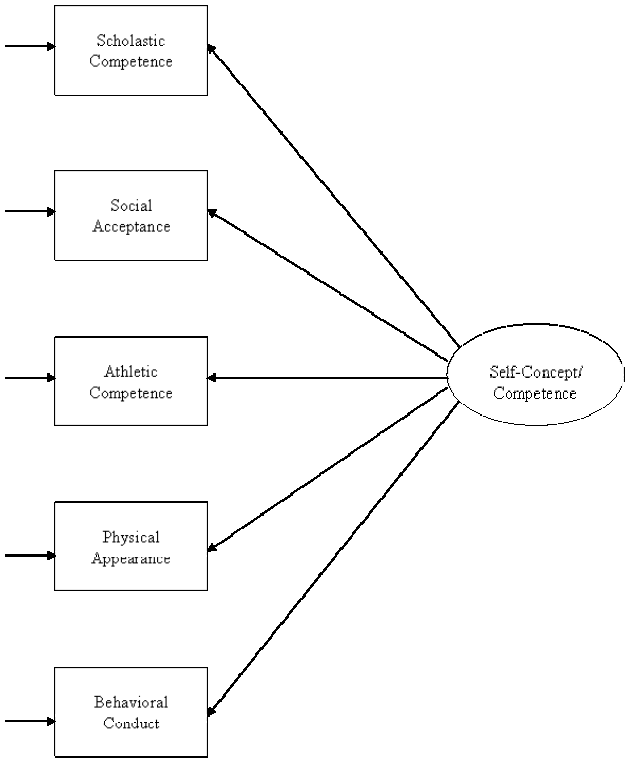

FIGURE 1 Higher order partial aggregation model of self-concept/competence (Harter, 1985) in which the mean of each subscale serves as a measured variable. Rectangles represent measured variables (or first-order, domain-specific subscales of self-concept/competence), and the ellipse represents the latent variable (or the second-order factor of self-concept/competence). Arrow-headed straight lines connecting the latent variable to measured variables represent subscale factor loadings. The small arrow-headed straight lines to each subscale represent unique error variance in each domain-specific subscale that is unrelated to total self-concept/competence.

Assessing Statistical Significance

A sequentially rejective Sidak (1967) Bonferroni-type multiple comparisons procedure was used to ensure an experiment-wise, Type I error rate of $p < .05$ (Sidak, 1967; Soltysik & Yarnold, 1993). With a traditional Bonferroni adjustment, one divides the desired alpha level (e.g., .05) by the total number of statistical comparisons to obtain an adjusted experiment-wise alpha level. With the Sidak procedure, in contrast, one first arranges all obtained p values in order of increasing magnitude and then beginning with the smallest value and moving in sequence to the largest, divides .05 by the number of statistical comparisons remaining to obtain an adjusted alpha level for assessing the significance of each

comparison. Although both procedures control experiment-wise error, the Sidak procedure is slightly less conservative. Effects that met the Sidak-adjusted, experiment-wise criterion were referred to as statistically significant. It is important to note that although the sequentially rejective Sidak Bonferroni-type multiple comparisons procedure protects against Type I errors (i.e., detecting a difference between groups that is not really meaningful), using an adjusted experiment-wise error rate increases the chances of committing Type II errors (i.e., failing to detect a significant difference across groups).

RESULTS

Overview

Our analyses addressed three basic questions about the structure of self-concept and adult-rated child competence in children with spina bifida and able-bodied children: (a) Are the domains of Scholastic Competence, Social Acceptance, Athletic Competence, Physical Appearance, and Behavioral Conduct relevant in defining total self-concept/competence for both groups?; (b) Do each of the five specific domains have the same loadings on the total self-concept/competence factor for both groups?; and (c) Are the error variances for each self-concept/competence domain the same for both groups?

We used single-group CFA to address the first research question concerning the equivalence of factor patterns across groups using four criteria to assess the model's goodness of fit for the set of measured indicators (see Hu & Bentler, 1998). Two of these criteria represent measures of absolute model fit: the root mean square error of approximation (RMSEA; Steiger, 1990) and the goodness-of-fit index (GFI; Jöreskog & Sörbom, 1996). The other two represent measures of relative model fit: the comparative fit index (CFI; Bentler, 1990) and the non-normed fit index (NNFI; Tucker & Lewis, 1973). RMSEA reflects the size of the residuals that result when using the model to predict the data, adjusting for model complexity, with smaller values indicating better fit. According to Browne and Cudeck (1989), RMSEA $\leq$ .05 represents "close fit"; RMSEA between .05 and .08 represents "reasonably close fit"; and RMSEA > .10 represents "an unacceptable model." Analogous to R^2 in multiple regression, GFI reflects the proportion of available variance–covariance information in the data that the given model explains, with larger GFI values representing better model fit. CFI and NNFI indicate how much better the given model fits the data relative to a "null" model, which assumes sampling error alone explains the covariation among observed measures (i.e., that there is no common variance among the items/scales of the instrument). Bentler and Bonett (1980) recommended that measurement models have GFI, CFI, and NNFI values of at least .90. In addition, the goodness-of-fit chi-square (χ^2) statistic was evaluated. Recall in a CFA, it is desirable to have a nonsignificant χ^2, as this indicates that our model does not differ significantly from the model proposed by Harter (1985).

We used multigroup CFA to evaluate the two remaining questions about factor structure. To test for invariance in factor loadings and unique error variances, we contrasted the goodness-of-fit chi-square values of two competing nested models: one constraining the parameters in question (i.e., loadings or unique variances) to be equal for both groups of children and the other omitting this invariance constraint. A nonsignificant difference in chi-square values indicates that the parameters in question are invariant across the two groups, whereas a significant difference in chi-square values indicates that the parameters in question are different for the two groups (Bryant & Yarnold, 1995; Jöreskog & Sörbom, 1996; Kline, 1998). If/when an overall structural difference across groups was found, we used multigroup CFA with equality constraints (Jöreskog & Sörbom, 1996) to pinpoint the source of these group differences.

Are the Same Domain-Specific Factors Relevant Across Samples?

SPPC. The one-factor model provided a reasonable goodness of fit to the SPPC data for both the spina bifida sample, $\chi^2(5, N = 62) = 8.13$, $p \geq .14$; RMSEA = .0995, GFI = .95, CFI = .95, NNFI = .90; and the able-bodied sample, $\chi^2(5, N = 68) = 3.13$, $p \geq .67$; RMSEA = .0000, GFI = .98, CFI = 1.0, NNFI = 1.0, although the model fit the data of the latter group somewhat better than it fit the data of the former group. Table 3 provides a summary of the model fit for all informants.

TRS. The one-factor model provided a reasonable goodness of fit to the TRS data for both the spina bifida sample, $\chi^2(5, N = 65) = 8.79$, $p \geq .11$; RMSEA = .1040, GFI = .95, CFI = .90, NNFI = .81; and the able-bodied sample, $\chi^2(5, N = 56) = 6.50$, $p \geq .26$; RMSEA = .0865, GFI = .95, CFI = .96, NNFI = .92. Similar to the analysis of the SPPC, the model fit the teacher data for the able-bodied sample somewhat better than it fit the teacher data for spina bifida sample.

PRS. Both mothers and fathers completed the PRS. Instead of combining parents' responses into a single model, however, mother and father reports were examined separately. When imposed on the mothers' PRS data for the spina bifida sample, the one-factor model initially yielded an inadmissible solution. Specifically, the Athletic Competence subscale had a negative unique error variance in the LISREL solution (i.e., a Heywood case, due to poor model fit). Con-

TABLE 3
Summary of the Structural Modeling Findings Across Informants for Both the Spina Bifida and Able-Bodied Samples

Scale and Sample	^{2}	*df*	*RMSEA*	*GFI*	*CFI*	*NNFI*
SPPC						
Spina bifida	8.13	5	.0995	.95	0.95	0.90
Able bodied	3.13	5	.0000	.98	1.00	1.00
TRS						
Spina bifida	8.79	5	.1040	.95	0.90	0.81
Able bodied	6.50	5	.0865	.95	0.96	0.92
PRS (mother)						
Spina bifida	10.66	5	.1040	.94	0.75	0.51
Able bodied	9.34	5	.1100	.95	0.84	0.68
PRS (father)						
Spina bifida	—	—	—	—	—	—
Able bodied	4.92	5	.0000	.97	1.00	1.00

Note. Due to an uncorrectable inadmissibility problem, fit statistics were not calculated for the PRS scores of fathers of children with spina bifida. RMSEA = root mean square error of approximation; GFI = goodness-of-fit index; CFI = comparative fit indx; NNFI = nonnormed fit index; SPPC = Self-Perception Profile for Children; TRS = Teacher's Rating Scale of Child's Actual Behavior; PRS = Parent's Rating Scale of Child's Actual Behavior.

straining the offending unique variance estimate to be nonnegative (see Wothke, 1993) solved the admissibility problem but yielded a relatively poor model fit to the PRS data for the spina bifida sample, $\chi^2(5, N = 66) = 10.66$, $p \geq .05$; RMSEA = .1040, GFI = .94, CFI = .75, NNFI = .51. Likewise, the one-factor model provided a relatively poor fit to the PRS data for the able-bodied sample, $\chi^2(5, N = 68) = 9.34$, $p \geq .09$; RMSEA = .1100, GFI = .95, CFI = .84, NNFI = .68. The pattern of high absolute fit but poor relative fit evidenced by mothers from both samples suggests that the hierarchical, one-factor model explains the available variance–covariance information well but that there is relatively little common variance to explain in the first place (see intercorrelations among the subscales for mothers' PRS data in Table 2). To support this conclusion, median subscale intercorrelations were compared across the SPPC and the PRS. Such intercorrelations for the SPPC versus the mothers' PRS were .34 versus .18 for the spina bifida sample and .36 versus .23 for the able-bodied sample.

Similar to the fit obtained for mothers in the spina bifida sample, when the one-factor model was imposed on the fathers' PRS data for the spina bifida sample, the one-factor model initially yielded an inadmissible solution (i.e., a Heywood case, due to poor model fit). For the fathers' data, the Physical Attractiveness subscale had a negative unique error variance in the LISREL solution. Constraining the offending unique variance estimate to be nonnegative (see Wothke, 1993), however, did not solve the admissibility problem for the spina bifida sample. Interestingly, in contrast to the previous finding for mothers, the model fit the fathers' PRS data well for the able-bodied sample, $\chi^2(5, N = 52) = 4.92$, $p \geq .42$; RMSEA = .0000, GFI = .97, CFI = 1.0, NNFI = 1.0. As was the case for mothers' PRS data, however, median subscale intercorrelations were higher for the SPPC data than for father's PRS data (i.e., .34 vs. .14 for the spina bifida sample and .36 vs. .25 for the able-bodied sample). Thus, across both samples, the proposed model seems less applicable to the parent-report data than to the child-report or teacher-report data. Given the tendency toward poor fit for the maternal and paternal PRS data (with the only exception emerging for fathers in the able-bodied sample), no further tests of factorial invariance were conducted for the PRS data.

Are the Factor Loadings the Same Across Samples?

SPPC. Multigroup CFA revealed that the magnitudes of the factor loading for the hierarchical, one-factor SPPC model were invariant across samples, $\Delta\chi^2(4, N = 130) = 8.49$, $p \geq .07$.

TRS. Unlike results for child-report data, multigroup CFA revealed that the magnitudes of the factor loadings for the one-factor TRS model were different for the spina bifida and able-bodied samples, $\Delta\chi^2(4, N = 121) = 12.15$, $p \leq .02$. Follow-up analyses using equality constraints revealed that only one of the domain-specific subscales had loadings that were different across the two samples—specifically, teacher-rated total child competence tended to have more to do with Scholastic Competence for the spina bifida sample than for the able-bodied sample, $\Delta\chi^2(1, N = 121) = 4.13$, $p \leq .05$.

Are the Unique Variances in Domain-Specific Self-Concepts the Same Across Samples?

SPPC. Multigroup CFA revealed that the unique variances in the domain-specific subscales for the hierarchical, one-factor SPPC model were nonequivalent across groups, $\Delta\chi^2(5, N = 130) = 19.07$, $p \leq .01$. Follow-up analyses using equality constraints revealed that both Athletic Competence, $\Delta\chi^2(1, N = 130) = 7.80$, $p \leq .01$, and Behavioral Conduct, $\Delta\chi^2(1, N = 130) = 5.32$, $p \leq .03$, had more unique error variance for the spina bifida sample than for the able-bodied sample.

TRS. Likewise, multigroup CFA revealed that the unique variances in the domain-specific subscales for the hierarchical, one-factor TRS model were nonequivalent across groups, $\Delta\chi^2(5, N = 121) = 25.27$, $p \leq .01$. Follow-up analyses using equality constraints revealed that more of the variance in Physical Appearance, $\Delta\chi^2(1, N = 121) = 11.98$, $p \leq .01$, was unrelated to total child competence in the spina bifida sample compared to the able-bodied group.

DISCUSSION

The purpose of this study was to determine whether the factor structure of Harter's (1985) measures of self-concept (i.e., the SPPC) and adult-rated child competence (i.e., the TRS and PRS) is the same across samples of children with and without spina bifida. Recall that Harter (1985) proposed that five first-order variables (Scholastic Competence, Social Acceptance, Athletic Competence, Physical Appearance, and Behavioral Conduct) were influenced by a single, higher order latent variable (self-concept/child competence). It was predicted that the Athletic Competence and Physical Appearance subscales would load more highly on self-concept/competence in the spina bifida sample when compared to the able-bodied sample. Although this specific structural hypothesis was not supported, findings revealed that the one-factor model fit the data reasonably well for the SPPC and the TRS.

Factorial validity, however, was not unanimously established for the PRS structure proposed by Harter (1985) when both mothers and fathers completed the instrument. Only the structure of the fathers PRS was validated with the able-bodied sample. Factorial validity for the other versions of the instrument was not established due to the instrument's relatively poor fit for the spina bifida sample for both mother and father report and the able-bodied group for mother report. One possible explanation of this poor fit is that parents may have been unclear as to whom they should compare their children with when answering questions on the PRS (particularly for parents of children with spina bifida). Whereas some parents may have compared their child to other children with spina bifida, other parents may have compared their child to able-bodied classmates. These differences in comparison groups could impact parents' rating of their children's competence and produce considerable "noise" in the data. For example, a child with spina bifida may be perceived as more athletically competent in comparison to other children with spina bifida (e.g., the child may play on a wheelchair basketball team) but regarded as less athletically competent when a comparison is made to able-bodied children (e.g., the child may not engage in other sports, such as track or baseball). This explanation, however, does not explain the low correlations among the subscales of the instrument (unless parents are inconsistent in their use of a comparison group across the different subscale items). Thus, a more plausible explanation may be that the competencies tapped by the PRS do not cohere as part of a second-order, total child competence construct for parent respondents. To help resolve this issue, future analyses with larger sample sizes should be conducted at the item level rather than at the scale level. Moreover, it may be that more items need to be generated for the adult version of this scale, given the possibility that three-item subscales are more likely to be psychometrically unstable than scales with greater numbers of items.

Are the Factor Loadings the Same Across Groups?

Further tests of factorial invariance were conducted on the two instruments for which factorial validity was established (the SPPC and the TRS). For both of these instruments, the spina bifida and able-bodied samples tended to have the same pattern and magnitude of factor loadings for each of the five self-concept/competence subscales. These findings parallel other studies that have examined the equivalence of factor structures across special populations. For example, Veerman et al. (1996) established factorial invariance of the SPPC across clinic and nonclinic samples.

Are the Unique Variances in Domain-Specific Self-Concepts the Same Across Samples?

The final step in establishing factorial invariance involved comparing the unique error variances for each of the self-concept/competence subscales. These results reveal that the Athletic Competence and Behavioral Conduct subscales had more unique error variance on the SPPC, and Physical Appearance had more unique error variance on the TRS, for the spina bifida sample compared to the able-bodied sample. In other words, the Athletic Competence and Behavioral Conduct subscales had less to do with child-rated self-concept, and Physical Appearance had less to do with teacher-rated child competence, for the spina bifida sample. These findings run contrary to the prediction that there would be less unique variance in the Athletic Competence and Physical Appearance domains among children with spina bifida compared to children without spina bifida.

These results indicate that there are factors beyond children's self-rated self-concept and teachers' rated child competence that influence these subscales. Other possible influences on children's Athletic Competence and Behavioral Conduct subscales may be the limited opportunities for children with spina bifida to engage in athletic activities or be disruptive, child personality factors, or level of depression (e.g., children with spina bifida report higher levels of depression than their able-bodied peers; Appleton et al., 1997; Holmbeck et al., in press). A primary additional influence on the teachers' Physical Appearance subscale may be teachers' desire to report socially appropriate responses to items referring to children's physical attractiveness, which was (anecdotally) reflected in teachers' reluctance to answer items related to students' physical appearance (particularly for the spina bifida sample).

GENERAL CONCLUSIONS

Considered together, these results suggest that child-rated self-concept and teacher-rated child competence as assessed with the SPPC and TRS, respectively, have largely

the same factor structure for children with and without spina bifida. Thus, we can be reasonably confident in using Harter's (1985) SPPC and TRS to compare children with and without this particular pediatric condition. Importantly, these results are consistent with previous findings of factorial invariance across other groups of children (e.g., clinic vs. nonclinic samples; Veerman et al., 1996). Irrespective of the results of this study, however, evaluation of the Harter instruments with pediatric populations should continue. Further testing with larger samples is necessary to cross-validate these findings and increase confidence in our conclusions regarding the factorial invariance of these instruments. Because of the small sample size in this study, we used partial aggregation models to assess factorial invariance. A larger sample size would enable tests of factorial invariance at the item level. The focus on children with spina bifida also limits the degree to which we can generalize to other populations of interest. Moreover, this study was unable to validate the factor structure of Harter's PRS. This finding could be related to the differential use of reference groups among parents when completing the instrument. Although the parent and teacher versions of the Harter instrument do not directly request comparisons of the target child to other children, such comparisons are inherent to this type of survey methodology. In the case of children with physical disabilities, it may be more difficult to complete such a measure (e.g., Does a parent/teacher assess a child's competence in comparison to all other children that age or does a parent/teacher compare a child only to other children with disabilities?). In defense of using this instrument, however, we point out that most children with spina bifida are mainstreamed in regular school programs, thus increasing the likelihood that comparisons are made to a target child's peers rather than only to other children with physical disabilities. Nonetheless, further research should be conducted to evaluate the impact of varying the reference group on parent and teacher responses to surveys such as the PRS and TRS.

ACKNOWLEDGMENTS

Completion of this manuscript was supported by Social and Behavioral Sciences Research Grants 12–FY93–0621, 12–FY95–0496, 12–FY97–0270, 12–FY99–0280, and 12–FY01–0098 from the March of Dimes Birth Defects Foundation and a Research Support Grant and Paid-Leave from Loyola University of Chicago. We thank Ann Walsh Johnson, Joy Ito, Barb Kania, Pat McGovern, Pat Braun, Caroline Anderson, David McLone, John Lubicky, the Illinois Spina Bifida Association, and the staff of the spina bifida clinics at Children's Memorial Hospital, Shriner's Hospital–Chicago, and Loyola University of Chicago Medical Center. We also thank numerous undergraduate and graduate students for their help with data collection and data entry. Most important, we thank the parents, children, teachers, and health professionals who participated in our study.

REFERENCES

Appleton, P. L., Ellis, N. C., Minchom, P. E., Lawson, V., Boll, V., & Jones, P. (1997). Depressive symptoms and self-concept in young people with spina bifida. *Journal of Pediatric Psychology, 22,* 707–722.

Bagozzi, R. P., & Edwards, J. R. (1998). A general approach for representing constructs in organizational research. *Organizational Research Methods, 1,* 45–87.

Bagozzi, R. P., & Heatherton, T. F. (1994). A general approach to representing multifaceted personality constructs: Application to state self-esteem. *Structural Equation Modeling, 1,* 35–67.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238–246.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88,* 588–606.

Browne, M. W., & Cudeck, R. (1989). Single-sample cross-validation indices for covariance structures. *Multivariate Behavioral Research, 24,* 445–455.

Bryant, F. B., & Yarnold, P. R. (1995). Principal-components analysis and exploratory and confirmatory factor analysis. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 99–136). Washington, DC: American Psychological Association.

Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Cohen, J. (1988). *Statistical power for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Cole, D. A., Gondoli, D. M., & Peeke, L. G. (1998). Structure and validity of parent and teacher perceptions of children's competence: A multitrait-multigroup investigation. *Psychological Assessment, 10,* 241–249.

Demo, D. H. (1992). The self-concept over time: Research issues and directions. *Annual Review of Sociology, 18,* 303–326.

Granleese, J., & Joseph, S. (1993). Factor analysis of the Self-Perception Profile for Children. *Personality and Individual Differences, 15,* 343–345.

Granleese, J., & Joseph, S. (1994). Reliability of the Harter Self-Perception Profile for Children and predictors of global self-worth. *The Journal of Genetic Psychology, 155,* 487–492.

Harter, S. (1985). *Manual for the Self-Perception Profile for Children.* Denver, CO: University of Denver.

Harter, S. (1998). The development of self-representations. In W. Damon (Series Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (5th ed., pp. 553–617). New York: Wiley.

Hollingshead, A. A. (1975). *Four factor index of social status.* Unpublished manuscript, Yale University, New Haven, CT.

Holmbeck, G. N., Belvedere, M. C., Christiansen, M., Czerwinski, A. M., Hommeyer, J. S., Johnson, S. Z., et al. (1998). Assessment of adherence with multiple informants in pre-adolescents with spina bifida: Initial development of a multidimensional, multitask parent-report questionnaire. *Journal of Personality Assessment, 70,* 427–440.

Holmbeck, G. N., Gorey-Ferguson, L., Hudson, T., Seefeldt, T., Shapera, W., & Uhler, J. (1997). Maternal, paternal, and marital functioning in families of preadolescents with spina bifida. *Journal of Pediatric Psychology, 22,* 167–181.

Holmbeck, G. N., Johnson, S. Z., Wills, K. E., McKernon, W., Rose, B., Erklin, S., et al. (2002). Observed and perceived parental overprotection in relation to psychosocial adjustment in preadolescents with a physical

disability: The mediational role of behavioral autonomy. *Journal of Consulting and Clinical Psychology, 70,* 96–110.

Holmbeck, G. N., Shapera, W., & Hommeyer, J. S. (2002). Observed and perceived parenting behaviors and psychosocial adjustment in pre-adolescents with spina bifida. In B. K. Barber (Ed.), *Intrusive parenting: How psychological control affects children and adolescents* (pp. 191–234). Washington, DC: American Psychological Association.

Holmbeck, G. N., Westhoven, V. C., Phillips, W. S, Bowers, R., Gruse, C., Nikolopoulos, T., et al. (in press). A multi-method, multi-informant, and multi-dimensional perspective on psychosocial adjustment in pre-adolescents with spina bifida. *Journal of Consulting and Clinical Psychology.*

Hu, L.-T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3,* 424–453.

Hull, J. G., Lehn, D. A., & Tedlie, J. C. (1991). A general approach to testing multifaceted personality constructs. *Journal of Personality and Social Psychology, 61,* 932–945.

Jöreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 294–316). Newbury Park, CA: Sage.

Jöreskog, K. G., & Sörbom, D. G. (1996). *LISREL 8: User's reference guide.* Chicago: Scientific Software.

Kline, R. B. (1998). *Principles and practice of structural equation modeling.* New York: Guilford.

MacCallum, R. C., & Tucker, L. R. (1991). Representing sources of error in common factor analysis: Implications for theory and practice. *Psychological Bulletin, 109,* 501–511.

Marsh, H. W., & Hocevar, D. (1985). The application of confirmatory factor analysis to the study of self-concept: First and higher order factor structures and their invariance across age groups. *Psychological Bulletin, 97,* 562–582.

McKernon, W. L., Holmbeck, G. N., Colder, C. R., Hommeyer, J. S., Shapera, W., & Westhoven, V. (2001). Longitudinal study of observed and perceived family influences on problem-focused coping behaviors of preadolescents with spina bifida. *Journal of Pediatric Psychology, 26,* 41–54.

Piers, E. V., & Harris, D. B. (1964). Age and other correlates of self-concept in children. *Journal of Educational Psychology, 55,* 91–95.

Satorra, A., & Saris, W. E. (1985). Power of the likelihood ratio test in covariance structure analysis. *Psychometrika, 50,* 83–90.

Schumann, B. C., Striegel-Moore, R. H., McMahon, R. P., Waclawiw, M. A., Morrison, J. A., & Schreiber, G. B. (1999). Psychometric properties of the Self-Perception Profile for Children in a biracial cohort of adolescent girls: The NHLBI growth and health study. *Journal of Personality Assessment, 73,* 260–275.

Sidak, Z. (1967). Rectangular confidence regions for the means of multivariate normal distributions. *Journal of the American Statistical Association, 62,* 626–633.

Soltysik, R. C., & Yarnold, P. R. (1993). *ODA 1.0: Optimal data analysis for DOS.* Chicago: Optimal Data Analysis.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25,* 173–180.

Trent, L. M. Y., Russell, G., & Cooney, G. (1994). Assessment of self-concept in early adolescence. *Australian Journal of Psychology, 46,* 21–28.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum-likelihood factor analysis. *Psychometrika, 38,* 1–10.

Veerman, J. W., ten Brink, L. T., Straathof, M. A. E., & Treffers, P. D. A. (1996). Measuring children's self-concept with a Dutch version of the "Self-Perception Profile for Children": Factorial validity and invariance across a nonclinic and a clinic group. *Journal of Personality Assessment, 67,* 142–154.

Wothke, W. (1993). Nonpositive definite matrices in structural modeling. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 256–293). Newbury Park, CA: Sage.

Grayson N. Holmbeck
Loyola University of Chicago
Department of Psychology
6525 North Sheridan Road
Chicago, IL 60626
E-mail: gholmbe@luc.edu

Received May 3, 2002
Revised February 22, 2003