School Psychology Forum:
RESEARCH IN PRACTICE

VOLUME 3 • ISSUE 1 • PAGES 1–12 • Winter 2009

Goal Attainment Scaling: A Progress-Monitoring Tool for Behavioral Interventions

Gina Coffee
Loyola University Chicago
Corey E. Ray-Subramanian
Northern Illinois University

ABSTRACT: Within a response-to-intervention framework, monitoring students' progress is essential for determining whether additional intervention is needed. Although progress monitoring is most commonly considered in the context of improving academic outcomes, it can be equally useful for measuring student progress toward important behavioral goals. However, behavioral assessment methods such as systematic direct observation by a consultant or repeated administration of broadband standardized behavioral rating scales do not offer the specificity and sensitivity necessary for monitoring individual students' behavioral goals within a multitier service delivery system. One such method that can be used effectively for behavioral progress monitoring is goal attainment scaling (GAS). This article describes the development of GAS as a clinical evaluation tool and its current uses within school psychology practice. A case example and analysis are also presented to illustrate the utility of GAS as a progress-monitoring tool for determining whether a student is responding adequately to a behavioral intervention.

Monitoring student progress during an intervention is critical for determining whether the intervention is producing the intended results. More broadly, progress monitoring is a key component of data-based decision making in schools within a response-to-intervention (RTI) framework. At the core of an RTI framework is a service delivery system in which students who are not progressing at the expected rate are given evidence-based interventions to improve their performance (Glover & DiPerna, 2007). A key component of RTI service delivery involves student assessment and data-based decision making (Glover & DiPerna), which is typically utilized to improve *academic* outcomes for students. Within this focus on academic performance, progress monitoring involves (a) establishing academic goals and measuring student progress toward the goals, (b) providing opportunities for universal screening to identify students who may be at risk for academic problems, and (c) offering accountability evidence for stakeholders regarding the effectiveness of instructional programs and academic interventions (Shapiro, 2008).

Some scholars have acknowledged that RTI as a model of service delivery can be used to address students' social or behavioral functioning in addition to academic performance (e.g., Barnett et al., 2006; Glover & DiPerna, 2007), potentially to help identify students with emotional and behavioral

Correspondence concerning this article should be directed to Gina Coffee, School of Education, Loyola University Chicago, 820 N. Michigan Avenue #1146, Chicago, IL, 60611.

Copyright 2009 by the National Association of School Psychologists, ISSN 1938-2243

disabilities (Gresham, 2007). However, there has not been much discussion in the literature regarding appropriate progress-monitoring tools for behavioral issues within a multitiered RTI framework. Systematic direct observation by consultants (e.g., school psychologists, problem-solving team members) has been incorporated into multigate standardized behavioral screening systems such as the Systematic Screening for Behavioral Disorders (Walker & Severson, 1990) and as a tool for measuring a student's response to behavioral intervention (Barnett et al., 2006), but frequent use of outside observers for progress monitoring is likely too resource intensive to be used in an RTI service delivery framework.

To evaluate the effectiveness of Tier 1 universal positive behavior support (PBS) systems, Horner, Sugai, and Lewis-Palmer (2005) created a template for use within a general evaluation framework that stresses the importance of measuring changes in school practices and student behavior rather than relying only on counts of staff training events and participant satisfaction ratings. Horner et al. (2005) identify several assessment methods that may be used to monitor whether an implemented PBS initiative has resulted in improved social or academic outcomes for students. Examples of these methods include regularly collecting frequency counts of office discipline referrals, suspensions, and expulsions, in addition to annual measures of school climate and safety. Although these measures may be useful for long-term progress monitoring of school-wide behavioral goals, they lack the sensitivity to measure change in an individual student's behavior over a relatively short period.

One option for progress monitoring of behavioral goals for an individual student at Tier 2 or Tier 3 is goal attainment scaling (GAS). This approach can be used by classroom teachers or older students, themselves, to monitor progress during an intervention and does not necessarily require ongoing systematic observation by an outside consultant. In addition to being efficient and cost-effective, GAS can be designed to be more sensitive to small increments of change as compared to repeated administrations of standardized behavior rating scales such as the Behavior Assessment System for Children, Second Edition (Reynolds & Kamphaus, 2004) or Social Skills Rating System (Gresham & Elliott, 1990). For these reasons, GAS is a progress-monitoring tool worth consideration when determining how to measure a student's response to a behavioral intervention within a multitiered service delivery system.

The purpose of this article is to present GAS as an option for progress monitoring within a problem-solving consultation framework and an RTI service delivery system. We describe GAS's origins as an evaluation tool, the basic components of GAS, and how it can be applied to school-based behavioral interventions. A case example and analysis using GAS is presented to illustrate its application as a progress-monitoring tool for behavioral interventions. We conclude with a discussion of the main advantages and limitations of GAS that should be taken into consideration prior to its selection as a tool for monitoring students' progress toward behavioral goals.

GOAL ATTAINMENT SCALING

Kiresuk and Sherman (1968) developed GAS as a method to evaluate mental health services and community interventions. Although the historical and political context within which GAS was created is complex, it ultimately grew out of a desire to learn the impact of interventions and service delivery on individuals' outcomes (Kiresuk & Sherman, 1968) because the available evaluation tools (e.g., the Minnesota Multiphasic Personality Inventory) were not adequate for this purpose. Since its development, GAS has been applied across multiple disciplines (education, medicine, rehabilitation, and mental health and social services) in specific areas such as learning disabilities, penitentiary medicine, vocational rehabilitation counseling, rape counseling, and substance abuse (Kiresuk, Smith, & Cardillo, 1994).

GAS is a criterion-referenced tool that facilitates individualized goal setting and the evaluation of outcomes for individuals, groups, or whole organizations/systems (Kiresuk et al., 1994). It is amenable to different types of goals (e.g., behavioral, social, affective, or cognitive) and the use of either quantitative or qualitative measurement. The primary intervention agent (e.g., the teacher) first identifies a goal (or goals)

and corresponding indicators, ideally in collaboration with the client (e.g., the student) and other stakeholders (e.g., the school psychologist, other school personnel, and/or family members). Then, five outcome levels are specified, beginning with the expected outcome level (designated by a value of 0), and including levels that are somewhat more (+1) and less (–1) than expected, and much more (+2) and much less (–2) than expected. In its traditional form, at least three goals per client are typically identified, and a scale is developed for each; however, fewer or more goals may be identified. Goals may then be weighted according to priority, and a goal attainment score may be calculated using procedures delineated by Kiresuk et al. (1994). The goal attainment score is presented as a *T* score and represents a client's overall progress toward attaining the set goals. Goal attainment is usually evaluated at a predetermined follow up by an independent rater. Figure 1 is an example of a goal attainment scale for a sixth-grade student who exhibits behavioral difficulties during math and may benefit from Tier 2 interventions. An extension of this case example with an analysis will be presented later within a conjoint behavioral consultation framework.

Figure 1. Example of a goal attainment scale.

Level of Attainment	Scale 1 Out of Seat	Scale 2 Calling Out	Scale 3 Homework Completion
(–2) Much less than expected	Jack is out of his seat without teacher permission more than 12 minutes during math. [specify number of minutes]	Jack calls out an answer without teacher permission 5 or more times during math. [specify number of times]	Jack completes less than 60% of math assignments per week. [specify percent]
(–1) Somewhat less than expected	Jack is out of his seat without teacher permission 10–12 minutes during math.	Jack calls out an answer without teacher permission 4 times during math.	Jack completes 60–69% of math homework assignments per week.
(0) Expected level of outcome	Jack is out of his seat without teacher permission 7–9 minutes during math.	Jack calls out an answer without teacher permission 3 times during math.	Jack completes 70–79% of math homework assignments per week.
(+1) Somewhat more than expected	Jack is out of his seat without teacher permission 4–6 minutes during math.	Jack calls out an answer without teacher permission 2 times during math.	Jack completes 80–89% of math homework assignments per week.
(+2) Much more than expected	Jack is out of his seat without teacher permission 3 or fewer minutes during math. [specify number of minutes]	Jack calls out an answer without teacher permission 1 or fewer times during math.	Jack completes 90–100% of math homework assignments per week.
Comments	Monitor daily.	Monitor daily.	Monitor weekly.

Note. Adapted from Kiresuk, T. J., Smith, A., & Cardillo, J. E. (1994). *Goal attainment scaling: Applications, theory, and measurement.* Hillsdale, NJ: Erlbaum.

Reviews of the psychometric properties have not produced consistent results. Whereas reviews by Cytrynbaum, Ginath, Birdwell, and Brandt (1979) and Donnelly and Carswell (2002) suggest questionable psychometric properties, a recent review conducted by Schlosser (2004) presents a more promising picture. In terms of its limitations, reports of GAS's reliability have been mixed, with Kiresuk and Sherman (1968) reporting intraclass correlation coefficients (ICC) between ICC = .59 to ICC = .65 and Cytrynbaum et al. reporting interrater reliabilities between $r = .51$ and $r = .95$ but indicating caution in interpretation. Similarly, concurrent validity has been reported as "low to moderate" (Donnelly & Carswell, 2002, p. 88), and Cytrynbaum et al. (1979) found the evidence to establish the content and construct validity of GAS

"largely unconvincing" (p. 35). Nevertheless, evidence for the social validity (Schlosser) and clinical utility (Donnelly & Carswell) has been established. In addition, Schlosser indicated that the reliability and validity of GAS is best determined on a case-by-case basis and reported recommendations for reducing bias. Despite the apparent simplicity of GAS, the process of scaling goals is fraught with pitfalls, especially for the inexperienced user. Common problems the GAS user may encounter include clerical problems (e.g., omissions of crucial details on the scales), the use of jargon, vagueness, overlapping intervals, gaps between levels, multidimensional scales, and blank levels (Kiresuk et al., 1994). One way to avoid the use of jargon, vagueness, and multidimensionality is by creating SMART goals that are specific, measureable, attainable, realistic, and timely or tangible. In addition, with regard to overlapping intervals, gaps between

Figure 2. Example of a goal attainment scale with common errors.

Level of Attainment	Scale 1 Out of Seat	Scale 2 Calling Out	Scale 3 Homework Completion
(–2) Much less than expected	Jack is out of his seat 2 or more times during math. [specify number of minutes]	Jack calls out an answer without teacher permission 11 times during math.	Jack completes 30% or less of his homework assignments with at least 85% accuracy.
(–1) Somewhat less than expected		Jack calls out an answer without teacher permission 5–10 times during math.	
(0) Expected level of outcome	Jack is out of his seat 1 time during math.	Jack calls out an answer without teacher permission 5 times during math.	Jack completes 70% of his homework assignments with at least 85% accuracy.
(+1) Somewhat more than expected		Jack calls out an answer without teacher permission 2–4 times during math.	
(+2) Much more than expected	Jack is out of seat his 0 times during math.	Jack calls out an answer without teacher permission 1 time during math.	Jack completes 100% of his homework assignments with at least 85% accuracy.
Comments	Monitor daily.	Monitor daily.	

Note. Adapted from Kiresuk, T. J., Smith, A., & Cardillo, J. E. (1994). *Goal attainment scaling: Applications, theory, and measurement.* Hillsdale, NJ: Erlbaum.

levels, and blank levels, users should keep in mind that the rater must have a place to mark any instance of behavior only one time. Figure 2 illustrates common errors in the construction of a goal attainment scale.

An analysis of each scale illustrated in Figure 2 reveals a multitude of problems. First, with regard to scale 1, the choice of measurement (number of times Jack is out of his seat versus the total amount of time he is out of his seat during math) is not appropriate for measuring the target behavior. For example, it is possible that Jack meets the expected outcome of only being out of his seat one time during math. However, during that one time, he may be out of his seat for 1 minute, 10 minutes, 30 minutes, or the entire class period. In addition, the benchmarks lack specificity and become vague with the omission of "without teacher permission." As written, it suggests that raters should mark all instances of out-of-seat behavior, with or without teacher permission. Alternatively, if multiple raters exist, it may be that some raters would infer "without teacher permission" while others would not and would mark all instances of out-of-seat behavior. Finally, the blank levels on scale 1 decrease its sensitivity to small, yet possibly meaningful, changes in Jack's behavior.

Scale 2 illustrates the common problems of unequal and overlapping intervals, and all instances of behavior are not accounted for. Unequal intervals are problematic in that the unequal intervals produce misleading data that may be artificially inflated or depressed. For example, for Jack, movement from the –1 level to the 0 level, and from the 0 level to the +1 level, does not represent the same amount of change. Overlapping intervals are also problematic in that an instance of behavior can be recorded at multiple levels, thus defeating the purpose of a goal attainment scale. As presented in Figure 2, if Jack calls out an answer without teacher permission five times during math, then he is functioning both at the expected outcome level and the somewhat less-than-expected level. Finally, scale 2 does not account for all instances of calling out behavior, as the rater does not have a place to mark less than 1 instance of calling out or more than 11 instances.

Errors on scale 3 include multidimensionality, blank levels, gaps between levels, and vagueness. As shown, the indicators for Jack's progress toward the homework completion goal are accuracy and completion. Essentially, as written, this goal can only be rated when Jack turns in homework assignments with 85% accuracy. Therefore, his progress toward meeting one goal (completion) is contingent upon an entirely different behavior (accuracy). Depending on his performance in math (it may be that he consistently completes his homework with 75% accuracy), this goal may not be scoreable. If homework accuracy and completion are areas of concern, it is best to develop separate scales for each. The blank levels and the gaps between levels create additional problems with scoring. In particular, in contrast to the blank levels on scale 1, the blank levels on scale 3 may encourage raters to infer its content, which can lead to inaccurate scoring. Similarly, as written, some instances of behavior (e.g., 71–99%) cannot be scored due to the gaps between levels. Finally, scale 3 is vague, as it does not provide an indication of the academic subject or of how frequently the goal must be monitored (e.g., daily or weekly). As emphasized by Kiresuk et al. (1994), adequate training of relevant personnel is the ideal way to prevent such problems. An overview of GAS can be found at a website developed by Marson and Dran (2008). The site includes information that may assist in training and in the application of GAS, including specific guidelines for creating a goal attainment scale, examples of goal attainment scales, and an Excel template for developing a goal attainment scale.

APPLICATION IN SCHOOL PSYCHOLOGY

The use of GAS in the field of school psychology is not a new concept. As noted previously, soon after its inception, it was, and continues to be, applied in various disciplines. Given its application in both mental health and education, utilizing GAS in school psychology practice seems a logical step. As such, GAS has been used as an evaluation tool for district, school, class, and individual interventions and, importantly, has been found to be acceptable to teachers and mental health professionals (Cardillo, 1994; Carr, 1979; Robertson-Mjaanes, 1999). For example, Robertson-Mjaanes reported that most teachers in her sample endorsed GAS as "feasible" or "very feasible" in its use with students, and the teachers believed most professionals would find GAS "easy" or "very easy" to use in progress monitoring and intervention outcome evaluation. Similarly, special education teachers have found GAS to be time efficient (Carr, 1979). Examples of GAS's uses to date in school psychology include the evaluation of school psychological service departments (Maher & Barbrack, 1981), the measurement of accountability in intervention-based school psychology services (Barnett et al., 1999; Bonner & Barnett, 2004), and the evaluation of interventions via individual progress monitoring and intervention evaluation (Imich & Roberts, 1990; Mitchell & Cusick, 1998; Oren & Ogletree, 2000; Sladeczek, Elliott, Kratochwill, Robertson-Mjaanes, & Stoiber, 2001).

It should be noted that the application of GAS in school psychology is slightly different than its original use. That is, Kiresuk and Sherman (1968) originally developed GAS as a tool to evaluate intervention or service delivery outcomes at a specified follow-up interview, typically following intervention termination. In school psychology, however, GAS has been modified to facilitate repeated measurement across time, thereby offering a viable means to not only evaluate overall intervention outcomes, for example, but to monitor student progress on a daily or weekly basis. To accommodate this change, the descriptors on the original 5-point scale changed from a *much less than expected* (–2)/*much more than expected* (+2) continuum to a *worst*

possible outcome (–2)/*best possible outcome* (+2) continuum, where initial assessments (or baseline) are designated a rating of 0 or *no change in behavior/baseline*.

The Academic Competence Evaluation Scales (ACES; DiPerna & Elliott, 2000) and Outcomes PME: Planning, Monitoring, Evaluation (Stoiber & Kratochwill, 2001) are examples of commercially available instruments that incorporate GAS. The ACES is a standardized rating scale and intervention-planning tool that measures students' performance in the two primary domains of academic competence: academic skills and academic enablers (i.e., study skills, interpersonal skills, motivation, and engagement). The ACES teacher form includes 33 items on which teachers rate students' specific academic skills in the areas of reading/language arts, mathematics, and critical thinking in comparison with grade-level expectations at their school. This form also includes 40 items on which teachers rate the frequency with which students exhibit particular behaviors related to study skills, interpersonal skills, motivation, and engagement. An academic skills total raw score and an academic enablers total raw score, along with corresponding confidence intervals, can then be calculated from the ratings.

In addition to providing a standardized assessment of a student's academic skills and academic enablers, the ACES includes a Linking Assessment Results to Intervention section that utilizes GAS as part of a five-step problem-solving process that allows for the development and evaluation of an intervention intended to improve a student's academic skills and/or academic enabling behaviors. The first step in the process is to identify concerns related to the student's performance using the provided Behavior Classification Scheme, which includes the identification of strengths, performance problems, and acquisition problems in the academic skill domains (reading/language arts, mathematics, and critical thinking) and for each of the academic enablers (interpersonal skills, engagement, motivation, and study skills). The second step involves analyzing the concerns by identifying a target behavior or behaviors and considering the instructional environment in which the student is having difficulty. The third step is to plan the intervention by identifying an intervention goal, a desired behavior, and a general intervention strategy, and by creating a goal attainment scale for monitoring behavior change. The 5-point scale ranges from –2 (worst possible behavior) to +2 (best possible behavior) and includes spaces for writing descriptive criteria for each benchmark. The fourth step in the process is to implement the intervention utilizing the Academic Intervention Monitoring System (Elliott, DiPerna, & Shapiro, 2001), a related intervention guidebook and progress-monitoring tool that also includes goal attainment scaling. The final step in the ACES problem-solving process is to graph the goal attainment scale results on the provided form to evaluate the outcomes of the intervention.

Outcomes PME: Planning, Monitoring, Evaluating is another instrument that incorporates GAS. Essentially, Outcomes PME is a comprehensive tool that guides stakeholders through the process of problem identification; goal and benchmark setting; intervention planning; progress monitoring; and the evaluation of academic, social, and behavioral problems in school, home, and community settings. Outcomes PME can also be used to address issues related to language–communication development and physical–motor performance. Outcomes PME is unique in that it provides a means for interpreting information from multiple sources through a method called convergent evidence scaling. Unlike the original scale descriptors posited by Kiresuk and Sherman (1968), the goal attainment scale included in Outcomes PME is a 7-point scale, including seven benchmarks ranging from –3 (worst possible outcome; much worse) to +3 (best possible outcome; much improved) or 0 (extremely unacceptable competence level; skill not evident) to +6 (exceeding or reaching goal; competent). Scale selection (–3 to +3 or 0 to +6) depends on the target behavior.

Although GAS has widespread applications in school psychology, its use in monitoring students' behavioral progress within a school-based consultation framework (Sladeczek et al., 2001) is the application of particular interest for the purposes of this article. As illustrated by Sladeczek et al. (2001), GAS is especially useful in consultation because it can provide a means with which to monitor intervention progress and evaluate behavioral outcomes in the context of the consultation process. Within a consultation framework,

GAS can facilitate communication and collaboration among stakeholders. In addition, when data are collected within a single-case design framework (i.e., providing information across baseline and interventions conditions), GAS can also provide useful data that inform decision making regarding intervention modifications, for example.

Given the individualized nature of GAS, it can be applied to a multitude of problems commonly encountered in the school setting. For example, the goals Sladeczek et al. (2001) delineated in their presentation of a conjoint behavioral consultation case surrounded the particular challenging behaviors of screaming and clinging, behaviors that were of concern to the preschool student's teacher and parent. However, GAS has also been used to evaluate intervention outcomes of students with disruptive behaviors (Imich & Roberts, 1990), autism (Oren & Ogletree, 2000), and traumatic brain injuries (Mitchell & Cusick, 1998). Given its versatility, a goal attainment scale can be developed for students of any age and for virtually any behavior of concern.

CASE EXAMPLE

To illustrate an application of GAS, we created the following fictitious case example. Consider the case of Jack, the sixth-grade student exhibiting challenging behaviors in math. Jack attends Windsor Middle School and has two primary teachers for his core academic subjects. Within this team-teaching arrangement, Jack has Ms. Stockton for language arts and science and Mr. Porter for math and social studies. This year Jack is earning Bs in most of his classes, but currently has a C- in math. Mr. Porter sought consultation from the school psychologist regarding Jack's performance in math.

Mr. Porter reports that Jack appears to understand the content being taught in the class, but his grade has been significantly affected by his sporadic homework completion. When Jack does submit a homework assignment on time, it is often incomplete or contains many careless errors. Jack has said that he "doesn't get the point of homework" and finds it "boring."

In addition to homework completion, Mr. Porter is, to a lesser degree, concerned about Jack's in-class behaviors that are disruptive to other students and detract Jack's attention from the math content being taught. Specifically, Mr. Porter said that Jack often blurts out answers without raising his hand or being called on first. He has also observed that Jack gets out of his seat frequently to sharpen his pencil unnecessarily, look out the window, or talk to other students. Mr. Porter indicated that several students seem to display these disruptive behaviors in math. When asked about these behaviors, Jack denies being aware that he and his peers have been disruptive in class.

The school psychologist observed Jack in several of his classes and interviewed Ms. Stockton and Mr. Porter to obtain their perspectives on Jack's behavior in his various classes. Although Ms. Stockton reported that Jack has missed a few homework assignments in language arts and science, she said that he does not typically engage in disruptive behavior in class. Her report was consistent with the school psychologist's direct observations and with Mr. Porter's report of Jack's behavior in social studies, which is noticeably better than his behavior in math.

During an interview with Jack's mother and stepfather, they stated that they were not aware of Jack's disruptive behavior in math nor did they realize how infrequently he has been turning in his homework. Jack's parents said that they are both at work until the early evening, so they are not present when Jack and his older sister are doing their homework after school. They describe Jack as an "active boy with a lot of energy" who likes to play basketball and video games and enjoys spending time with his parents and sister. Jack's parents stated that they were willing to meet with Jack's teachers to talk about Jack's homework completion and behavior in math.

The school psychologist held a meeting with Jack, his parents, and his teachers to discuss Jack's homework completion and his disruptive behavior in math class and to establish performance goals and a plan for helping Jack reach these goals. Jack's sister also attended the meeting, as his parents thought it might be helpful to include her given that she and Jack typically work on their homework together after school. During the meeting, Jack, his parents, and his teachers prioritized their concerns; analyzed the identified problems using data collected by Jack's parents, teacher, and school psychologist; and agreed to begin by implementing an evidence-based, individualized intervention at home and school to address Jack's homework completion. To facilitate generalization, and given that Jack has missed some homework assignments in all of his classes, Ms. Stockton and Mr. Porter thought it might be beneficial to implement the intervention in all of Jack's classes. In addition to developing an individualized intervention for homework completion, Mr. Porter expressed an interest in implementing a class-wide intervention to address the disruptive behaviors of Jack and his peers.

Jack's behavioral progress in response to both interventions can be monitored by GAS. The GAS worksheet and progress-monitoring data for homework completion in math are presented in Figures 3, 4, and 5.

Figure 3. Goal attainment scale worksheet.

Dates: 1/29/08–4/7/08 School: Windsor Middle School Student: Jack Teacher: Mr. Porter

Goal attainment scaling (GAS) provides a method for quantifying teachers' reports of treatment progress about a target behavior and problem situation. The basic elements of a GAS are a 7-point scale ranging from a –3 to +3 and descriptions of the target behavior and problem situation. By using the numerical ratings for each of the 7 descriptive categories of behavioral functioning, you should be able to provide a weekly report of treatment progress. Please place a mark in the box that corresponds with the benchmark your student meets each week.

+3																
+2																
+1																
0																
–1																
–2																
–3																
Week	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16

Target Behavior: Jack completes at least 90% of math homework assignments per week.

(+3) (Best Possible Outcome; Much Improved): Jack completes at least 90–100% of math homework assignments per week.

(+2): Jack completes 80–89% of math homework assignments per week.

(+1): Jack completes 70–79% of math homework assignments per week.

(0) (No Change in Behavior; Baseline): Jack completes 60–69% of math homework assignments per week.

(–1): Jack completes 50–59% of math homework assignments per week.

(–2): Jack completes 40–49% of math homework assignments per week.

(–3) (Worst Possible Outcome; Much Worse): Jack completes less than 40% of math homework assignments per week.

Figure 4. Progress monitoring data collected by Mr. Porter for math.

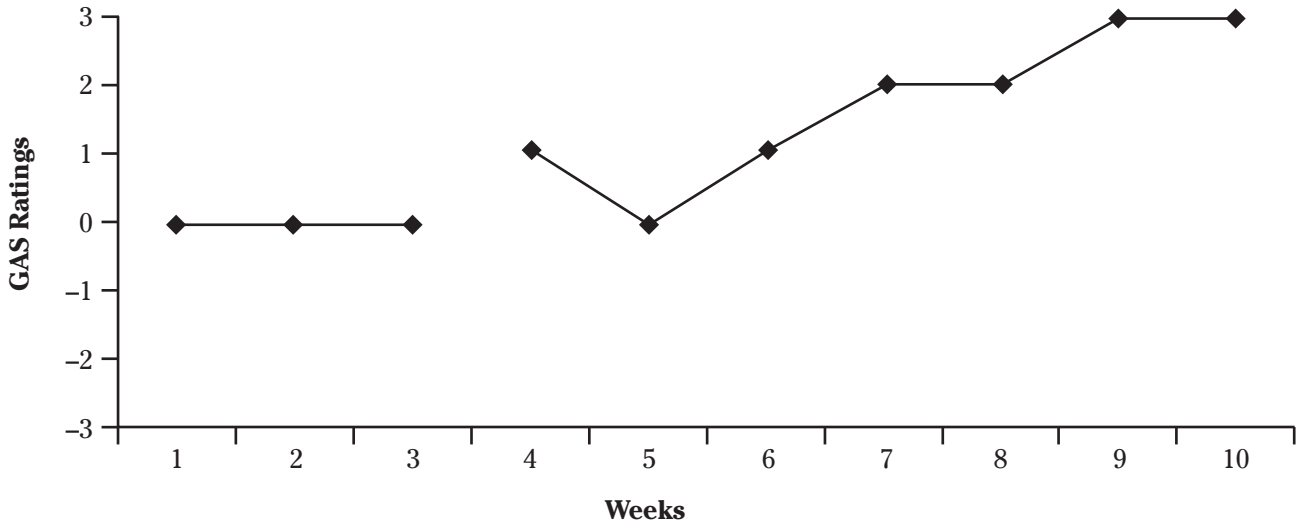
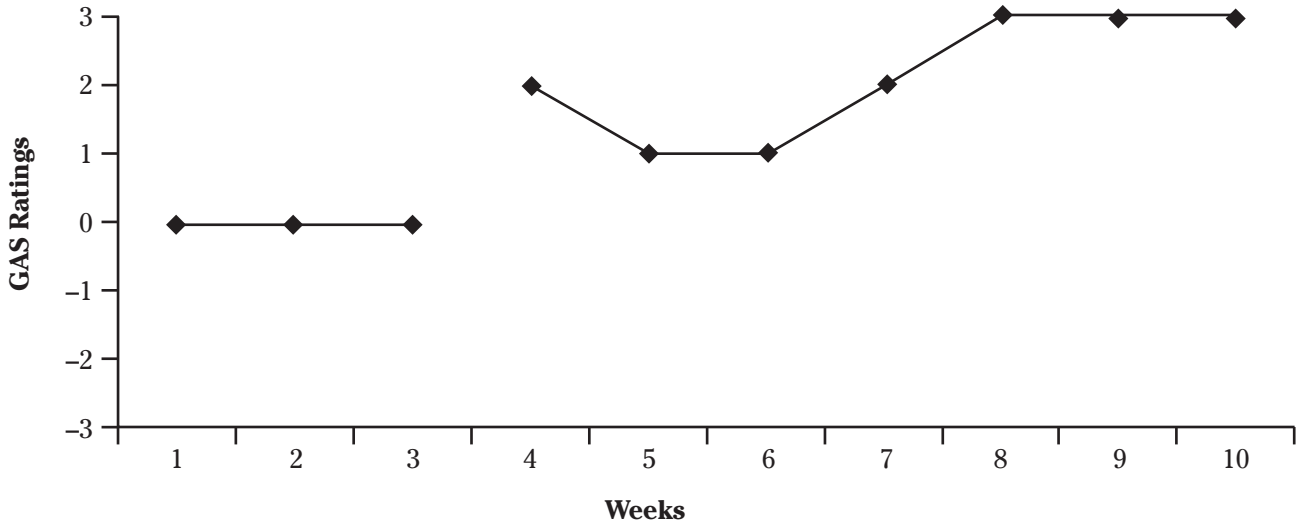


Figure 5. Self-monitoring data collected by Jack for math.



The team first identified a goal and corresponding benchmarks along a 7-point scale. They determined that Jack's teachers would implement the selected intervention daily during core academic subjects and that Jack's progress would be monitored individually by each teacher and Jack (self-monitoring) each week. Using permanent products (completed homework assignments), Mr. Porter and Ms. Stockton kept track of Jack's homework completion in their grade books and transferred the weekly percentage of completed homework assignments to the GAS worksheet. Jack kept track of the number of homework assignments he completed and turned in to his teachers in his assignment notebook and transferred the weekly number to a separate self-monitoring GAS worksheet. Jack and his teachers met briefly on Friday afternoons to convert the data Jack collected to percentages, to compare the weekly percentages obtained by Jack and his teachers, and to discuss Jack's progress. Jack was praised for monitoring his behavior and received an additional reward when he and his teachers' ratings were within three percentage points of each other. The school psychologist, Jack's teachers, and Jack's family kept in regular contact to discuss Jack's progress

and issues related to intervention implementation and treatment integrity.
After 7 weeks of intervention implementation, the school psychologist, Jack, his family, and his teachers met to evaluate the overall outcomes of the intervention. The data presented in Figures 4 and 5 represent the progress Jack made in completing his homework for math. Although discrepancies in ratings were noted in weeks 4, 5, and 8, the data indicated an upward trend in homework completion, and Jack met his goal. Data from his other courses revealed similar progress. Jack, his teachers, and his family reported positive perceptions of the consultation process and the data collection procedures, and developed a plan for maintenance.

IMPORTANT CONSIDERATIONS

The main advantage of GAS in school psychology is its clinical utility. That is, in practice GAS provides school personnel (and sometimes students themselves) with a viable option for measuring students' progress in response to behavioral Tier 1, Tier 2, and Tier 3 interventions. Broadband standardized behavior rating scales typically do not measure the target behavior meant to have an impact by the selected intervention and, consequently, are not sensitive enough to monitor progress over time. In contrast, GAS allows for the development of timely and individualized goals specific to the behavior of concern, as well as benchmarks surrounding those goals. As such, GAS's sensitivity as a progress-monitoring tool is enhanced. In addition, GAS can be used to monitor multiple behaviors of concern (Kiresuk et al., 1994; Stoiber & Kratochwill, 2001), though Stoiber and Kratochwill advise no more than two goal attainment scales per intervention.

Consistent with its clinical utility, GAS is particularly advantageous in problem-solving consultation, a process that often involves multiple change agents and multiple sources of data. As noted by Sladeczek et al. (2001), the use of GAS can promote communication and collaboration among the persons involved in the consultation process and has been shown to be a user-friendly and acceptable form of data collection and progress monitoring (Cardillo, 1994; Carr, 1979; Robertson-Mjaanes, 1999). Essentially, using GAS prompts stakeholders to work together to develop behavioral goals and benchmarks, monitor students' progress toward meeting those goals, and evaluate intervention outcomes using a common metric.

The primary limitations of GAS are its questionable psychometric properties (Cytrynbaum et al., 1979; Donnelly & Carswell, 2002). In response to limitations in reliability and validity, Schlosser (2004) has recommended that users of GAS establish reliability and validity on a case-by-case basis. In addition, he has suggested the following methods for reducing bias: (a) conducting adequate team training of those involved in the GAS process, (b) collecting subjective and objective data in the determination of attainment levels, (c) operationalizing definitions of outcome criteria, (d) utilizing an independent observer in the evaluation of goal attainment, and (e) ensuring adequate treatment integrity.

CONCLUSION

GAS is a useful tool for monitoring students' progress toward behavioral goals and for facilitating data-based decisions regarding behavioral interventions. With proper training in goal setting and scale construction, GAS can be incorporated into a multitier service delivery system designed to prevent behavioral problems and appropriately address behavioral concerns.

REFERENCES

Barnett, D. W., Daly III, E. J., Hampshire, E. M., Hines, N. R., Maples, K. A., Ostrom, J. K., et al. (1999). Meeting performance-based training demands: Accountability in an intervention-based practicum. *School Psychology Quarterly, 14,* 357–379.

Barnett, D. W., Elliott, N., Wolsing, L., Bunger, C. E., Haski, H., McKissick, C., et al. (2006). Response to intervention for young children with extremely challenging behaviors: What it might look like. *School Psychology Review, 35,* 568–582.

Bonner, M., & Barnett, D. W. (2004). Intervention-based school psychology services: Training for child-level accountability; preparing for program-level accountability. *Journal of School Psychology, 42,* 23–43.

Cardillo, J. E. (1994). Goal setting, follow up, and goal monitoring. In T. J. Kiresuk, A. Smith, & J. E. Cardillo (Eds.), *Goal attainment scaling: Application, theory, and measurement* (pp. 39–60). Hillsdale, NJ: Erlbaum.

Carr, R. A. (1979). Goal attainment scaling as a useful tool for evaluating progress in special education. *Exceptional Children, 46,* 88–95.

Cytrynbaum, S., Ginath, Y., Birdwell, J., & Brandt, L. (1979). Goal attainment scaling: A critical review. *Evaluation Quarterly, 3,* 5–40.

DiPerna, J. C., & Elliott, S. N. (2000). *ACES: The Academic Competence Evaluation Scales.* San Antonio, TX: The Psychological Corporation.

Donnelly, C., & Carswell, A. (2002). Individualized outcome measures: A review of the literature. *Canadian Journal of Occupational Therapy, 69,* 84–94.

Elliott, S. N., DiPerna, J. C., & Shapiro, E. S. (2001). *AIMS: The Academic Intervention Monitoring System.* San Antonio, TX: The Psychological Corporation.

Glover, T. A., & DiPerna, J. C. (2007). Service delivery for response to intervention: Core components and directions for future research. *School Psychology Review, 36,* 526–540.

Gresham, F. M. (2007). Response to intervention and emotional and behavioral disorders: Best practices in assessment for intervention. *Assessment for Effective Intervention, 32,* 214–222.

Gresham, F. M., & Elliott, S. N. (1990). *Social Skills Rating System (SSRS).* Circle Pines, MN: American Guidance Service.

Horner, R., Sugai, G., & Lewis-Palmer, T. (2005). *School-Wide Positive Behavior Support Evaluation Template.* Eugene, OR: University of Oregon. Retrieved May 31, 2008, from http://www.pbis.org/tools.htm

Imich, A., & Roberts, A. (1990). Promoting positive behavior: An evaluation of a behavior support project. *Educational Psychology in Practice, 5,* 201–209.

Kiresuk, T. J., & Sherman, R. E. (1968). Goal attainment scaling: A general method for evaluating community health programs. *Community Mental Health Journal, 4,* 443–453.

Kiresuk, T. J., Smith, A., & Cardillo, J. E. (1994). *Goal attainment scaling: Applications, theory, and measurement.* Hillsdale, NJ: Erlbaum.

Maher, C. A., & Barbrack, C. R. (1981). Evaluating the individual counseling of conduct problem adolescents: The goal attainment scaling method. *Journal of School Psychology, 22,* 285–297.

Marson, S. M., & Dran, D. (2008). *Goal attainment scaling.* Pembroke, NC: The University of North Carolina at Pembroke. Retrieved May 10, 2008, from http://www.marson-and-associates.com/GAS/GAS_index.html

Mitchell, T., & Cusick, A. (1998). Evaluation of a client-centered pediatric rehabilitation program using goal attainment scaling. *Australian Occupational Therapy Journal, 45,* 7–17.

Oren, T., & Ogletree, B. T., (2000). Program evaluation in classrooms for students with autism: Student outcomes and program processes. *Focus on Autism and Other Developmental Disabilities, 15,* 170–175.

Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavior Assessment System for Children, Second Edition (BASC-2).* Circle Pines, MN: American Guidance Service.

Robertson-Mjaanes, S. L. (1999). *An evaluation of goal attainment scaling as an intervention monitoring and outcome evaluation technique.* Unpublished doctoral dissertation, University of Wisconsin-Madison.

Schlosser, R. W. (2004). Goal attainment scaling as a clinical measurement technique in communication disorders: A critical review. *Journal of Communication Disorders, 37,* 217–239.

Shapiro, E. S. (2008). Best practices in setting progress monitoring goals for academic skill improvement. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V (pp. 141–158).* Bethesda, MD:

National Association of School Psychologists.
Sladeczek, I. E., Elliott, S. N., Kratochwill, T. R., Robertson-Mjaanes, S., & Stoiber, K. C. (2001). Application of goal attainment scaling to a conjoint behavioral consultation case. *Journal of Educational & Psychological Consultation, 12,* 45–58.
Stoiber, K. C., & Kratochwill, T. R. (2001). *Outcomes PME: Planning, monitoring, evaluating.* San Antonio, TX: The Psychological Corporation.
Walker, H. M., & Severson, H. (1990). *Systematic Screening for Behavioral Disorders (SSBD).* Longmont, CO: Sopris West.