

# The Structure of the School Social Behavior Scales: A Confirmatory Factor Analysis

---

Susan L. Crowley  
*Utah State University*

Kenneth W. Merrell  
*University of Oregon*

The School Social Behavior Scales (SSBS; Merrell 1993a) is a behavior rating instrument designed for use by teachers and other school personnel in evaluating social and antisocial behavior of children in grades K through 12. Published psychometric data for the SSBS have been promising to date. Although exploratory factor-analytic techniques were used to define its subscales, a confirmatory methodology has not been used with this measure. The present research sought to confirm the structure of the SSBS using the national data collected for the standardization and norming of the instrument. The standardization sample was split to allow for post-hoc model fitting, if necessary. The initial model tested for both the Social Competence and Antisocial Behavior scales was in line with the scoring recommendations of the author, and results suggested some model mis-specification. Alternative models were tested to identify a model that would provide a better fit to the data. After identification of an acceptable model, the invariance of the new model was tested with the remaining half of the standardization sample. Generally, both scales performed well and resulted in models very similar to those originally proposed by the author of the instrument. These findings support the continued use of the SSBS for both clinical and research purposes. Results are discussed from the perspectives of screening, assessment, and intervention planning with children and youth in K-12 educational settings.

The social behavior of children and youth has always been a topic of great interest for educators. However, there has probably never been a time in the history of American education where more attention has been paid to this topic than the present. More than ever educators are now being asked to assume the role of protector, socializing agent, and caregiver for the students they serve (Walker, Colvin, & Ramsey, 1995).

One of the major reasons cited for the current focus on social skills is that when families are ineffective in their socialization responsibilities, schools must deal with increased social skill deficits and higher levels of antisocial behavior among students. But as noted by Flanagan, Alfonso, Primavera, Povall, and Higgins (1996), the increased attention to social behavior in school settings may be traced to several additional influences. For example, important changes are occurring in this area due to the expansion of the learning

disability definition to include social skill deficits and inclusion of a social competence component in the definition of mental retardation. In addition, intervention programs and rating instruments designed specifically to measure social functioning have improved. The inclusion of social skill deficits in disability language potentially presents teachers with the need to target remediation efforts as an additional content area.

Another reason for concern about the social behavior of children and youth is the growing awareness that early social deficits may lead to significant social problems later in life. Patterson, DeBaryshe and Ramsey (1989) proposed a developmental model of antisocial behavior, arguing that antisocial behavior appears to be a developmental trait beginning early in life and often continuing into adolescence and adulthood. In addition, they cited numerous studies showing that antisocial behavior in children contributes to

the later development of alcoholism, divorce, and psychiatric illness, among other maladaptive social behaviors. Models such as those proposed by Patterson et al. are consistent with the notion of developmental pathways, which has emerged as an important concept in recent years (e.g., Loeber, 1991; Loeber et al., 1993; Nurcombe et al., 1991; Rubin, Hymel, Mills, & Rose-Krasnor, 1991). Basically, a developmental pathway is a particular pattern of behavior that increases the probability that other behaviors in the "pathway" will occur in the future. For example, infants and toddlers who are fussy, irritable, noncompliant, and who engage in excessive attention-seeking behaviors are at increased risk for exhibiting aggressive behavior and other conduct problems when they enter the school setting in preschool or kindergarten. Thus, these difficult behaviors early in life might be considered a developmental pathway to future problem behaviors in the school setting, which may in turn lead to adjustment problems later in life.

In responding to the various reasons for concern about the social behavior of their students, it is critical that teachers, school psychologists, counselors, and professionals in related fields use appropriate and effective social behavior assessment measures for screening and evaluating the social and antisocial behavior of students. Such tools provide a basis for effective interventions, and are also essential for other important tasks, such as early identification, determining service eligibility, and gauging the progress of students. For screening and assessing social behavior in school settings, it has been suggested that the most ecologically valid and clinically useful assessment methods include direct behavioral observation, interviewing, sociometric techniques, and the use of behavior rating scales (Merrell, 2000). Other methods, such as self-report measures and projective-expressive techniques, appear to have less utility for assessing social behavior. Each method of assessment has specific advantages

and limitations, which have been addressed previously in many sources (e.g., Martin, 1988; Merrell, 1999).

The focus of this investigation was on behavior rating scales, a popular and cost-effective way to screen and assess the social-emotional behavior of children and youth. More specifically, within the general domain of child behavior rating scales, the School Social Behavior Scales (SSBS; Merrell 1993a) was our specific focus. The SSBS was designed for use by teachers and other school personnel in evaluating social and antisocial behavior of children in grades K through 12. The SSBS was designed for the following specific purposes:

- as a screening tool for early identification of behaviorally at risk students
- as part of a multi-source, multi-method assessment for classification and determination of special program eligibility
- for use in developing intervention plans and to provide initial descriptive information relevant for planning functional behavior assessments
- as an evaluation tool to monitor behavior change during and following intervention

Psychometric evidence for the SSBS (which is reviewed in more detail in the Method section of this article) has been encouraging, and the instrument has generally been positively reviewed in the professional literature (e.g., Demaray, Ruffalo, & Carlson, 1995; Hooper, 1998; Kreisler, Mangione, & Landau, 1997; Welsh, 1998). However, instrument validation and construct refinement is an ongoing process rather than something that may be considered to be "finished" once an instrument and a few supporting studies have been published (e.g., Gregory, 1996). It is incumbent on instrument developers and measurement researchers to continue to evaluate the psychometric characteristics of instruments, so that a larger body of evidence may accrue. In turn, this enlarged body of evidence should

demonstrate in more detail the validity of measures for specific purposes, as well as the integrity of their internal structure and properties. Thus, this article focuses on some specific aspects of the continued refinement and validation of the SSBS.

The SSBS items were developed using a rational-theoretical approach to item inclusion consistent with contemporary theories of social competence (i.e., Merrell & Gimpel, 1998) and antisocial behavior (i.e., Loeber et al., 1993; Walker et al., 1995). Each of the two major scales consists of three factor or subscale scores areas. For example, Scale A, Social Competence, includes 32 items that describe adaptive, prosocial behavioral competencies. Subscale A1 (Interpersonal Skills) includes 14 items reflecting social skills that are important in establishing positive relationships with and gaining social acceptance from peers (e.g., "Offers help to other students when needed" and "Interacts with a wide variety of peers"). Subscale A2 (Self-Management Skills) includes 10 items related to social skills involving self-restraint, cooperation, and compliance with the demands of school rules and expectations (e.g., "Responds appropriately when corrected by teacher" and "Shows self-restraint"). Subscale A3 (Academic Skills) consists of eight items relating to competent performance and engagement on academic tasks (e.g., "Completes individual seatwork without being prompted" and "Completes assigned activities on time"). In contrast to the positive social competence items included in Scale A, Scale B, Antisocial Behavior, includes 33 items describing problematic behaviors that are either other-directed in nature, or are likely to lead to negative social consequences such as peer rejection or strained relationships with the teacher. Subscale B1 (Hostile-Irritable) consists of 14 items that describe behaviors considered to be self-centered and annoying, and that will therefore likely lead to peer rejection (e.g., "Will not share with other students" and

"Argues and quarrels with other students"). Subscale B2 (Antisocial-Aggressive) consists of 10 behavioral descriptors relating to overt violation of school rules and intimidation or harm to others (e.g., "Gets into fights" and "Takes things that are not his/hers"). Finally, Subscale B3 (Disruptive-Demanding) includes nine items that reflect behaviors likely to disrupt ongoing school activities and place excessive and inappropriate demands on others (e.g., "Is overly demanding of teacher's attention" and "Is difficult to control"). The SSBS is somewhat unique among behavior rating scales in focusing exclusively on *social* behavior and including separate comprehensive screens for both adaptive social behavior and antisocial behavior.

Although exploratory factor-analytic techniques were used to define the SSBS subscales, a confirmatory methodology was not originally used in scale construction, nor in subsequent published research. However, various researchers (e.g., Crowley & Fan, 1997) have described confirmatory factor analysis as ideal for use in instrument development and related measurement situations. Whereas exploratory procedures are recommended during the early stages of instrument development, confirmatory procedures may be best used later when a specific structure for the instrument has been identified. As noted by Pedhazur and Schmelkin (1991), confirmatory factor-analytic procedures are "eminently suited for internal and cross-structure analysis in the process of construct validation" (p. 632). Given the importance of this type of analysis and the fact that it had not yet been applied with the SSBS, the present investigation sought to confirm, using confirmatory factor analysis, the structure of the SSBS as proposed by the author of the instrument.

## Method

### Sample

The data used for the present analysis were the national standardization sample of the

School Social Behavior Scales, as well as additional normative data collected after the instrument was published. These data included teacher SSBS ratings of 2,047 students in grades K-12, or ages 6 to 18 (mean age = 11.45,  $SD = 3.53$ ). The sample was collected from public schools across the United States from 1990 to 1993. Teachers at the participating data collection sites were asked to randomly select three students from their class rosters, and to rate them in an anonymous manner using the SSBS. The sample consisted of 57% boys and 43% girls, and includes participant representation from each of the four U.S. geographical regions. Eighty-two percent of the sample was European American (Caucasian); the remainder was made up of members of various ethnic minority groups. The sample is well stratified with respect to socioeconomic status, special education participation, and other important demographic characteristics, which are presented in detail in the SSBS user's guide.

### *Instrument*

The SSBS includes two separate scales (Social Competence and Antisocial Behavior) with a total of 65 items that describe both positive and negative social behaviors that commonly occur in school settings. Items are rated using a 5-point scale with three anchors, "never," "sometimes," and "frequently." Each of the two scales yields a total score, which may be converted to standard scores and percentile ranks for normative comparisons. The two scales each have three empirically derived (through exploratory factor analyses) subscales. Raw subscale scores are converted to one of four Social Functioning Levels, "High Functioning," "Average," "Moderate Problem/Deficit," and "Significant Problem/Deficit." These levels were devised by setting cutoff points that identify ratings of children and youth whose SSBS scores are either in the lowest 5% of Social Competence total scores or the highest

5% of Antisocial Behavior total scores, based on the national norming sample.

A number of studies and procedures concerning the psychometric properties and validity of the instrument are reported in the SSBS manual and subsequent publications. Briefly, the scales have strong internal consistency (coefficient alpha) and split-half reliability (.91 to .98). Test-retest reliability at three-week intervals has been reported at .76 to .83 for the Social Competence scores, and .60 to .73 for the Antisocial Behavior scores. Interrater reliability between resource room teachers and paraprofessional aides was found to range from .72 to .83 for the Social Competence scores, and from .53 to .71 for the Antisocial Behavior scores (Merrell, 1992, 1993b).

Validity of the scales has been demonstrated in several ways, including statistically significant convergent and discriminant validity with other behavior rating scales (Emerson, Crowley, & Merrell, 1994; Merrell, 1993b); evidence of strong sensitivity to theoretically based group differences (Merrell, 1992, 1993b; Merrell & Gill, 1994; Merrell, Sanders, & Poppinga, 1993; Robbins & Merrell, 1998); convergence with other types of assessment such as sociometric procedures, self-report instruments, and direct behavioral observation (Merrell, 1993b; Merrell, Cedeno, & Johnson, 1993); and strong classification accuracy of students from various special education and clinical groups (Merrell, 1992, 1993b; Merrell, Sanders, & Poppinga, 1993). Additionally, a parent report version of the SSBS, the Home and Community Social Behavior Scales, was recently developed (Merrell, 2002).

The SSBS has been positively reviewed in the professional literature (Demaray et al., 1995; Elksnin & Elksnin, 1995; Hooper, 1998; Kreisler et al., 1997; Welsh, 1998). However, the absence of verifying evidence for testing the three-subscale factor structure of each of the two SSBS scales through confirmatory factor analysis or structural model-

ing techniques has been noted as an area in which the SSBS is lacking.

### Results

For the present study, the two SSBS scales, Social Competence and Antisocial Behavior, were considered in separate analyses in line with the author's recommended use (the scales are derived and scored separately). Analyses were conducted using Lisrel 8.3 (Joreskog & Sorbom, 1996). Analyses were conducted on a correlation matrix of the observed variables, which are available from the author upon request. In all cases, the initial model tested was a first-order model with the items loading on the requisite latent variable (instrument subscale). Initially, no error terms were freed to correlate and the metric of the latent variables was constrained to be 1.0. If the model did not provide an acceptable fit to the data, post-hoc model fitting was used to improve the fit between the model and the data. Because of the problems inherent in post-hoc model fitting, half of the standardization sample was held in reserve to be used for cross-validation of the final model identified.

Several authors have noted the difficulty in conducting confirmatory factor analyses with item-level data (see Floyd & Widaman, 1995, for a discussion). Briefly, while exploratory factor analyses seek to identify what variance can be accounted for, confirmatory procedures focus more on the amount of variance that cannot be accounted for. Thus, the unreliability of the observed, or measured, variables becomes even more important. Individual items are quite unreliable, making confirmatory procedures difficult. Gorsuch (1997) outlined four major limitations of item-level data, compared to scales, in analyses. First, as previously stated, items have lower reliabilities than scales (which, he notes, is the reason why scales are developed to begin with). Second, there are often additional sources of confounding variance in

items besides the construct purportedly being measured. Third, the distribution of each item score varies, which subsequently reduces the relationship between items. Fourth, items are generally ordinal-level data. The limited number of possible responses artificially lowers the measured correlations between items.

To deal with the difficulties involved with analyses of item-level data, several authors have recommended what Gorsuch calls (1997) "indirect item analyses" (Collins & Gleaves, 1998; Comrey & Lee, 1992). In this approach, item packets, testlets, or mini-scales are used as the observed variables. Testlets are a combination of several items (e.g., two to four), creating a mini-scale that is more reliable than the individual items. For instance, rather than having 14 items as predictors of a subscale, 4 testlets would act as predictors. According to Gorsuch, this procedure has been used for 40 years, beginning with Cattell's early work on the 16PF. Based on the compelling arguments identifying the limitations of item-level data analyses, the item packet approach was used for the present analyses. In all cases, items were combined based on general content as judged by the authors to create three or four testlets for each subscale (Comrey & Lee, 1992; Gorsuch, 1997).

#### *Scale A: Social Competence*

Testlets created for the Social Competence subscales are presented in Table 1. As previously mentioned, analyses for the initial model were conducted on the odd cases of an even-odd split of the standardization sample. A plethora of fit statistics have been created over the past two decades to assess model fit. Byrne for example, recommends (1998) a subset of fit indices that use various approaches to model fit and have received strong support in the literature. Specifically, these indices include the  $\chi^2$ , root mean square error of approximation (RMSEA), expected cross-validation index (ECVI), standardized root

**Table 1** Items Packets for Scale A, Social Competence

Interpersonal 1	<p>offers help to other students when needed</p> <p>understands other students' problems and needs</p> <p>is sensitive to feelings of other students</p>
Interpersonal 2	<p>effectively participates in group discussions and activities</p> <p>appropriately enters ongoing activities with peers</p> <p>is appropriately assertive when he/she needs to be</p>
Interpersonal 3	<p>has skills or abilities that are admired by peers</p> <p>has good leadership skills</p> <p>is sought out by peers to join activities</p> <p>is "looked up to" or respected by peers</p>
Interpersonal 4	<p>listens to and carries out directions from teacher</p> <p>interacts with a wide variety of peers</p> <p>is skillful at initiating or joining conversations with peers</p> <p>compliments others' attributes or accomplishments</p>
Self-Management 1	<p>cooperates with other students in a variety of situations</p> <p>is accepting of other students</p> <p>will compromise with peers when appropriate</p>
Self-Management 2	<p>remains calm when problems arise</p> <p>controls temper when angry</p> <p>shows self-restraint</p>
Self-Management 3	<p>follows classroom rules</p> <p>behaves appropriately in a variety of school settings</p> <p>adjusts to different behavioral expectations across school settings</p>
Academic 1	<p>completes individual seatwork without being prompted</p> <p>accomplishes assignments and other tasks independently</p> <p>produces work of acceptable quality for his/her ability level</p>
Academic 2	<p>listens to and carries out directions from teacher</p> <p>asks for clarification of instructions in an appropriate manner</p> <p>appropriately asks for assistance as needed</p>
Academic 3	<p>appropriately transitions between classroom activities</p> <p>completes assigned activities on time</p>

mean residual (RMR), goodness of fit (GFI), adjusted goodness of fit (AGFI), and the comparative fit index (CFI). Fit statistics for the initial Scale A model presented in Table 2 were as follows:  $\chi^2 = 844.18$  with 32 df ( $p < .0001$ ), RMSEA = .15, EVCI = .81, RMR = .056, GFI = .87, AGFI = .77, and CFI = .94.

For interpretive purposes, goodness of fit indices (e.g., GFI, CFI) above .90 are generally considered to provide an "acceptable" fit of the model to the data. Conversely, a low RMR, generally below .05, is suggestive of an adequate fit. Similarly, a low RMSEA, usually below .10, is indicative of an adequate fit of the model to the data. The EVCI assesses the likelihood that the model in question will cross validate with another, similar sample. Theoretically, the model with the lowest EVCI is most likely to replicate. The EVCI for the model in question can be compared to the ECVI for an independence model (assuming independence of all parameters, zero correlation between variables) and a saturation model (a minimally restrictive, just-identified model). Finally, the AGFI is one of a family of fit statistics that considers the number of

parameters to be estimated and, therefore, considers the parsimony of the model.

A review of the fit statistics for our initial model suggests some misspecification. The  $\chi^2$  value is, not surprisingly for the sample size, statistically significant. The RMSEA and RMR are close to desired values, but suggest misspecification. Similarly, the GFI and AGFI are below the desired .90. The EVCI value for our current model, .81, fell between that of the independence model, 12.88, and that of the saturated model, .11. These results suggest that freeing additional parameters could improve the likelihood of cross validation. Of course, the decision to free parameters must be made on substantive as well as statistical criterion. Only the CFI suggests an acceptable fit between the model and the data. Based on these findings, subsequent analyses were conducted on the first-order model to improve model fit

A series of three additional models were tested to identify the "best" fit. In Model 1, the error terms for mini-scales Academic 1 and Academic 3 were allowed to correlate. Allowing error terms to correlate has been the

**Table 2 Summary Statistics for Scale A Models**

	Author's	Post-hoc model fitting			Cross-validation	
	initial	Model 1	Model 2	Model 3	pattern	values
change		TD(10,8)	LX(5,1)	TD(5,1)		
$\chi^{2a}$	844.18	622.31	485.88	389.55	288.58	295.98
df	32	31	30	29	29	37
$\Delta \chi^2$		221.87	136.43	96.33		
RMSEA	0.15	0.13	0.12	0.11	0.088	0.078
ECVI	0.81	0.6	0.53	0.41	0.31	0.3
RMR	0.056	0.042	0.03	0.027	0.023	0.026
GFI	0.87	0.9	0.91	0.93	0.95	0.95
AGFI	0.77	0.82	0.84	0.87	0.91	0.93
CFI	0.94	0.95	0.97	0.97	0.98	0.98

<sup>a</sup> All  $\chi^2$  values are statistically significant at  $p < .001$ .

subject of much debate in the literature. For example, Byrne (1998) stated that correlated errors are not uncommon in social psychological research, these measurement "error covariances represent systematic, rather than random, measurement error in item responses and may derive from characteristics specific either to the items, or to the respondents . . . . Another type of method effect that can trigger correlated errors is a high degree of overlap in item content" (p.147). A review of the item content in academic test packets 1 and 3 suggests that the latter is likely the case. In fact, the two mini-scales are highly correlated ( $r = .91$ ). In Model 2, Self-Management 1 was allowed to load on both the Self-Management and the Interpersonal subscales. A review of the item content suggested this to be a substantively supportable model improvement, as the items were focused on cooperation and compromise with other students. In the Final model, another error term between Self-Management 1 and Interpersonal 1 was allowed to correlate. The changes in fit statistics for each model are presented in Table 2. As illustrated, each change in the hypothesized model results in noteworthy increases in fit statistics and nearly all suggest an adequate fit between the model and the data. The  $\chi^2$  remained statistically significant and the RMSEA was slightly higher than desirable. However, the RMR, GFI, AGFI, and CFI all suggest a good to excellent fit.

Obviously, the fit of a model to the data can be almost endlessly increased by merely continuing to free parameters in the model. In the present analyses, a balance was sought between providing an acceptable fit and having a model that would replicate well in other samples, while seeking to retain, if possible, the original structure of the instrument. The Final model is presented in Figure 1. In all cases, path values were statistically significant ( $p < .05$ ) and were relatively reliable predictors of the latent variables (note that the path values presented are standardized). The one exception to this is Self-Management 1,

which had a relatively lower, but still statistically significant, loading on the Interpersonal Skills subscale ( $r = .34$ ) and the Self-Management subscale ( $r = .56$ ). The items loaded most highly on those recommended by the author of the instrument.

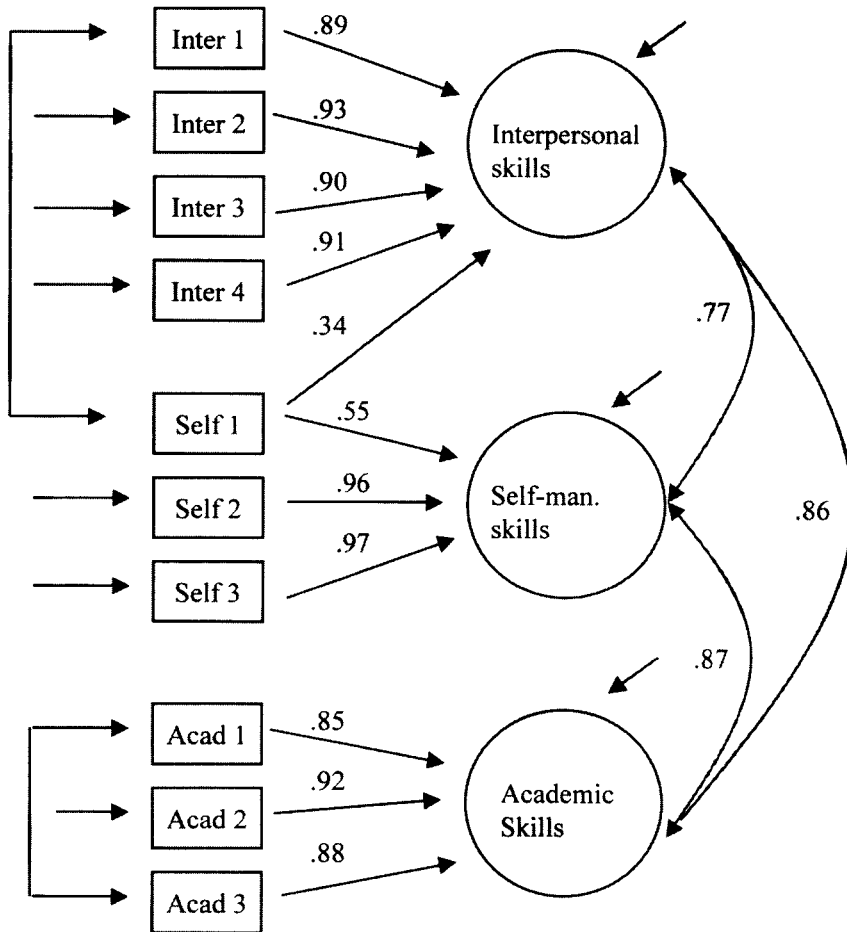
Once post-hoc model fitting has been done, the best test of the robustness of a mode is to test it on a second, independent sample. In the present study, the remaining half of the standardization sample, the even split, that had been held out in the previous analyses acted as a cross-validation sample ( $N = 1018$ ). The Final model was cross-validated in two ways. First, only the general pattern identified in the Final model was specified for the new sample (called the Pattern Model in Table 2). Second, the specific path values and correlations among latent variables were specified to be equal to those from the Final model (called the Values model in Table 2). This is a more restrictive and rigorous test of invariance. Note that the correlated errors were not specified as having the same values as in the Final model (Byrne, 1989). The fit statistics for the Pattern and Value models are presented in Table 2. The new hypothesized model replicated extremely well. That is, even with a more restrictive test, all fit statistics suggest a good to excellent fit between the model and the data. In fact, the model seems to fit the cross-validation sample better than the initial sample with stronger fit statistics in all areas. Factor loadings for the observed variables were all statistically significant and generally strong, ranging from .85 to .97. The exception, Self-Management 1, had more modest loadings of .34 on Interpersonal skills and .55 on Self-Management Skills. The generally strong values suggest that the observed variables reliably measured the latent construct under study, and that the model is likely to be stable and invariant across samples.

### *Scale B: Antisocial Behavior*

The approach to testing the hypothesized model for Scale B was followed as reported for



**Figure 1 Final model for Social Competence Scale of the SSBS.**



Scale A. Testlets for the Antisocial Behavior subscales are presented in Table 3. The initial model tested was the one proposed by the authors of the instrument; it consisted of three correlated factors. The fit statistics for the Initial model are presented in Table 4. Although the model provided a promising fit to the data, several fit statistics were in the “unacceptable” range (i.e., RMSEA, AGFI) or minimally acceptable (i.e., GFI); therefore attempts at model improvement commenced.

The first model improvement was made by allowing test packet Aggressive 1 to load on the Demanding subscale. A review of the item content of both the test packet and the subscale revealed that defiance and lying can easily be

considered to be in line with items loading on the Demanding subscale (e.g., ignores teacher). This resulted in a statistically significant change in  $\chi^2$  using a one degree of freedom test ( $\Delta \chi^2 = 69.97, df = 1, p < .01$ ; Byrne, 1989), and movement in the desired direction for nearly all fit statistics. The remaining modifications allowed three error terms to correlate in models 2 through 4. These were the errors between the following mini-scales: Aggressive 2 and Hostile 3, Hostile 4 and Hostile 2, and Hostile 4 and Hostile 1. The  $\chi^2$  value remained statistically significant in the Final model, but all other indices showed adequate to excellent fit between the model and the data. The factor loadings for Scale B were statistically significant

**Table 3 Testlets for Scale B, Antisocial Behavior**

Hostile 1	blames other students for problems disregards feelings and needs of other students is boastful; brags
Hostile 2	teases and makes fun of other students will not share with other students insults peers argues and quarrels with peers
Hostile 3	is disrespectful or "sassy" acts as if he/she is better than others whines and complains
Hostile 4	is easily provoked; has a short fuse has temper outbursts or tantrums is easily irritated
Aggressive 1	defies teacher or other school personnel lies to teacher or other school personnel gets in trouble at school
Aggressive 2	gets into fights destroys or damages school property is physically aggressive
Aggressive 3	takes things that are not his/hers cheats on schoolwork or in games threatens other students, is verbally aggressive swears or uses obscene language
Demanding 1	ignores teacher or other school personnel cannot be depended on unproductive, achieves very little
Demanding 2	is overly demanding of teacher's attention bothers and annoys other students demands help from other students
Demanding 3	is difficult to control disrupts ongoing activities acts impulsively or without thinking

**Table 4 Summary Statistics for Scale B Models**

	Author's		Post-hoc model fitting			Cross-validation	
	initial	Model 1	Model 2	Model 3	Model 4	pattern	values
Change		LX(5,3)	TD(6,3)	TD(4,2)	TD(4,1)		
$\chi^2$ <sup>a</sup>	581.03	511.06	445.93	370.94	310.23	393.1	398.2
df	32	31	30	29	28	28	36
$\Delta\chi^2$		69.97	65.13	74.99	60.71		
RMSEA	0.13	0.12	0.12	0.11	0.098	0.11	0.1
ECVI	0.62	0.53	0.49	0.41	0.35	0.43	0.42
RMR	0.025	0.023	0.021	0.02	0.018	0.02	0.02
GFI	0.9	0.91	0.92	0.93	0.94	0.93	0.93
AGFI	0.82	0.84	0.85	0.87	0.89	0.86	0.89
CFI	0.96	0.96	0.97	0.97	0.98	0.97	0.97

<sup>a</sup> All  $\chi^2$  values are statistically significant at  $p < .001$ .

(using a t-statistic,  $p < .05$  for all); nearly all were above .84 (range .84 to .92). The Aggressive 1 testlet was allowed to double load and had more modest path values of .59 on the Antisocial-Aggressive latent variable, and .37 on the Demanding-Disruptive latent variable. Again, these values indicate that the observed variables measured the latent construct of Antisocial Behavior well. The Final model is depicted in Figure 2. As in Figure 1, all path values are standardized.

Once again, the Final model was cross-validated on the even split of the standardization sample using the procedure described for Scale A. The model replicated well when only the pattern of fixed and freed parameters was constrained, or when the path values were fixed to equal those from the odd sample. There was a slight decrease in some fit statistics, as is expected when cross-validating. However, the changes were minor, usually less than .01.

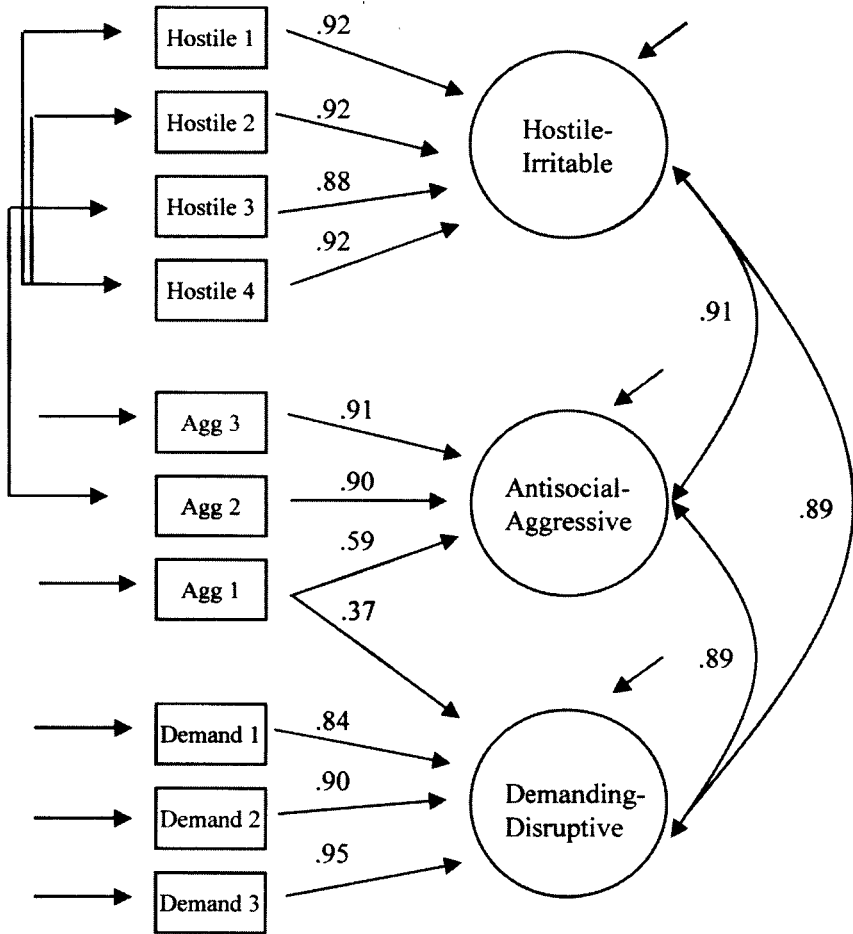
### Discussion

Results from the present investigation add to the body of literature on the psychometric

properties of the SSBS. The structure of the instrument, as proposed by the author, was supported by the current findings. The model modifications made were relatively minor and likely of more interest to researchers and psychometricians than to practitioners. From a clinical perspective, these modifications are of limited importance and do not suggest the need for any changes in the clinical usage of the instrument. In fact, findings from the confirmatory analysis indicate that the testlets, and indirectly the items, are quite reliable, all strong predictors of the latent constructs under study.

The analyses unquestionably provide support for the continued use of both the Social Competence and Antisocial Behavior subscales of the SSBS. Of the two, however, the model for the Antisocial Behavior scale provided a relatively better fit to the data. There are several possible explanations for this, including that this was merely a chance finding. However, externalizing or "acting-out" behavior has been a focus of investigation within the field of psychology for decades and

**Figure 2 Final model for the Antisocial Behavior Scale of the SSBS.**



the problematic behaviors children may exhibit have been fairly well identified. In addition, external behaviors can be more easily seen by an external observer, often call attention to themselves (as compared to children engaging in prosocial behavior), and translate more easily into pencil-and-paper questions. Thus, the stronger fit for Scale B may reflect a better ability to measure problematic behaviors both on this instrument and in the field generally.

In addition to supporting the scale structure of the SSBS, our results contribute to the overall understanding of theories and models of social behavior of children and youth. Although some practitioners and researchers

may assume that social competence and anti-social behavior are polar opposites of each other, and may infer the presence of one from the absence of the other, such is not the case. The SSBS was designed to provide co-normed scales to measure the constructs of social competence and antisocial behavior simultaneously but separately. These results indicate that each of the major constructs Social Competence and Antisocial Behavior are complex dimensional constructs that are superordinate to various subdimensions. For example, the Social Competence scale, which was built using items designed to reflect peer-related, teacher-related, and self-related forms of social adjustment, reflects each of these

areas. The Antisocial Behavior scale, which was constructed with both covert and overt forms of general antisocial behavior descriptors, appears to reflect a multidimensional construct composed of various types of antisocial, aggressive, oppositional, and disruptive behavioral characteristics that are social in nature or have socially-related repercussions.

In considering the contribution of the SSBS and the present study to theoretical knowledge about antisocial behavior, it is important to recognize that the Antisocial Behavior scale was not designed to be a broad-band representation of problem social behavior, but to reflect the antisocial behavior construct in a general manner. Other types of problem social behavior not measured by the SSBS are likely to relate to the internalizing dimensions of anxiety, depression, and social withdrawal. Because the SSBS was designed specifically to measure the general constructs of social competence and antisocial behavior as they are typically exhibited in school settings, these results are not necessarily generalizable across the full spectrum of social behavior problems, and do not necessarily inform the theoretical foundations of these other areas.

The results of this investigation have useful implications for clinical and educational practice, especially when considered in view of the overall literature on the SSBS. The Social Competence and Antisocial Behavior scales of the SSBS clearly are multidimensional, and enable users not only conduct a broad-band assessment of the general constructs of social competence and antisocial behavior, but to look at the narrow-band components of these constructs in more detail. Assessment results may be used not only to form general hypotheses about the social skills and antisocial problem behavior of a given student, but to make inferences regarding more highly specified subdimensions of these constructs, such as interpersonal skills, self-management, and antisocial-aggressive behavior. Such information may prove useful in making diagnostic or classification decisions, as well as in

developing intervention plans that are tailored to the specific needs of a student. A best practice in linking assessment to intervention is to match intervention goals to specifically identified deficits or problem areas, rather than using a "one size fits all" generic intervention plan (Peacock Hill Working Group, 1991).

When used as a screening tool to identify students who are potentially at risk because of social-behavioral problems, total scores rather than subscale scores should be the focus, and the screening criteria for SSBS scores should be set loose enough so that there are few if any "false-negative" errors. Further, when used as an assessment tool for service eligibility and detailed intervention planning, the SSBS should be used in conjunction with other instruments and methods within the context of a multi-method, multi-source, multi-setting design (Merrell, 1999) to reduce error variance and produce an aggregated view of social and emotional behavior.

As stated, instrument validation is an ongoing process. At no time can an instrument be considered "valid" or "reliable," but evidence can be gathered regarding the subjects and situations in which scores from an assessment tool are reliable and valid. In that continuing process, replication of the present findings will be important. In extending the current findings, invariance across psychologically relevant variables such as age, gender, and ethnicity will further our understanding of the strengths and limitations of the SSBS. At the present time, however, empirical evidence supports the continued use of the SSBS as a screening and assessment tool for social and antisocial behavior of school-aged children.

## References

- Byrne, B.M. (1989). *A primer of LISREL: Basic applications and programming for confirmatory factor analytic models*. New York: Springer-Verlag.
- Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts,*

- applications, and programming. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Collins, J.M., & Gleaves, D.H. (1998). Race, job applicants, and the five-factor model of personality: Implications for black psychology, industrial/organizational psychology, and the five-factor theory. *Journal of Applied Psychology*, 83, 531-544.
- Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2<sup>nd</sup> ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Crowley, S. L., & Fan, X. (1997). Structural equation modeling: Basic concepts and applications in personality assessment research. *Journal of Personality Assessment*, 68, 508-531.
- Demaray, M. K., Ruffalo, S. L., & Carlson, J. (1995). Social skills assessment: A comparative evaluation of six published rating scales. *School Psychology Review*, 24, 648-671.
- Elksnin, L. K., & Elksnin, N. (1995). *Assessment and instruction of social skills*. San Diego: Singular.
- Emerson, E. N., Crowley, S. L., & Merrell, K. W. (1994). Convergent validity of the School Social Behavior Scales with the Child Behavior Checklist and Teacher's Report Form. *Journal of Psychoeducational Assessment*, 12, 372-380.
- Flanagan, D. P., Alfonso, V. C., Primavera, L. H., Povall, L., & Higgins, D. (1996). Convergent validity of the BASC and SSRS: Implications for social skills assessment. *Psychology in the Schools*, 33, 13-23.
- Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, 7, 286-299.
- Gorsuch, R. (1997). Exploratory factor analysis: Its role in item analysis. In J. A. Schinka & R. L. Greene (Eds.), *Emerging issues and methods in personality assessment* (pp. 309-338). Mahwah, NJ: Erlbaum.
- Gregory, R. J. (1996). *Psychological testing: History, principles, and applications* (2<sup>nd</sup> ed.). Boston: Allyn and Bacon.
- Hooper, S. R. (1998). Review of the School Social Behavior Scales. In J. C. Impara & B. S. Plake (Eds.), *The thirteenth mental measurements yearbook* (pp. 878-880). Lincoln: University of Nebraska, Buros Institute of Measurements.
- Joreskog, K. G., & Sorbom, D. (1996). *LISREL 8: User's reference guide* (2<sup>nd</sup> ed.). Chicago: Scientific Software International.
- Kreisler, T. A., Mangione, C., & Landau, S. (1997). Review of the School Social Behavior Scales. *Journal of Psychoeducational Assessment*, 15, 182-190.
- Loeber, R. (1991). Questions and advances in the study of developmental pathways. In D. Cicchetti & S. L. Toth (Eds.), *Rochester symposium on developmental psychopathology, Vol. 3: Models and integrations* (pp. 97-116). Rochester, NY: Rochester University Press.
- Loeber, R., Wung, P., Keenan, K., Giroux, B., Stouthamer-Loeber, M., Van Kammen, W. B., et al. (1993). Developmental pathways in disruptive child behavior. *Development and Psychopathology*, 5, 101-132.
- Martin, R. P. (1988). *Assessment of personality and behavior problems*. New York: Guilford.
- Merrell, K. W. (1992). The utility of the school social behavior scales in differentiating students with behavioral disorders from other students with disabilities. *Severe Behavior Disorders of Children and Youth*, 15, 27-34.
- Merrell, K. W. (1993a). *School Social Behavior Scales*. Eugene, OR: Assessment-InterventionResources (<http://www.assessment-intervention.com>).
- Merrell, K. W. (1993b). Using behavior rating scales to assess social skills and antisocial behavior in school settings: Development of the School Social Behavior Scales. *School Psychology Review*, 22, 115-133.
- Merrell, K. W. (1999). *Behavioral, social, and emotional assessment of children and adolescents*. Mahwah, NJ: Erlbaum.
- Merrell, K. W. (2000). Informant report: Rating scale measures. In E. S. Shapiro & T. R. Kratochwill (Eds.), *Conducting school-based assessment of child and adolescent behaviors* (pp. 203-234). New York: Guilford Publications.
- Merrell, K. W. (2002). *Home & Community Social Behavior Scales*. Eugene, OR: Assessment-Intervention Resources (<http://www.assessment-intervention.com>).
- Merrell, K. W., Cedeno, C. J., & Johnson, E. R. (1993). The relationship between social behavior and self-concept in school settings. *Psychology in the Schools*, 30, 293-298.
- Merrell, K. W., & Gill, S. J. (1994). Using teacher ratings of social behavior to differentiate gifted from non-gifted students. *Roeper Review*, 16, 286-289.
- Merrell, K. W., & Gimpel, G. A. (1998). *Social skills of children and adolescents: Conceptualization, assessment, treatment*. Mahwah, NJ: Erlbaum.
- Merrell, K. W., Sanders, D. E., & Poppinga, M. (1993). Teacher ratings of social behavior as a predictor of special education status: Discriminant validity of the School Social Behavior Scales. *Journal of Psychoeducational Assessment*, 11, 220-231.
- Nurcombe, B., Baker, L., Cantwell, D. P., Singer, J. L., Wolff, S., Yates, A., et al. (1991). Developmental pathways. In M. Lewis (Ed.), *Child and adolescent psychiatry: A comprehensive textbook* (pp. 160-221). Baltimore, MD: Williams & Wilkins.
- Patterson, G. R., DeBaryshe, B. D., & Ramsey, E. (1989). A developmental perspective on antisocial behavior. *American Psychologist*, 44, 329-335.
- Peacock Hill Working Group. (1991). Problems and

- promises in special education and related services for children and youth with emotional or behavioral disorders. *Behavioral Disorders*, 16, 299-313.
- Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach*. Mahwah, NJ: Erlbaum.
- Robbins, R., & Merrell, K. W. (1998). Cross-informant comparisons of the Home and Community Social Behavior Scales and the School Social Behavior Scales. *Diagnostique*, 23, 204-218.
- Rubin, K. H., Hymel, S., Mills, R. S. L., & Rose-Krasnor, L. (1991). Conceptualizing different developmental pathways to and from social isolation in childhood. In D. Cicchetti & S. L. Toth (Eds.), *Internalizing and externalizing expressions of dysfunction* (pp. 91-122). Hillsdale, NJ: Erlbaum.
- Walker, H. M., Colvin, G., & Ramsey, E. (1995). *Antisocial behavior in school: Strategies and best practices*. New York: Brooks Cole.
- Welsh, L. A. (1998). Review of the School Social Behavior Scales. In J. C. Impara & B. S. Plake (Eds.), *The thirteenth mental measurements yearbook* (pp. 880-882). Lincoln: University of Nebraska, Buros Institute of Measurements.

### Authors' Note

A previous version of this paper was presented at the Kansas Conference on Clinical Child Psychology, Lawrence, KS, October 1998.