

## RESEARCH BRIEF

---

### Assessing Social Validity of School-wide Positive Behavior Support Plans: Evidence for the Reliability and Structure of the Primary Intervention Rating Scale

Kathleen Lynne Lane, Jemma Robertson Kalberg, Allison Leigh Bruhn,  
Steven A. Driscoll, Joseph H. Wehby, and Stephen N. Elliott  
*Vanderbilt University*

*Abstract.* This study provides initial evidence for the reliability and structural validity of scores from the Primary Intervention Rating Scale (Lane, Robertson, & Wehby, 2002), an adapted version of the Intervention Rating Profile-15 (Witt & Elliott, 1985) designed to assess faculty's perceptions of social validity of primary prevention plans prior to intervention onset. Results indicated the Primary Intervention Rating Scale is a one-factor instrument, with high internal consistency and utility. These results were found to replicate across educators from elementary, middle, and high schools. In addition, there was a significant, positive relation between social validity and treatment integrity when examining data at the school-site level. Limitations and future directions are offered.

Recently a series of articles was written to delineate core quality indicators to guide the research and teaching communities in evaluating the scientific rigor of experimental, quasi-experimental (Gersten, Fuchs, Compton, Coyne, Greenwood, & Innocenti, 2005), and single-

case (Horner, Carr, Halle, McGee, Odom, & Wolery, 2005) methodologies used to investigate educational practices. These indicators were used to evaluate individual studies and provide guidelines for determining if the practice under investigation is indeed evidence-

---

The authors thank Joe Witt and Steve Elliott for the work on the IRP-15 and Children's Intervention Rating Profile. These measures have been essential in our school-based intervention work. This project was funded in part by the following sources: Office of Special Education Programs (OSEP) Directed Research Grant H324D020048; OSEP Field Initiated Grant H324C030044B; the Henderson Training Series; and National Institute of Child Health and Human Development Grant P30HD15052 to Vanderbilt University.

Correspondence regarding this article should be addressed to Kathleen Lynne, Vanderbilt University, Peabody College, Department of Special Education, Nashville, TN 37203; E-mail: kathleen.lane@vanderbilt.edu

Copyright 2009 by the National Association of School Psychologists, ISSN 0279-6015, which has nonexclusive ownership in accordance with Division G, Title II, Section 218 of P. L. Law 110-161 and NIH Public Access Policy.

based. One of these core quality indicators is social validity.

Social validity refers to the extent to which consumers (e.g., teachers, parents, and students) view a given practice as addressing socially significant goals, socially acceptable treatment procedures, and socially important intervention outcomes (Wolf, 1978). Ideally, social validity should be assessed before intervention onset and immediately following intervention completion, with the latter being a more common practice (Lane, Kalberg, & Menzies, in press). Assessing social validity before conducting an intervention may provide useful information for identifying intervention features potentially in need of revision; determine the level of commitment or buy-in; and perhaps predict the likelihood that the intervention will be implemented as designed (treatment integrity) and sustained over time. Namely, if the person completing the social validity measure does not view the intervention as targeting a meaningful or realistic goal, if the procedures are viewed as too cumbersome, and/or outcomes (e.g., increased academic performance) seem improbable, it is highly likely that the intervention will not be implemented with fidelity (Reimers, Wacker, & Koeppel, 1987). If the intervention is not implemented as planned, then it is possible that desired outcomes are unlikely to be realized and, consequently, the practice will not be sustained over time (Noell & Gresham, 1993). In sum, social validity may affect implementation, initial effectiveness, maintenance, and future use (Elliott, 1988).

Despite the importance of social validity for all levels of prevention, it mainly has been assessed in targeted interventions such as secondary or tertiary levels of prevention, with limited application in primary prevention efforts. For example, in the literature focusing on academic-based secondary prevention efforts and function-based interventions at the tertiary level for students with and at risk for emotional and behavioral disorders (EBD), social validity often has been assessed using standardized ratings scales (Lane, 2004). Standardized instruments such as the Treatment Evaluation Inventory (Kazdin, 1980), Treat-

ment Acceptability Rating Form—Revised (Reimers, Wacker, Cooper, & DeRaad, 1992), the Intervention Rating Profile-15 (IRP-15; Witt & Elliott, 1985), and the Behavior Intervention Rating Scale (VonBrock & Elliott, 1987) have been used to assess social validity from the teacher's perspective. In addition, the IRP-15 and Behavior Intervention Rating Scale have been designed for use with parents and psychologists. The IRP-15 also has a downward extension, the Children's Intervention Rating Profile, developed for use with children. In general, these factor analytically derived measures require teachers to rate each item on a Likert-type scale. Next, composite scores are created, with higher scores associated with higher treatment acceptability and/or effectiveness per the constructs measured. These instruments have been developed with evidence of strong internal consistency.

Although such standardized measures of social validity have been used with targeted attention, less attention appears to have been devoted to assessing social validity of primary prevention efforts such as school-wide positive behavior support plans. Recently, two systematic reviews of primary prevention programs with a behavioral component have been conducted, with one focusing on the elementary level (Lane, Kalberg, & Edwards, 2008) and another focusing on middle and high schools (Lane, Robertson, & Graham-Bailey, 2006). The findings of these studies suggest that approximately one-third of the school-wide primary prevention efforts included in the review at the elementary level mentioned and reported social validity. Similarly, findings of the review conducted at the middle school and high school levels revealed that slightly more than half of the studies included in the review mentioned social validity from the teacher, student, and/or parent perspective. Although social validity was assessed using a variety of methods in these studies of primary prevention programs, it was most often assessed using surveys with unknown psychometric qualities. In fact, based on a review of published research, instruments comparable to the Treatment Evaluation Inventory, Treatment Acceptability Rating Form—Revised,

**Table 1**  
**Teacher Characteristics**

| Variable                                   | Level             | Elementary<br><i>N</i> = 329 | Middle School<br><i>N</i> = 86 | High School<br><i>N</i> = 202 |
|--|-------------------|------------------------------|--------------------------------|-------------------------------|
| Gender <i>n</i> (%)                        | Male              | 27 (9.03)                    | 12 (14.46)                     | 66 (34.38)                    |
|  | Female            | 272 (90.97)                  | 71 (85.54)                     | 126 (65.63)                   |
| Highest degree attained <i>n</i> (%)       | Bachelor's        | 132 (45.36)                  | 34 (40.48)                     | 94 (48.45)                    |
|  | Master's          | 113 (38.83)                  | 35 (41.67)                     | 72 (37.11)                    |
|  | Master's + 30     | 33 (11.34)                   | 8 (9.52)                       | 17 (8.76)                     |
|  | Eds               | 4 (1.37)                     | 3 (3.57)                       | 3 (1.55)                      |
|  | PhD               | 4 (1.37)                     | 2 (2.38)                       | 3 (1.55)                      |
|  | Other             | 5 (1.72)                     | 2 (2.38)                       | 5 (2.58)                      |
| Program taught <i>n</i> (%)                | General education | 211 (72.76)                  | 56 (70.00)                     | 136 (73.91)                   |
|  | Special education | 38 (13.10)                   | 16 (20.00)                     | 31 (16.85)                    |
|  | Other             | 41 (14.14)                   | 8 (10.00)                      | 17 (9.24)                     |
| Waivers <i>n</i> (%)                       |                   | 2 (0.67)                     | 1 (1.18)                       | 4 (2.11)                      |
| Teaching experience <i>M</i> ( <i>SD</i> ) |                   | 13.41 (9.76)                 | 12.78 (9.75)                   | 11.42 (9.93)                  |

*Note.* Frequency counts and percentages are computed based on the number of respondents who completed the items. Ed.S. refers to educational specialists. Ph.D. refers to a doctorate of philosophy. The term *waiver* refers to those teachers who have not yet completed teaching certification. Teaching experience is indicated in number of years.

IRP-15, and Behavior Intervention Rating Scale have not been developed and validated to assess teachers' perceptions of the social validity of school-wide programs. Given the intended effect of such programs on large numbers of students, it is surprising that their social validity has not been an explicit, measured variable.

### Purpose

This study provides initial evidence for the reliability and validity of scores from the Primary Intervention Rating Scale (PIRS; Lane et al., 2002). The PIRS is an adapted version of the IRP-15, a widely used measure of social validity. In this report, we explain how the IRP-15 was adapted to provide the field with a brief teacher-completed rating scale to assess social validity of primary prevention programs before intervention onset.

### Method

#### Participants

Participants were 617 teachers representing 11 elementary (329 teachers), 3 mid-

dle (86 teachers), and 5 high schools (202 teachers) in middle Tennessee, who participated in a year-long training series designed to construct a three-tiered model of positive behavior support. Teachers were predominantly female, with slightly more than half of teachers at each school level having earned educational degrees beyond the bachelor's level (see Table 1). Most teachers were general educators, with the majority of teachers having earned teaching credentials. The level of teaching experience was 13.41 (*SD* = 9.76) years at the elementary level, 12.78 (*SD* = 9.75) at the middle school level, and 11.42 (*SD* = 9.93) at the high school level.

The sample of teachers was from schools representing a range of geographic locales within middle Tennessee and included both public and private schools, although the majority of participating schools were in the public sector. Student enrollment ranged from 44 to 1838. The economic disadvantage rate for the public schools ranged from 1.0% to 84.4%, with 4 elementary schools receiving Title 1 services. The diversity of the student population also varied, with the percentage of

Caucasian students ranging from 7.6% to 98.1% (Tennessee Department of Education, 2006).

## Procedures

Each of the 19 schools participated in a year-long training series in positive behavior support (PBS) conducted at Vanderbilt University by the first author. As part of this institutional review board approved study, each school established a PBS team that included, at a minimum, an administrator with decision-making authority, two general education teachers, a special education teacher, a parent, and a student. The team attended the PBS training series to design a plan containing primary, secondary, and tertiary levels of prevention. As part of the design process, each team received input on the primary prevention component of the school-wide plan. The PBS team presented an overview of the proposed PBS plan during a regularly scheduled faculty meeting. The team passed out copies of the drafted plan, which included the following: the mission and purpose statement; an overview of the components constituting the behavioral, academic, and social domains; behavioral expectation matrix; procedures for teaching; procedures for reinforcing; and procedures for monitoring. Faculty and staff were encouraged to ask questions and participate in the discussion led by the PBS team. At the close of the meeting, they completed PIRS: Pre-Implementation, the adapted version of the IRP-15, to assess social validity of the primary plan from the teachers' perspective. Faculty completed the PIRS anonymously, turning in the completed survey after the faculty meeting. Data were entered into a database. Research assistants assessed reliability of entry for 100% of the surveys, with reliability >99%.

## Measures

The PIRS: Pre-Implementation was adapted from the IRP-15 to measure faculty perceptions of the social significance of intervention goals, social acceptability of intervention procedures, and likelihood of socially important outcomes. The IRP-15 is a 15-item, factor analytically derived survey used to mea-

sure treatment acceptability and perceived effectiveness. Each item is rated on a 6-point Likert-type scale ranging from 1 (strongly disagree) to 6 (strongly agree). Internal consistency reliabilities range from .88 to .98. The IRP-15 was adapted for use in obtaining teacher social validity ratings of the school-wide plan by (a) modifying the item's wording to reflect a school-wide, primary prevention program rather than a targeted intervention, and (b) adding two additional items regarding monitoring procedures (Item 15: "The monitoring procedures are manageable"; Item 16: "The monitoring procedures will give the necessary information to evaluate the plan"). Each of the 17 items constituting the PIRS was rated on the same 6-point Likert-type scale, yielding a total score ranging from 17 to 102, with higher scores suggesting higher acceptability. We developed three versions of the PIRS, one for the elementary, middle, and high school levels.

## Treatment Integrity

Treatment integrity of the primary prevention plan was assessed during the first year of program implementation for schools electing to participate in program evaluation activities ( $N = 14$ ). One method of monitoring treatment integrity common among all schools participating in the evaluation study was teacher-completed interval ratings. Teachers who agreed to participate in an overall study of program implementation to evaluate the effect of the primary plan completed a brief treatment integrity scale, Positive Behavior Support Plan: Primary Level, either quarterly or monthly to evaluate the degree to which they implemented components of the primary plan. The component checklist contained specific items depicting procedures for teaching and reinforcing the school-wide expectations (e.g., giving a student a ticket paired with behavior-specific praise tied to one of the school-wide expectations). At the end of the quarter or month, teachers rated the extent to which they addressed each component on a 3-point Likert-type scale (not at all = 0, part of the time = 1, all of the time = 2). Com-

posite scores for teacher-completed interval ratings were computed by averaging either the 4 quarterly or 10 monthly session integrity scores collected across the academic year for each teacher. From these composite scores, an average score was computed for the entire school.

### Experimental Design and Statistical Analysis

To examine the internal structure of the PIRS for each school level, an exploratory factor analysis of the 17 items was conducted using squared multiple correlations as prior communality estimates. The principal factor method was used to extract the factors rather than a principal components analysis given that the goal was not data reduction (Fabrigar, Wegener, MacCallum, & Strahan, 1999). Instead, the intent was to estimate the pattern of relations between the common factors and each variable measured, with the overall goal of identifying latent constructs underlying these variables. A varimax rotation was planned to maintain the orthogonality of the reference frame; however, rotation was not conducted given that only one factor was retained. We used three procedures to determine the number of factors to be retained. First, we used the Kaiser criterion that involved computing the eigenvalues for the correlation matrix and identifying eigenvalues  $>1$ . Second, we examined scree plots. The eigenvalues of the correlation matrix were plotted in order of descending values. Then, the “elbow” of the last substantial drop in the magnitude of the eigenvalue was identified. The number of factors to be retained is the last number before the last substantial drop. Both the Kaiser criterion and scree plot procedures led to identical decisions with respect to the number of factors to be retained. Finally, the makeup of the items for a factor was determined when item loadings were  $\geq 0.55$  for one factor and  $< 0.55$  on a subsequent factor.

To estimate internal consistency for each school-level PIRS form, alpha coefficients were computed for each administration for elementary, middle, and high school teachers who com-

pleted the survey before implementing the school-wide PBS plan. We also reported Cronbach coefficient alpha with deleted variables, item total correlations, and intercorrelations.

Finally, we examined predictive validity by examining the extent to which PIRS scores at the end of the training year predicted the level of treatment fidelity during the first year of program implementation. We computed Pearson correlation coefficients between the schools’ mean PIRS scores with mean treatment integrity scores for the first year of program implementation. It was necessary to conduct these analyses at the school level rather than the individual teacher level given that PIRS data were collected anonymously to increase the likelihood of receiving more candid responses. Thus, it was not possible to link individual teachers’ PIRS and treatment integrity scores.

In sum, we examined basic internal structure and relationship aspects of the PIRS across three levels of schools. This approach provided a replication of the structural evidence with comparable samples of educators and evidence for the results of the PIRS.

## Results

### Factor Structure

In each of the three exploratory factor analyses conducted using squared multiple correlations as prior communality estimates, there was one factor retained for the elementary, middle, and high school version of the PIRS, as determined using the scree plot and an eigenvalue ( $>1.0$ ) criteria. Two- and three-factor solutions were tested. Yet, none of the items loaded on the second or third factors. Factor loadings on the one-factor solution exceeded the 0.40 values typically required, ranging from 0.62 to 0.90 at the elementary, 0.64 to 0.92 at the middle, and 0.57 to 0.92 at the high school levels. Explained variance was 70.47%, 72.41%, and 70.47%, respectively. See Table 2.

### Internal Consistency

**Elementary school level.** Based on 329 elementary teacher ratings representing 11

**Table 2**  
**Survey Items and Corresponding Factor Loadings by School Level**

| Item  | Level                        |                         |                        |
|---|------------------------------|-------------------------|------------------------|
|   | Elementary<br><i>N</i> = 329 | Middle<br><i>N</i> = 86 | High<br><i>N</i> = 202 |
| 1. This would be an acceptable intervention for the elementary <sup>a</sup> school.             | 0.87                         | 0.85                    | 0.89                   |
| 2. Most teachers would find this intervention appropriate.                                      | 0.84                         | 0.85                    | 0.80                   |
| 3. This intervention should prove effective in meeting the purposes.                            | 0.89                         | 0.91                    | 0.90                   |
| 4. I would suggest the use of this intervention to other teachers.                              | 0.88                         | 0.76                    | 0.90                   |
| 5. The intervention is appropriate to meet the school's needs and mission.                      | 0.89                         | 0.89                    | 0.86                   |
| 6. Most teachers would find this intervention suitable for the described purposes and mission.  | 0.86                         | 0.88                    | 0.79                   |
| 7. I would be willing to use this intervention in the school setting.                           | 0.83                         | 0.87                    | 0.81                   |
| 8. This intervention would <i>not</i> result in negative side-effects for the students.         | 0.62                         | 0.64                    | 0.57                   |
| 9. This intervention would be appropriate for a variety of students.                            | 0.81                         | 0.84                    | 0.85                   |
| 10. This intervention is consistent with those I have used in school settings.                  | 0.71                         | 0.65                    | 0.62                   |
| 11. The intervention is a fair way to fulfill the intervention purposes.                        | 0.87                         | 0.89                    | 0.88                   |
| 12. This intervention plan is reasonable to meet the stated purposes.                           | 0.89                         | 0.91                    | 0.92                   |
| 13. I like the procedures used in this intervention.  | 0.89                         | 0.91                    | 0.89                   |
| 14. This intervention is a good way to meet the specified purpose.                              | 0.90                         | 0.92                    | 0.92                   |
| 15. The monitoring procedures are manageable.   | 0.74                         | 0.86                    | 0.83                   |
| 16. The monitoring procedures will give the necessary information to evaluate the plan.         | 0.81                         | 0.85                    | 0.83                   |
| 17. Overall, this intervention would be beneficial for elementary <sup>a</sup> school students. | 0.89                         | 0.91                    | 0.89                   |

*Note.* More recent versions of the instrument used the term *PBS plan* rather than the term *intervention*.

<sup>a</sup>School level was referred to as elementary, middle, or high school in each of the respective versions.

elementary schools prior to onset of a school-wide PBS program, the Cronbach's alpha value was .97. Cronbach's coefficient alphas with deleted variables were also .97 for all items, except for Item 8, which was .98. Item-total correlations ranged from .61 to .88, suggesting strong internal consistency at the elementary level (see Table 3). Intercorrelations were moderately to highly positive (.46–.88) and statistically significant at the  $p < .0001$  level.

**Middle school level.** Based on 86 middle school teacher ratings representing three middle schools prior to onset of a school-wide PBS program, the Cronbach's alpha value was .98. Cronbach's coefficient alphas with deleted variables ranged from .97 to .98 for all items. Item-total correlations ranged from .63 to .91. As with the elementary findings, intercorrelations were moderately to highly positive (.42–.88) and significant at the  $p < .0001$  level.

**Table 3**  
**Cronbach's Coefficient Alpha With Deleted Variables**

| Item | Elementary $N = 329$ |      |                        |       | Middle School $N = 86$ |      |                        |       | High School $N = 202$ |      |                        |       |
|------|----------------------|------|------------------------|-------|------------------------|------|------------------------|-------|-----------------------|------|------------------------|-------|
|      | $M$                  | $SD$ | Standardized Variables |       | $M$                    | $SD$ | Standardized Variables |       | $M$                   | $SD$ | Standardized Variables |       |
|      |                      |      | $r$ With Total         | Alpha |                        |      | $r$ With Total         | Alpha |                       |      | $r$ With Total         | Alpha |
| 1    | 5.16                 | 0.82 | .86                    | .97   | 4.88                   | 0.96 | .85                    | .98   | 4.62                  | 0.99 | .88                    | .97   |
| 2    | 4.98                 | 0.76 | .83                    | .97   | 4.71                   | 1.00 | .84                    | .98   | 4.34                  | 0.85 | .79                    | .97   |
| 3    | 4.93                 | 0.90 | .88                    | .97   | 4.75                   | 0.99 | .90                    | .97   | 4.40                  | 0.90 | .88                    | .97   |
| 4    | 4.93                 | 0.92 | .87                    | .97   | 4.65                   | 1.02 | .75                    | .98   | 4.40                  | 0.96 | .89                    | .97   |
| 5    | 5.07                 | 0.90 | .87                    | .97   | 4.74                   | 1.03 | .88                    | .97   | 4.56                  | 0.94 | .85                    | .97   |
| 6    | 4.98                 | 0.83 | .85                    | .97   | 4.58                   | 1.06 | .86                    | .97   | 4.32                  | 0.83 | .78                    | .97   |
| 7    | 5.23                 | 0.85 | .82                    | .97   | 4.92                   | 1.05 | .85                    | .98   | 4.74                  | 0.96 | .80                    | .97   |
| 8    | 4.93                 | 0.91 | .61                    | .98   | 4.74                   | 1.01 | .63                    | .98   | 4.47                  | 0.92 | .57                    | .98   |
| 9    | 4.98                 | 0.88 | .80                    | .97   | 4.74                   | 0.97 | .83                    | .98   | 4.50                  | 0.96 | .84                    | .97   |
| 10   | 4.90                 | 0.88 | .70                    | .97   | 4.38                   | 1.22 | .64                    | .98   | 4.15                  | 1.16 | .60                    | .98   |
| 11   | 5.06                 | 0.85 | .86                    | .97   | 4.73                   | 0.94 | .88                    | .97   | 4.46                  | 0.89 | .87                    | .97   |
| 12   | 5.02                 | 0.83 | .88                    | .97   | 4.69                   | 0.96 | .90                    | .97   | 4.48                  | 0.91 | .91                    | .97   |
| 13   | 4.95                 | 0.99 | .88                    | .97   | 4.58                   | 1.14 | .90                    | .97   | 4.30                  | 1.05 | .88                    | .97   |
| 14   | 4.94                 | 0.89 | .88                    | .97   | 4.76                   | 1.01 | .91                    | .97   | 4.38                  | 0.98 | .91                    | .97   |
| 15   | 4.67                 | 1.00 | .73                    | .97   | 4.43                   | 1.21 | .84                    | .98   | 4.15                  | 0.98 | .82                    | .97   |
| 16   | 4.74                 | 0.92 | .80                    | .97   | 4.53                   | 1.08 | .83                    | .98   | 4.27                  | 0.94 | .81                    | .97   |
| 17   | 5.05                 | 0.85 | .88                    | .97   | 4.70                   | 1.13 | .89                    | .97   | 4.47                  | 1.01 | .88                    | .97   |

**High school level.** Based on 202 high school teacher ratings representing five high schools prior to onset of a school-wide PBS program, the Cronbach's alpha value was .97. Cronbach's coefficient alphas with deleted variables were also .97 for all items, except for Items 8 and 10, which were .98. Item total correlations ranged from .57 to .91. Intercorrelations were moderately to highly positive (.40–.88) and statistically significant at the  $p < .0001$  level.

#### Evidence for the Predictive Validity of PIRS

Of the 19 schools participating in the training series, 14 went on to implement the PBS plan and participate in a program evaluation study during next year. School sites' mean PIRS values ranged from 68.13 ( $SD = 15.32$ ) to 96.83 ( $SD = 6.27$ ; see Table 4). School sites' mean treatment integrity scores ranged from 67.45 ( $SD = 15.16$ )

to 96.62 ( $SD = 4.24$ ). The Pearson correlation coefficient between schools' mean PIRS and treatment integrity scores for schools participating in a program evaluation study was  $r = .71$ ,  $p = .005$ , suggesting a strong association and significant positive correlation between social validity and treatment integrity.

#### Discussion

As we move towards developing three-tiered models of prevention to meet the academic, behavioral, and social needs of an increasingly diverse student body, it will be important to design models that adhere to the core quality indicators necessary to draw valid inferences regarding intervention outcomes. One of these core quality indicators is social validity. To date, limited attention has been devoted to assessing social validity of primary prevention programs at the elementary, middle, or high school levels. This limited atten-

**Table 4**  
**Mean PIRS and Treatment Integrity**  
**Scores for Schools by Grade Level**

| Level      | PIRS     |           | Treatment Integrity |           |
|------------|----------|-----------|---------------------|-----------|
|            | <i>M</i> | <i>SD</i> | <i>M</i>            | <i>SD</i> |
| Elementary |          |           |                     |           |
| E1         | 82.85    | 16.50     | a                   | a         |
| E2         | 90.63    | 11.20     | a                   | a         |
| E3         | 80.13    | 13.28     | a                   | a         |
| E4         | 83.89    | 12.22     | 94.22               | 7.70      |
| E5         | 96.83    | 6.27      | 96.62               | 4.24      |
| E6         | 86.67    | 10.61     | 86.56               | 11.75     |
| E7         | 68.13    | 15.32     | a                   | a         |
| E8         | 73.36    | 16.35     | 78.88               | 17.66     |
| E9         | 88.11    | 9.30      | 91.12               | 7.66      |
| E10        | 83.97    | 11.78     | 86.45               | 12.55     |
| E11        | 81.00    | 13.15     | 85.94               | 17.16     |
| Middle     |          |           |                     |           |
| M1         | 82.58    | 10.28     | 85.60               | 12.10     |
| M2         | 73.75    | 22.26     | 67.85               | 14.17     |
| M3         | 78.55    | 12.89     | 86.23               | 12.15     |
| High       |          |           |                     |           |
| H1         | 73.83    | 13.95     | 71.48               | 16.34     |
| H2         | 71.38    | 15.71     | 81.38               | 15.49     |
| H3         | 79.79    | 18.46     | 67.45               | 15.16     |
| H4         | 81.58    | 9.11      | 74.87               | 14.48     |
| H5         | 82.79    | 10.79     | a                   | a         |

Note. PIRS = Primary Intervention Rating Scale.

<sup>a</sup>School did not elect to participate in the program evaluation during the following academic year.

tion to assessing social validity of primary prevention plans may be due, in part, to the absence of standardized instruments.

In this study, we addressed this void by providing initial evidence for the score reliability of the PIRS, an adapted version of the IRP-15 available for use in evaluating primary prevention plans across the kindergarten through 12th grade continuum. Results of the three replications of the analyses suggested that like the IRP-15, the PIRS is a one-factor instrument explaining approximately 70% of the variance at each school level, a substantial amount of variance. Furthermore, the PIRS demonstrated strong internal consistency

across the kindergarten through 12th grade span with overall internal consistency estimates of .97, .98, and .97, at the elementary, middle, and high school levels. The internal consistency of the PIRS actually exceeded the initial internal consistency of the IRP-15 and far exceeded the minimum criteria of .70 for internal consistency. Perhaps most significantly, this is the first study examining the extent to which social validity predicted treatment integrity of a school-wide, primary prevention program. Findings suggest a significant, positive relation indicating that higher social validity is predictive of a higher level of treatment fidelity, when examining data at the school-site level. In this case, findings appear to run contrary to those of Noell, Duhon, Gatti, and Connell (2002) and Noell et al. (2005) suggesting that teachers tend to rate the social validity of interventions highly irrespective of whether they implemented the program. Findings of this study suggested that social validity ratings collected before putting the primary plan in place were predictive of the level of program implementation during the first year. However, it is important to recognize that in this study treatment integrity was assessed using a self-report technique whereas the Noell studies used a direct observation technique.

Although this study provides initial evidence to support the use of the PIRS for assessing social validity of primary prevention programs before implementation, it is imperative that additional research be conducted to substantiate these findings before drawing definitive conclusions as to the functional utility of the PIRS. Also, the current findings should be interpreted in light of the following limitations.

### Limitations and Future Directions

First, although there was a range of locales represented in all three samples, all schools were located in middle Tennessee. Thus, replication in other regions that include more economically and ethnically diverse populations is necessary. Second, although the elementary and high school samples exceed the minimum criteria of 100 participants re-

quired for most factor analytical techniques, the middle school sample contained only 86 respondents.

Third, in this study the PIRS was completed before intervention onset to determine the feasibility of a primary prevention plan containing academic, social, and behavioral components. We encourage additional inquiry to determine the reliability of the PIRS when administered 1, 2, and 3 years following plan implementation. It is possible that perceptions of social validity will shift following implementation exceeding or falling short of initial expectations. Thus, we need to determine if the factor structure and internal consistency still hold when the PIRS is administered at different time points with a goal of capturing changes in social validity. If the PIRS is suitable for use across various phases of implementation, then this measure could be used to not only quantify social validity over time, but also to determine if social validity predicts (a) treatment integrity, (b) student performance, and (c) sustainability.

Fourth, this study examined the factor structure and internal consistency of the PIRS when completed by faculty. Additional research is needed to determine the reliability of this instrument when completed by the parents and students. This is necessary to determine if this instrument is also appropriate to measure social validity from other treatment agents' perspectives, as with measures of classroom level intervention such as the Behavior Intervention Rating Scale (VonBrock & Elliott, 1987).

Fifth, analyses pertaining to predictive validity should be interpreted cautiously given that (a) data were analyzed at the school-site level rather than at the teacher level, which may mask individual teacher differences, and (b) treatment integrity data were collected using teacher self-report techniques that often suggest higher levels of treatment integrity than direct observation techniques. Future research is needed to establish the generalizability of the findings, with studies examining data at the individual teacher level and using more direct fidelity measures.

Despite these limitations, the study provides initial evidence for the reliability and

validity of the PIRS ratings for use with teachers in evaluating the social validity of primary prevention plans. Following replication of these findings and additional inquiry in this area, we view the PIRS as potentially valuable in that it may provide a feasible, reliable tool for use in evaluating the social validity of primary prevention plans before intervention onset. This information may be used to shape plan revisions and possibly predict fidelity of program implementation, buy-in, and sustainability. For cases in which social validity ratings are low prior to plan implementation, this information can be used to revise school-wide plans that are viewed as targeting more appropriate goals, including more feasible procedures, and anticipating desired outcomes. By revising the plans prior to implementation, schools can save valuable resources and set the stage for high levels of treatment fidelity.

## References

- Elliott, S. N. (1988). Acceptability of behavioral treatments: A review of variables that influence treatment selection. *Professional Psychology: Research and Practice, 19*, 68–80.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*, 272–299.
- Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children, 71*, 149–164.
- Horner, R. H., Carr, E. C., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children, 71*, 165–179.
- Kazdin, A. E. (1980). Acceptability of alternative treatments for deviant child behavior. *Journal of Applied Behavior Analysis, 13*, 259–273.
- Lane, K. L. (2004). Academic instruction and tutoring interventions for students with emotional/behavioral disorders: 1990 to present. In R. B. Rutherford, M. M. Quinn, & S. R. Mathur (Eds.), *Handbook of research in emotional and behavioral disorders* (pp. 462–486). New York: Guilford Press.
- Lane, K. L., Kalberg, J. R., & Edwards, C. (2008). An examination of school-wide interventions with primary level efforts conducted in elementary schools: Implications for school Psychologists. In D. H. Molina (Ed.), *School Psychology: 21st century issues and challenges* (pp. 253–278). New York: Nova Science Publishers.
- Lane, K. L., Kalberg, J. R., & Menzies, H. M. (in press). *Developing schoolwide programs to prevent and man-*

- age problem behaviors: A step-by-step approach. New York: Guilford Press.
- Lane, K. L., Robertson, E. J., & Graham-Bailey, M. A. L. (2006). An examination of school-wide interventions with primary level efforts conducted in secondary schools: Methodological considerations. In T. E. Scruggs & M. A. Mastropieri (Eds.), *Applications of research methodology: Advances in learning and behavioral disabilities* (Vol. 19, pp 157–199). Oxford, UK: Elsevier.
- Lane, K. L., Robertson, E. J., & Wehby, J. H. (2002). *Primary Intervention Rating Scale*. Unpublished rating scale.
- Noell, G. H., Duhon, G. J., Gatti, S. L., & Connell, J. E. (2002). Consultation, follow-up, and implementation of behavior management interventions in general education. *School Psychology Review, 31*, 217–234.
- Noell, G. H., & Gresham, F. M. (1993). Functional outcome analysis: Do the benefits of consultation and prereferral interventions justify the costs? *School Psychology Quarterly, 8*, 200–226.
- Noell, G. H., Witt, J. C., Slider, N. J., Connell, J. E., Gatti, S. L., Williams, K. L., et al. (2005). Treatment implementation following behavioral consultation in schools: A comparison of three follow-up strategies. *School Psychology Review, 34*, 87–106.
- Reimers, T. M., Wacker, D. P., Cooper, L. J., & DeRaad, A. O. (1992). Clinical evaluation of the variables associated with treatment acceptability and their relation to compliance. *Behavioral Disorders, 18*, 67–76.
- Reimers, T. M., Wacker, D. P., & Koepl, G. (1987). Acceptability of behavioral treatments: A review of the literature. *School Psychology Review, 15*, 212–227.
- Tennessee Department of Education. (2006). *Report Card 2006*. Web site: <http://www.k-12.state.tn.us/rp/rcrd06/>
- VonBrock, M., & Elliott, S. N. (1987). The influence of treatment effectiveness information on the acceptability of classroom interventions. *Journal of School Psychology, 25*, 131–144.
- Witt, J. C., & Elliott, S. N. (1985). Acceptability of classroom intervention strategies. In T. R. Kratochwill (Ed.), *Advances in school psychology* (Vol. 4, pp. 251–288). Mahwah, NJ: Erlbaum.
- Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis, 11*, 203–214.

Date Received: March 17, 2008

Date Accepted: December 7, 2008

Action Editor: James DiPerna ■

Kathleen Lynne Lane, PhD, BCBA, is an associate professor in the Department of Special Education of Peabody College at Vanderbilt University. Her research program primarily focuses on designing, implementing, and evaluating multilevel, school-based interventions to prevent the development of learning and behavior problems for students at risk for EBD and remediate the deleterious effects of existing problems exhibited by students with EBD.

Jemma Robertson Kalberg, MEd, earned her master's degree in special education from Peabody College at Vanderbilt University. Her research interest focuses on supporting students who are at risk for and identified with EBD as well as designing, implementing, and evaluating school-wide positive behavior support plans with academic, behavior, and social components at the elementary, middle, and high school level.

Allison Leigh Bruhn, MA, is a doctoral student at Vanderbilt University. Her research interests include identifying characteristics of responders and nonresponders to Positive Behavior Support, improving classroom management through function-based interventions, and the relation between reading deficits and behavior problems.

Steven A. Driscoll, MEd is the coordinator of Project WRITE at Vanderbilt University. His research interests include school-wide behavior interventions, as well as focused learning interventions

Joseph Wehby, PhD, is an associate professor in the Department of Special Education of Peabody College at Vanderbilt University. His research interests include teacher–student interactions and academic interventions for students with EBD.

Stephen N. Elliott is the Dunn Family Professor of Educational and Psychological Assessment and Director of the Learning Sciences Institute at Vanderbilt University.