Goal Attainment Scaling: An Efficient and Effective Approach to Monitoring Student Progress

Andrew T. Roach • Stephen N. Elliott

Then, this article is for you!

Goal attainment scales (GAS) provide an individualized, criterion-referenced approach to describing changes in the performance of students and can be very useful in documenting changes involving academic and social behaviors.

The Original Goal Attainment Scaling

Kiresuk and Sherman (1968) developed the original GAS method for use in evaluating the effectiveness of mental health services at the Hennepin County Mental Health Center in Minnesota. They were interested in identifying an alternative to the variety of rating scales, projective techniques, and clinical interviews used by clinicians to document client change. The original conceptualization and implementation of GAS resulted in a technique that not only promoted individualized, idiographic measurement of clients’ actual progress towards treatment goals; but also allowed for comparisons between different therapists and treatment programs (Kiresuk, 1994).

Although there has been substantial investigation and implementation of GAS in a variety of mental health and medical fields over the past 35 years, there has been less extensive research and application of GAS by school psychologists and special educators. For example, GAS was not included in a recent review of assessment methods commonly used by school psychologists (Wilson & Reschly, 1996). Fuchs’s chap-
This lack of familiarity with GAS as an assessment technique is unfortunate because of its utility as an element of multidimensional behavioral assessment (Shapiro & Kratochwill, 2000). Within the continuum of behavioral assessment developed by Shapiro and Browder (1990), student GAS self-ratings can function as either self-monitoring (a form of direct assessment, completed as behavior occurs) or self-report (a less direct measure of an individual’s perception of their behavior). GAS ratings completed by teachers and other adults function as informant reports, which are also indirect measures because they represent an observer’s retrospective perceptions of behavior. Moreover, because of its emphasis on operationalizing target behaviors and on-going (i.e., time-series) evaluation of academic or behavioral progress, GAS is a particularly useful tool for monitoring students’ progress and for verifying the need for additional support or intervention (Elliott, DiPerna, & Shapiro, 2001).

Although psychologists and special educators who embrace the concept of behavioral assessment may prefer low-inference measures such as CBM and direct observations, the rejection of less direct methods of assessment may limit their abilities to efficiently collect meaningful and useful data. As noted by Shapiro and Kratochwill (2000):

It is especially important to realize that data collected from one method are not inherently better than data collected from others. That is, data obtained through an indirect method from a parent (such as a rating scale) are not “less true” than data obtained by directly observing a student within a natural setting. The key to good assessment is to find conceptual links and relationships between methods and modalities of assessment. Each form of behavioral assessment contributes potentially unique ele-
By contributing additional and potentially unique information regarding students’ academic needs and progress, GAS ratings represent an important “piece of the assessment puzzle.”

**An Alternate Approach to Goal Attainment Scaling for Use in Educational Settings**

Drs. Thomas Kratochwill and Stephen Elliott used a modified version of GAS as part of a U.S. Department of Education research project entitled “An Experimental Analysis of Teacher/Parent Mediated Interventions for Preschoolers with Behavioral Problems.” Robertson-Mjannes (1999) outlined the differences between Kratochwill and Elliott’s modified GAS and the original approach developed by Kiresuk and Sherman. Beyond its utility as a follow-up measure of clinician or program effectiveness, Kratochwill and Elliott conceptualized GAS as a forma-
Case Examples From Current Research

The following cases were gathered as part of an investigation of the utility of the Academic Competence Evaluation Scales (ACES; DiPerna, & Elliott, 2000) and Academic Intervention Monitoring System (AIMS; Elliott et al., 2001) for developing and monitoring prereferral interventions in elementary school settings. To protect the confidentiality of the participants, we have changed any identifying information about the cases.

Case #1: Jacob

Jacob is an 8-year-old boy who is in second grade at Red Hawk Elementary School. Jacob was originally referred to the student assistance team as a kindergartner due to concerns about his ability to remember and retaining basic concepts skills. As a kindergartner, Jacob was referred by his teacher to the student assistance team—a small group of general and special education teachers, support staff, and administrators—for help in addressing her concerns about his ability to remember and retain basic concepts and skills.

The results of a psychoeducational evaluation, conducted when Jacob was a first grader, indicated he had average intelligence and academic skills in all areas. After consultation between his parents and school staff, it was decided Jacob should repeat first grade, and he was referred to a physician for an ADHD evaluation.

Early in second grade, Jacob’s teacher referred him to the student assistance team because of concerns with attention and academic performance, particularly in the area of reading and language arts. His teacher reported having used a variety of strategies to improve his performance including using preferential seating, modified materials and assignments, and pretesting/teaching important concepts. A consideration of work samples and a teacher-completed ACES rating scale resulted in the identification of the following target behavior: To increase Jacob’s ability to read second-grade vocabulary (i.e., sight words).

Jacob’s initial performance in reading second-grade words was at a level of less than 50% accuracy. Interventions included guided practice of text and sight word reading, frequent checks of Jacob’s independent work, and monitoring his progress through informal assessments with a goal of achieving at least 80% accuracy in the target behavior. Jacob’s teacher and members of the student assistance team collaborated to construct a GAS rating scale (see Figure 2) for monitoring Jacob’s progress in reading grade-level text.

Data on Jacob’s progress toward the target behavior was gathered periodically (via teacher-completed GAS rating) for approximately 4 weeks. Examination of the GAS ratings demonstrated improvements on the target behavior on 40% of the assessment dates. Subsequent consultation with Jacob’s teacher indicated that he had made moderate overall progress in reading during the assessment period and that additional practice opportunities and frequent informal assessments appeared to be more successful intervention practices than monitoring his independent work.

Case #2: Wilton

Wilton is a 9-year-old boy who is in fourth grade at Green Meadows Elementary School. Prior to this year, Wilton had never been referred to the student assistance team. Wilton’s third-grade teacher referred him to the student assistance team because of her concerns about his reading and writing skills, which were both at least one grade below grade level. His teacher reported having used a variety of strategies to improve his performance such as reducing the length of assignments, peer tutoring, one-on-one tutoring outside of his reading group, and giving him extended time to complete his written work. A consideration of work samples and a teacher-completed ACES rating scale resulted in the identification of the following target behavior: To increase Wilton’s ability to identify the main idea of a story passage.

At the beginning of progress monitoring, Wilton could identify the main idea from a grade-level passage approximately 60% of the time. Interventions included reducing the reading level of texts, working with a lower reading group, allowing Wilton to take home reading materials, and monitoring his progress through informal assessments with a goal of achieving at least 90% accuracy in the target behavior. Wilton’s teacher and members of the student assistance team collaborated to construct a GAS rating scale (see Figure 3) for monitoring Wilton’s progress in reading grade-level text.

Data on Wilton’s progress towards the target behavior was gathered periodically (via teacher-completed GAS rating) for approximately 6 weeks. Examination of the GAS ratings demonstrated improvements on the target behavior on 10% of the assessment dates with no overall improvement in the target behavior. Subsequent consultation with the classroom teacher indicated that the recommended interventions had been seen as acceptable and were implemented with integrity, and also that Wilton’s overall reading performance had not improved during the assessment period. In light of the assessment results suggesting resistance to intervention, the student assistance team referred Wilton for additional evaluation with “confidence that testing may be indicated.”
Using GAS as a repeated measure of students’ progress required changes in the original scaled descriptions developed by Kiresuk and Sherman. Although Kratochwill and Elliott maintained the 5-point scale, the modified GAS included initial assessments to establish a baseline, which was subsequently assigned a score of “0.” Thus, the new scale levels ranged from the best possible outcome (+2) to the worst possible outcome (-2). In addition, the modified GAS allowed for ratings of both over- and underattainment of behavioral or academic objectives (Kratochwill, Elliott, & Rotto, 1995).

**Figure 2. Goal Attainment Scale Ratings for the Case of Jacob**

<table>
<thead>
<tr>
<th>Goal Attainment Scale Worksheet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student: Jacob Jones</td>
</tr>
<tr>
<td>Date of Birth: 8-17-1984</td>
</tr>
<tr>
<td>Today’s Date: 10-19-02</td>
</tr>
<tr>
<td>Grade: 2nd</td>
</tr>
<tr>
<td>Gender: X Male</td>
</tr>
<tr>
<td>School: Red Hawk Elementary</td>
</tr>
</tbody>
</table>

**Target Behavior(s):**
- Jacob will improve his accuracy in reading 2nd grade vocabulary words.

**Goal Attainment Scale with Descriptive Criteria for Monitoring Academic Change**

+2 Jacob reads 2nd grade vocabulary with greater than 75% accuracy.
+1 Jacob reads 2nd grade vocabulary with 56%-75% accuracy.
0 Jacob reads 2nd grade vocabulary with 45%-55% accuracy.
-1 Jacob reads 2nd grade vocabulary with 30%-44% accuracy.
-2 Jacob reads 2nd grade vocabulary with less than 30% accuracy.

**Graph of Academic Progress**

Previous Research on Educational and Clinical Applications of GAS Ratings

Maher’s (1982) evaluation of the Goal-Oriented Approach to Learning (GOAL) program indicated that students who completed GAS self-ratings achieved higher goal attainment scores than peers who did not self-evaluate (58.0 to 33.3), and were more likely to achieve the “Expected Outcome” level for their goals (88% to 0%). A follow-up investigation (Maher, 1983) demonstrated that student and adult (i.e., guidance counselor) GAS ratings of school adjustment and achievement “did not differ significantly in terms of estimating goal attainment” (p. 533).

Single-case studies utilizing GAS ratings have been conducted with students with autism (Oren & Ogletree, 2000), traumatic brain injuries (Mitchell & Cusick, 1998), and conduct problems (Sladeczek, Elliott, Kratochwill, ...
Robertson-Mjaanes, & Stoiber, 2001). In each case, consultants and consultees reported GAS ratings helped to clarify intervention goals and document student progress towards those goals.

In their study of family therapy efficacy, Santa-Barbara et al. (1977) reported that GAS was “the only clinical outcome measure that [was] consistently related to measures of intellectual functioning and academic achievement obtained at follow-up ratings” (p. 54).

Follow-up studies on family therapy clients conducted by Woodward et al. (1981) demonstrated that GAS ratings...
were a more effective predictor than other outcome measures.

**GAS Ratings Are Time-Efficient and User-Friendly**

Although many researchers and policymakers consider CBM the “gold standard” for assessing students’ academic performance, practitioners often view these techniques as labor-intensive and unmanageable in general, and inappropriate for students in middle school or higher who are working on more advanced academic skills. For example, Fuchs and Elliott (1997) asserted CBM can be “time-consuming for teachers or school psychologists to collect, score, and analyze” (p. 228). Foegen, Espin, Allinder, and Markell (2001) suggested the following reasons for teacher’s reluctance to use CBM methods: (a) lack of time to implement and score CBM probes, (b) insufficient mastery of the skills needed to use CBM, and (c) practitioners’ beliefs about the validity of CBM.

Acceptability research with school and mental health professionals suggests that GAS is perceived as highly acceptable and useful in a variety of settings. Robertson-Mjaanes (1999) asked professionals (including teachers, n = 36) to rate the ease of use and acceptability of GAS. The results indicate that most teachers (94%) endorsed GAS as “very feasible” or “feasible” to use with their students. Moreover, Robertson-Mjaanes found that 70% of the teachers believed that most professionals would find GAS “easy” or “very easy” to use for monitoring intervention progress and outcomes.

According to Carr (1979), special education teachers who participated in an evaluation project described the time necessary to construct an individual GAS as “worthwhile” and valuable for specifying goals and facilitating the inclusion process. Cardillo (1994) reported that the existing research on the implementation of GAS in mental health settings suggests study participants experienced the GAS process as (a) “an unexpected, but beneficial learning experience, with respect to both goal setting and treatment planning”; and (b) one that lead “to an enhanced awareness of and respect for documentation requirements” (p. 59). In an evaluation of two adult and family literacy programs, Sherow (2000) identified the following positive outcomes of using GAS ratings:

- Encouraged cooperative goal setting between staff and clients.
- Facilitated the development of individualized programs for clients.
- Clarified and communicating program goals.

---

**Table 2. Advantages and Disadvantages to Utilizing GAS Ratings to Monitor Students’ Academic Growth**

<table>
<thead>
<tr>
<th>Advantages of GAS Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Time efficient</td>
</tr>
<tr>
<td>• Personalized/ individual assessment</td>
</tr>
<tr>
<td>• Conceptually consistent with behavioral consultation</td>
</tr>
<tr>
<td>• Requires minimal skills to collect data</td>
</tr>
<tr>
<td>• Nonintrusive assessment method</td>
</tr>
<tr>
<td>• Can be used as a self-assessment</td>
</tr>
<tr>
<td>• Can be used by multiple informants across settings (e.g., home, school, community)</td>
</tr>
<tr>
<td>• Can be used repeatedly to monitor perceptions of intervention progress</td>
</tr>
<tr>
<td>• Can be used to document perceptions of intervention outcomes</td>
</tr>
<tr>
<td>• Inexpensive</td>
</tr>
<tr>
<td>• Requires minimal skills to interpret data</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Disadvantages of GAS Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Limited published, empirical research on the school-based use of the method</td>
</tr>
<tr>
<td>• Subjective summary of observations collected over time</td>
</tr>
<tr>
<td>• Not norm-referenced</td>
</tr>
<tr>
<td>• Guidelines for interpretation are determined by parties involved with the intervention, thus subject to bias</td>
</tr>
<tr>
<td>• Global (i.e., less discrete) accounting of behavior</td>
</tr>
</tbody>
</table>
In addition to measuring academic knowledge and skills, educators may also choose to use GAS ratings to measure students’ progress in social behaviors or academic enablers.

- Provided both formative and summative evaluations of program effectiveness. Table 2 provides a summary of the advantages and disadvantages to using GAS ratings to monitor students’ academic growth.

Final Thoughts

In addition to measuring academic knowledge and skills, educators may also choose to use GAS ratings to measure students’ progress in social behaviors or academic enablers (i.e., “attitudes and behaviors that allow a student to participate in, and ultimately benefit from academic instruction”; DiPerna & Elliott, 2002, p. 294). A substantial body of research exists which demonstrates the relationship between students’ social behavior and their academic performance (DiPerna, Volpe, & Elliott, 2002; Malecki & Elliott, 2002; Wentzel, 1993). Figure 4 provides an example of a complete GAS rating for Ellen, a student with a social behavior that directly contributes to her educational success. Readers should note how this GAS rating scale operationalizes a group of behaviors (or “response class”) that pertains to the educational goal: successfully collaborating with classmates on a group project.

According to Smith and Cardillo (1994), GAS ratings are especially suited “as a measure of one important aspect of outcome: treatment- (or instruction-) induced change. If interest should be in the change that has been produced during (an intervention)...GAS can be recommended as a particularly sensitive method of measuring that change” (p. 272). The initial results from our case studies of prereferral interventions suggest GAS ratings can provide efficient and accurate assessments of students’ academic and behavior progress. Moreover, in our work designing academic interventions in collaboration with teachers, we have found GAS ratings to be a user-friendly and meaningful way for conceptualizing and communicating change over the course of a multiweek intervention. GAS ratings are relatively easy-to-use and understand, making them particularly appropriate tools for consulting and collaborating with general education teachers, paraprofessionals, and parents.

GAS can be utilized to meet a variety of assessment needs. As part of self-monitoring intervention, students themselves can complete GAS ratings. Conversely, educators and parents can use GAS ratings to record their observations and perceptions of students’ academic and behavior progress. In comparison to direct observation or CBM probes, teachers and paraprofessionals may find GAS ratings easier to collect within the complex ecologies of their classrooms. In addition, GAS promotes clearly operationalized intervention goals and on-going (i.e., time-series) evaluation of student progress, making it a potentially useful tool for special educators and school psychologists working within a responsiveness-to-intervention (RTI) model of special education identification.

References


Andrew T. Roach, Research Coordinator, Center for Assessment and Intervention Research; and Stephen N. Elliott, Dunn Professor of Assessment, Department of Special Education, Peabody College, Vanderbilt University, Nashville, Tennessee.

Address all correspondence concerning this article to Andrew T. Roach, Center for Assessment and Intervention Research, 405 Wyatt, 1930 South St., Nashville, TN 37203 (e-mail: roach@vanderbilt.edu)

TEACHING Exceptional Children, Vol. 37, No. 4, pp. 8-17.

Copyright 2005 CEC.