

Standardized Assessments for the Management of Children with Motor Disorders

GOAL ATTAINMENT SCALE (GAS)

Assessment Authors: Kiresuk and Sherman 1968, Turner-Stokes 2009

Description of Assessment

Purpose: The Goal Attainment Scale (GAS) is a functional scale used to measure the extent of progress towards individual goals during the course of an intervention. Each subject has their own outcome measures but they are scored in a standardized way to allow statistical analysis. Scale items are based on current and expected levels of performance.

Rating System

Several steps are required for the effective use of the GAS:

- Specific, measurable, attainable, realistic, and time related individual goals are set.
- Expected outcomes are defined for each achievement level before treatment is started.
- The goals are weighted for importance and difficulty.
- The baseline (current level) achievement is rated.
- Treatment is performed.
- The post-treatment achievement (performance level) is rated.
- The T-score is calculated.

Baseline scores (current level) are usually rated -1, unless there is no plausible worse condition with respect to that goal in which case the baseline rating is -2 (e.g. baseline shows maximum pain rating, not able to wear braces at all, not able to perform the goal at all).

Post treatment

If the patient achieves the expected level, this is scored at 0.

If patient achieves a better than expected outcome this is scored at:

- +1 (more than expected)
- +2 (much more than expected)

If the patient achieves a worse than expected outcome this is scored at:

- 1 (less than expected) or
- 2 (much less than expected)

A practical weighting scale for importance and difficulty was suggested by Turner-Stokes (2009)

Importance (by the patient/family)

- 0 = not at all important
- 1 = a little important
- 2 = moderately important
- 3 = very important

Difficulty (by the treating team)

- 0 = not at all difficult
- 1 = a little difficult
- 2 = moderately difficult
- 3 = very difficult

The GAS can be used for individual monitoring the progress of subjects undergoing interventions. GAS can also be used for larger populations (with differing individual goals) by using a formula. An excel sheet is available to calculate the GAS scores (See attached).

Formula

$$T = 50 + \frac{10 \sum w_i x_i}{\sqrt{(1-\rho) \sum w_i^2 + \rho (\sum w_i)^2}}$$

W_i is the weight given to the goal, and is the product of the rated importance of the goal and its difficulty. X_i is the numeric value of the attainment level of the goal, often scored from -2 to +2, where, for example, -2 would indicate an attainment level much less than expected. It is assumed that the mean (T-)value is 50 with a standard deviation of 10. The expected overall intercorrelation is indicated with ρ . In general, it accepted that ρ is set at 0.3 (Kiresuk and Sherman 1968).

If $T > 50$ = goal achieved. If $T < 50$ = goal not achieved.

Goal Attainment Scale:	
+2	Much more than expected
+1	Somewhat more than expected
0	Expected outcome
-1	Somewhat less than expected
-2	Much less than expected

Importance Rating Scale:	
0	Not at all important
1	A little important
2	Moderately important
3	Very important

Difficulty Rating Scale:	
0	Not at all difficult
1	A little difficult
2	Moderately difficult
3	Very difficult

Background / History

Development of the Assessment

The Goal Attainment Scale was developed for evaluation of psychological individual goals. Kiresuk and Sherman (1968) developed the technique in 1968. Since then the technique has been applied in many fields, especially in the evaluation of goals in rehabilitation setting. Measuring the effectiveness of treatments in rehabilitation using traditional standardized measures has always created major challenges due to the heterogeneity of the populations under study as well as the wide variety of subject goals and the importance of these goals to the subject and caregiver.

Turner-Stokes utilized the scale in multiple studies, clarified the appropriate use of the scale, and published a guide on the use of the GAS in rehabilitation (Turner-Stokes 2009).

Goals should be *Specific, Measurable, Attainable, Realistic and Time-Related (SMART)*. See the [GAS Article SMART Goals](#) article below. Suitable goals in the rehabilitation setting should be selected within the ICF context for body functions, activity and participation.

Reliability

In a systematic review of 58 articles reporting on the GAS, Gaasterland et al. (2016) found 12 studies reporting on inter-rater reliability. The inter-rater reliability was generally high (0.80-0.95 ICC) but 4 studies rated the reliability as poor. However, their review did not identify any articles that reported on intra-rater reliability.

In a literature review by Krasny-Pacini et al. (2013), the authors concluded that inter-rater reliability (IRR) is good but does vary according to the precision with which the levels are described, the person scoring the scale, and the field in question. IRR is improved when written by the physician/therapist treating the patient as well as when the raters observe the patients directly rather than videotaped assessments.

Vu and Law (2012) reviewed 14 published studies in the pharmacy literature on GAS psychometric properties. 13 studies reported high reliability. Correlation coefficients were greater than 85% in 9 studies.

Key studies:

Palisano (1993) examined the pre-study inter-rater reliability of 2 examiners scoring videotapes of the performance of 10 infants with motor delays. Agreement was 90% and the Kappa coefficient was 0.89. During the study, 16 goals were scored and agreement was 88% and the Kappa coefficient was 0.75.

Steenbeek et al. (2005) compared a video scoring and scoring by a physiotherapist for n=11 children with cerebral palsy undergoing botulinum toxin type A treatment. They reported Kappa=0.63. Only 5 out of 33 goal scores differed significantly.

Steenbeek et al. (2010) conducted a study to determine the inter-rater reliability of the GAS for n=8 physical therapists, n=8 occupational therapists, and n=4 speech therapists for n=23 children with cerebral palsy. The results showed an interrater reliability for the therapists of 0.82. The inter-rater reliability for independent raters was 0.64. The main reason for disagreement between raters was discrepancies in the professional's interpretation of the children's capacities versus their actual performance during assessment.

Stolee et al. (1992), n=15 geriatric patients, reported that 82% of goals were identified independently by two geriatricians. The physician-nurse inter-rater reliability was 0.87 ICC.

Stolee et al. (1999), n=61, found ICC=0.93 for the GAS follow-up score and ICC=0.89 of the separate goals when checked whether the goals had been attained.

Mailloux et al. (2007) stated that reliability concerns for the GAS include: (a) consistent goal selection, wording, and measurability of the scaled goals, (b) consistency in establishing increments between the levels of the scaled goals, and (c) interrater reliability among providers both within and across sites. They recommended training programs and testing of inter-rater reliability. Providers and clients need to be able to identify individual goals that can be expected to change as the result of the intervention and scale the goals into discrete levels of outcomes.

Validity

One of the major problems in the evaluation of the validity of the GAS is that the GAS does not measure one clear construct, since the individual goals differ for each patient. Furthermore, Gaasterland et al. (2016) concluded that there is insufficient information to assess the validity of GAS due to “the poor quality of the validity studies. However, the overall impression is that the GAS is a valuable tool for research in heterogeneous and small samples”.

Content Validity

In a review by Gaasterland et al. (2016), 5 studies in their review examined the content validity of the GAS, with results mostly “good” or “intermediate”.

Key studies:

Palisano (1993) conducted a study to examine (1) content validity of the GAS, (2) the responsiveness of the GAS compared to a behavioral objective, and (3) concurrent validity of the GAS and the Peabody Developmental Gross Motor Scale. Ten physical therapists rated 10 GAS goals on three dimensions for n=21 infants with motor delays. The results indicated that between 77% and 88% of the ratings for each dimension met the criterion for content validity. The rating of the 10 therapists did not differ significantly from each other. The results also indicated that the GAS was a more responsive measure of motor change compared to the behavioral objective. The authors concluded that validity is dependent on the judgment of the person or group who determine the goals. They must accurately assess the potential for change and the impact of the planned intervention to select levels of goal attainment that the subject is capable of achieving.

Stolee et al. (2012), n=90 subjects served by geriatric day hospitals, reported a “good overall usefulness” of the GAS in a study where clinicians rated the use of the GAS on a 5-point scale. Common goals included mobility, community reintegration, activities of daily living, medical issues, communication, and home safety.

Construct Validity

Gaasterland et al. (2016), identified 18 studies in their review which reported on the construct validity of the GAS by examining correlations with other instruments. Fourteen reported significant correlations. In contrast, Krasny-Pacini et al. (2013) reported poor correlations in the fields of geriatrics, cognition, neurologic disease, orthoses, and pediatrics.

Key studies:

Fisher and Hardie (2002), n=149 patients enrolled in a pain management program, correlated the GAS with improvements in walking, a general health questionnaire, Oswestry Low Back Pain Disability Questionnaire (OLBPDQ), NRS and change stand-sit and change PAIRS were measured. There was a significant correlation between the GAS and improvements for walking ($r=0.47$), the general health questionnaire ($r=0.25$), and the OLBPDQ ($r=-0.31$) with $p<0.01$ for all three. No significant correlations were found between the GAS and the NRS and change sit-stand and change PAIRS.

Steenbeek et al. (2011), n=23 children with cerebral palsy, found that the GAS, PEDI, and MFM-66 were complementary in their ability to measure individual change over time. They found a correlation with the Pediatric Evaluation of Disability Inventory Functional Status Score Mobility $r=0.64$ ($p<0.01$), correlation with PEDI Selfcare and social function was not significant.

Stolee et al. (1992), n=15 geriatric patients, correlated the GAS with the Barthel Index ($r=0.86$) and the global clinical outcome rating ($r=0.82$).

In an investigation by Stolee et al. (1999), n=173, change and follow-up scores of the GAS were correlated with the Barthel Index, Older American Resource Scale Instrumental Activities of Daily Living, Mini-Mental State Examination, Global Rating, and the Nottingham Health Profile. The correlations varied from -0.31 to 0.67,

Turner-Stokes et al. (2009), n=164, correlated the GAS with the Functional Independent Measure and the Functional Assessment Measure. Correlations with the FIM+FAM scores were moderate: 0.36-0.43 for raw scores, 0.41-0.49 for GAS transformed FIM+FAM scores.

Turner-Stokes et al. (2010), n=90, performed correlations between the GAS and MAS, Global Benefit patient report, Global Benefit investigator report, Hospital Anxiety and Depression Scale depression, Pain at rest, Pain on movement, Assessment of Quality of Life,

Patient Disability Score, and Carer burden score. Significant correlations between the GAS and MAS (0.35), Global Benefit patient report ($r=0.46$), and Global Benefit investigator report ($r=0.41$) were reported. Other correlations were not significant.

Turner-Stoles et al. (2013), $n=456$, found significant correlations between the GAS and MAS ($r=0.28$, $p<0.0001$) and with global assessment of benefit ($r=0.45$, $p<0.0001$).

Cusick et al. (2006) found that the GAS was poorly correlated to the Canadian Occupational Performance Measure (COPM).

Responsiveness

Gaasterland et al. (2016) found 14 studies examining responsiveness. Using different techniques such as the Relative Efficiency or the Standardized Response Mean (SRM) the GAS was shown to be responsive to change and often more than on other measurement instruments.

Key studies:

Cusick et al. (2006), $n=41$, found the ability to detect change over time and the ability to detect difference in change between groups measured with regression coefficients and effect sizes.

Stolee et al. (2012), $n=90$, found that GAS was able to detect meaningful change on all three measures of responsiveness.

Palisano et al. (1992) and Palisano (1993) investigated the validity of the GAS as a measure of motor change (responsiveness) in $n=65$ infants with motor delays. The results indicated that the GAS was responsive to change in individualized motor goals.

Turner-Stokes et al. (2010), $n=90$, found that a change in GAS score from baseline predicted a positive response with 52% sensitivity, 85% specificity, 81% positive predictive value, and 60% negative predictive value.

Sensitivity to Change

Ashford and Turner-Stokes (2006), $n=18$ disabled patients with brain injury, found that the GAS provided a useful measure of functional gains in response to treatment with BTX and was more sensitive to change than global measures such as the Barthel Index (Sensitivity=91% and Specificity=86%).

Cusick et al. (2006) compared the GAS to the Canadian Occupational Performance Measure (COPM) as outcome measures in a study of $n=41$ children with spastic hemiplegic CP. Both instruments were sensitive to within group change and detected significant between group change.

Tilton et al. (2017) in a study of $n=241$ children CP undergoing treatment with abobotulinumtoxinA, found that using the GAS confirmed good functional benefits in contrast with traditional standardized assessments that often failed to show improvement in patient function following injections.

Principles

The Goal Attainment Scale (GAS) is a functional scale used to measure the extent of progress towards individual goals during the course of an intervention. Each subject has their own outcome measures but they are scored in a standardized way. Scale items are based on current and expected levels of performance.

Individual goals are defined for each individual by the health care provider and the subject/caregivers. These goals are ranked according to their importance to the subject/caregivers. Then the health care providers rate the level of difficulty of each chosen goal.

Goals should be **Specific, Measurable, Attainable, Realistic and Time-Related (SMART)**.

Suitable goals in the rehabilitation setting should be selected within the ICF context for body functions, activity and participation.

Pros & Cons



Pros

- This measurement technique involves using individualized outcome indicators to construct detailed and comprehensive indexed measures that enable outcome evaluation of complex and multidimensional problems. It can be used to assess multiple individualized goals for each subject over time and program efficacy as a whole comparing performance across subjects within the same program (Vu and Law 2012).
- Ensures communication and collaboration between the multi-disciplinary team and the subject/caregivers.
- Helps to focus the subject/caregivers on realistic goals. Helpful especially for parents to see what is considered to be reasonable short-term achievement for their child on specific tasks, the sequence of anticipated progress, and the time-frame for their accomplishment (Maloney et al. 1978).
- Helps to focus the healthcare providers on goals that are important to the subject/caregivers.
- Subjects are more motivated and more likely to achieve goals when they are involved in the process.
- The GAS facilitates ongoing feedback for both providers and subjects and their caregivers.
- The GAS may help provide evidence of increments of change of individual goals reflected in a general measure required by third-party payers to support continued treatment.
- Avoids some of the issues of standardized measures such as floor and ceiling effects.
- Extremely flexible since goals are individualized and may be set for any area, any skill, for any time-frame.
- High sensitivity, for instance in cases where one specific item improved which is lost in the overall score.
- A GAS Practical Guide has been published by Turner-Stokes (2009) and is now available to download from the web at no charge.
- An electronic GAS calculation sheet is now available to download from the web at no charge.

Cons

- The most common error is establishing goals that are not SMART.
- Health care providers must have some experience in accurately predicting future performance.
- Training and experience is required to achieve inter-rater reliability.
- Estimating importance and difficulty (indicating weight in the GAS formula) is a major issue.
- Validity is dependent on the judgment of the person or group who determine the goals. They must accurately assess the potential for change and the impact of the planned intervention to select levels of goal attainment that the subject is capable of achieving (Palisano 1993).
- Setting goals requires extra time in the clinical setting and involves input from the inter-disciplinary team as well as the subject/caregivers.
- GAS calculation requires the use of a complex formula.
- The use of an ordinal scale in a complex mathematical formula (which requires interval level scale variables), is a concern for many investigators (Tennant 2007).

The developers of this website shall not be responsible for any injury and/or damage to persons or property as a matter of liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in this courseware. This website is strictly an educational forum about various assessments for the management of childhood motor disorders. The information on this page was developed using the best research evidence combined with the expertise of clinicians.