Goal Attainment Scaling: Outcome Measure of Consultation and IEP Progress

Lisa Ruble
University of Kentucky

John McGrew
Indiana University-Purdue University

NASP 2011

PA351 - Wed

Objectives/Learning Outcomes:

1. The use of GAS as a tool for:
   - Monitoring individual and group student progress and evaluating IEP outcomes
     - Teacher consultation
   - Conducting educational outcome research
2. How to take IEP objectives or other educational objectives and translate them into a GAS template.

Objectives/Learning Outcomes:

3. Understanding the assumptions of GAS and the strategies developed to test these assumptions GAS
4. A case study of GAS as an outcome measure applied to a randomized controlled study of teacher consultation

Progress monitoring and accountability

- IDEA and NCLB
  - Curriculum-based measurement
  - Classroom and large scale assessments

BUT……

How do we account for pivotal skills that impact all learning and yet are not well accounted for by standards-based assessment and accountability systems.

Challenge

- How do we monitor effectiveness of interventions and educational programs when
  - Skills represent
    - Social
    - Communication
    - Independent skills
  - The learning objectives (outcomes) are personalized for each student
  - Students start at different levels
  - The intervention varies

Goal Attainment Scaling

- GAS as a measurement system that can be applied as an objective approach for progress and outcome monitoring
  - Individual
  - Aggregated
- Sensitivity to the intervention being applied

So what about GAS?

- Allows for measurement of outcomes that are customized to the context/individual.
  - Been applied to district, school, classroom, and child level outcome assessment
  - Can be used with students who have different intervention outcomes and plans.
- Produces a Goal Attainment Score (GAS) which allows you to track progress and compare progress between groups.
- Allows you to weight goals according to importance.

Overview of Goal Attainment Scaling

- Developed by Kiresuk and Sherman (1968)
  - For mental health practitioners initially
  - Used by an array of disciplines today
- Used as the main outcome in studies on consultation effectiveness and is considered to be a standard (e.g., Sladeczek, et. al., 2001).
- Blends well with IEP goals (Oren & Ogletree, 2000).

Goal Attainment Scale

-2 Present level of perfor-mance	-1 Progress	0 Expected level of outcome (goal)	+1 Somewhat more than expected	+2 Much more than expected

Steps to Create GAS: Standard description

1. Identify expected goals/outcomes
2. Weight the goals according to priority (more severe, more importance = higher weighting)
3. Identify continuum of benchmarks
4. Determine baseline performance
5. Implement intervention
6. Monitor progress
7. Evaluate final goal attainment

Indicator	Value
Much less than expected outcome	-2
Less than expected outcome	-1
Expected outcome	0
More than expected outcome	+1
Much more than expected outcome	+2

Expected Outcomes

- Start by identifying the most likely outcome.
- Might be based on annual outcomes or broken down for every 6 weeks
- The middle level is the most probable and successful level of goal attainment
- Above this level is even greater response
- Below this level is less successful response

Level at intake: * Date _____

Level at followup: ✓ Date _____

Goal Attainment Scale

!

Check whether or not scale has been mutually negotiated between client and therapist	Goal Headings and Goal Weights				
	Yes___No___	Yes___No___	Yes___No___	Yes___No___	Yes___No___
Goal Attainment Levels	Goal 1	Goal 2	Goal 3	Goal 4	Goal 5
most unfavorable treatment outcome thought likely (-2)	!	!	!	!	!
less than expected success with treatment (-1)	!	!	!	!	!
expected level of treatment success (0)	!	!	!	!	!
more than expected success with treatment (+1)	!	!	!	!	!
best anticipated success with treatment (+2)	!	!	!	!	!
Comments:	!	!	!	!	!

More about goal writing:

- The outcomes should be:
  - measurable and specific
  - not too easily accomplished or too difficult to obtain
- Two independent observers should be able to agree on whether it has been obtained
  - Adequate interobserver reliability

Scoring Goal Attainment Scales

- If you are working with <u>individuals</u>:
  - Each time you measure goal attainment mark the box which indicates the individual's current status.
  - To obtain a score, add up the scores for each goal.
  - Scores on each goal range from -2 to +2 with 0 indicating treatment success.
  - The total across all goals represents the goal attainment score.

Scoring

- All scores are standardized and converted to T scores (i.e., *M* = 50; *SD* = 10) using the Kiresuk-Sherman formula (Kiresuk, et. al., 1994).

If you are comparing or collecting scores from a number of individuals with different numbers of goals:

- Convert the GAS into a standard score.
- Refer to the GAS conversion table.
- Find the row which indicates an individual's goal attainment score and find the column which represents the number of goals.
- The standardized GAS score is where the row and column intersect.

Goal Attainment Score Conversion Table

Total raw score (Sum of scale scores)	Number of Goals in Followup Guide					
	1	2	3	4	5	6
-12						19
-11						22
-10					20	24
-9					23	27
-8				21	26	29
-7				25	29	32
-6			23	28	32	35
-5			27	32	35	37
-4		25	32	35	38	40
-3		31	36	39	41	42
-2	30	38	41	43	44	45
-1	40	44	45	46	47	47
0	50	50	50	50	50	50
1	60	56	55	54	53	53
2	70	62	59	57	56	55
3		69	64	61	59	58
4		75	68	65	62	60
5			73	68	65	63
6			77	72	68	65
7				75	71	68
8				79	74	71
9					77	73
10					80	76
11						78
12						81

Assumptions of GAS

- GAS is a continuous measure
- Comparability of GAS scores across individuals
  - Equal interval between each scaled description
  - Measurability or objectivity of the targeted skill
  - Degree of difficulty in obtaining the skill due to different levels of starting abilities
  - Argued as a nonstandard measure (Mackay, 1996)
- Reliability of progress
  - Subjectivity of ratings (parent & teacher report)

Suggestions

- Conduct training on writing GAS
  - degree of difficulty and equality of intervals
- Operationalize definitions of outcome criteria ⟶ measurability
- Collect subjective and objective data in the determination of attainment levels
- Utilize an independent observer to code GAS
- Ensure adequate treatment integrity

Coffee, G. & Ray-Subramanian, C. (2009)

Case Example

- The limitations of GAS and strategies to test the assumptions of GAS as a standard approach are presented within a randomized controlled study of teacher consultation for students with autism.

Our adaptation for use of consultation outcome research on IEP goal attainment

- A 5-point response scale is used:
  - -2 (child's present levels of performance = worst outcome)
  - -1 (progress)
  - 0 (expected level of outcome)
  - +1 (somewhat more than expected)
  - +2 (much more than expected)
- For each goal, specific behavioral descriptors were developed delineating degrees of progress toward the goal.
- Completed before intervention began.

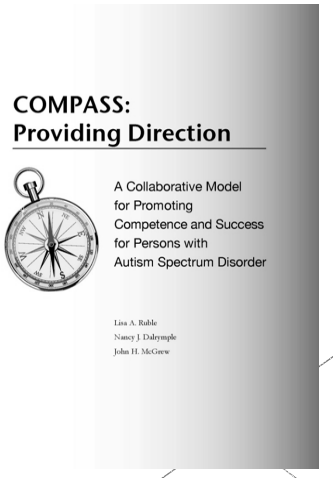
Goal Attainment Scale

-2 * Present level of perfor-mance	-1 Progress	0 Expected level of outcome (goal)	+1 Somewhat more than expected	+2 Much more than expected

*Present levels is used because the children are not expected to regress as recommended by Schlosser, 2004

Creating a GAS form

- A protocol was developed to ensure that the descriptions were measurable and of equal interval.
  - Completed after the fact, as a means for ensuring forms were similar between groups



Addressing Untested Assumptions of GAS

1. Comparability:

   Developed 3 dimensions coded by an independent rater:
   - Measurability
   - Level of Difficulty
   - Degree of equidistance between intervals

Addressing Untested Assumptions of GAS

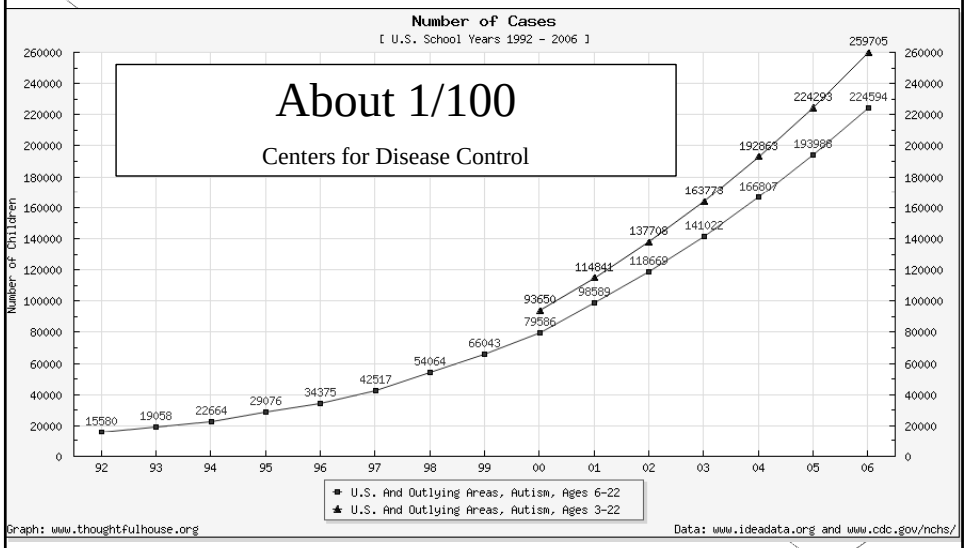
2. Objectivity/Reliability:
   - GAS coding was based on direct observation
   - Reliability of scores from live coding vs taped coding
     - Reliability of scores between teacher made tapes and live coding by researcher
       - Do teachers provide the "best scenario" tapes

What Problem are We Trying to Address Using GAS?

- Autism is described as a "national health emergency" (Interagency Autism Coordinating Committee, 2009)
- Annual costs are $35-90 billion (Jarbrink & Knapp 2001; Montes & Halterman, 2008)
- Total lifetime costs for individuals of 3.2 million (Ganz, 2006)

The Problem

- Compared to other children with special health care needs, children with ASD unmet needs for
  - specific health care services,
  - family support services,
  - delayed or foregone care,
  - difficulty receiving referrals, and
  - care that is not family centered
- African American and Latino children
  - receive later diagnosis,
  - have greater symptom severity, and
  - receive significantly lower numbers of services (Liptak, 2008; Mandell, et al., 2006, 2009; Morrier, et al., 2008)
- Children living in rural areas are underserved (Chen et al., 2008; Farmer, et al., 2005)



Solutions?

- Where are children getting services?
- What are mandated services?

Educational system is the *only* public funded and mandated agency to provide services

But…

- National shortage of special educators
- Educators lack knowledge of evidence based practices (Hess et al., 2008; Stahmer et al., 2005)
  - Services are often insufficient, lack specificity and intensity
  - Geographic disadvantages
- Little research available from community-based settings on effectiveness

The Question

- How can we start to reduce the research-to-practice gap by addressing the educational outcomes of children with autism?

Consultation

- Teacher consultation is effective (Busse et al., 1995; Medway & Updyke, 1985; Sheridan et al., 1996)
- Models of consultation
  - Behavioral (Noell et al., 2005)
  - Conjoint behavioral (Freer & Watson, 1999; McDougal, Nastasi, & Chafouleas, 2005; Sheridan & Steck, 1995; Sheridan, Clarke, Knoche, & Edwards, 2006; Sheridan, Eagle, Cowan, & Mickelson, 2001; Sladeczek, et. a., 2001; Wilkinson, 2005)
  - Collaborative (Denton, Hasbrouck, & Sekaquaptewa, 2003; Erchul, Hughes, Meyers, Hickman, & Braden, 1992; Givens Ogle, Christ, & Idol, 1991; Ray, Skinner, & Watson, 1999; Yocom & Staebler, 1996)
  - Systems consultation (Denton et al., 2003)

Types of problems

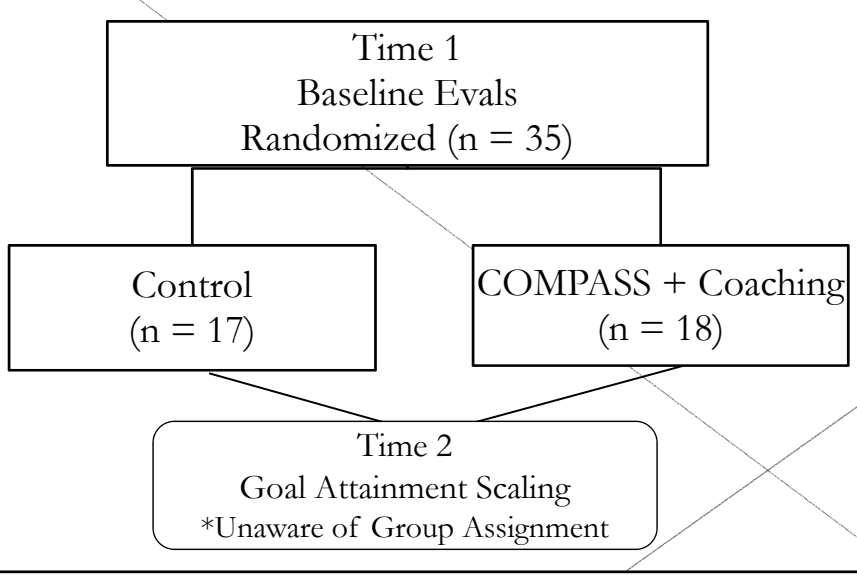
- Student achievement (Givens Ogle et al.,1991)
- Student disruptive behavior (Denton et al., 2003; McDougal, et., al., 2005; Ray et al., 1999; Sheridan et al., 2001; Sladeczek et al., 2001; Wilkinson, 2005)
- Teacher behavior (Cossairt, Vance Hall, & Hopkins, 1973; Meyers, Freidman, & Gaughan, 1975; Noell et al., 2005; Sparks, 1988; White & Fine, 1976)
- Parent-teacher relationships (Sheridan et al., 2006).

Primary outcome measure

- Goal attainment scaling
  - Start at different levels, different outcomes, different interventions
- Address core symptoms of autism
  - Impaired communication
  - Impaired social interaction

Randomized Single Blind Controlled

(Ruble, Dalrymple, & McGrew, 2010)



Group Comparison

- Comparison Group
  - Services as usual
  - Final evaluation
- Intervention Group
  - 3 hour COMPASS consultation (parent and teacher)
    - 3 IEP objectives
      - Specific to needs of students with autism
      - Measurable
      - Teaching plans
  - 4 teacher coaching sessions (1.5 hour every 4-6 weeks)
  - Final evaluation

Between Group Comparisons

Characteristics	Control *M (SD)*	Experimental *M (SD)*	*t*(df)	*p*
Child				
Age	5.98 (1.5)	6.18 (1.9)	-.34(33)	.74
Childhood Autism Rating Scale	41.43 (8.2)	36.38 (9.9)	1.55(30)	.13
Differential Abilities Scale	39.47 (18.4)	53.78 (27.1)	-1.81(33)	.08
Oral and Written Language Scales	41.13 (19.0)	51.56 (17.2)	-1.68(32)	.10
Adaptive Behavior Scales	62.29 (9.2)	64.88 (16.7)	-.56(32)	.58
BASC (externalizing composite)	59.53 (8.5)	59.83 (7.0)	-.11(31)	.91
Teacher				
Total Number of Children Taught	8.85 (11.5)	4.56 (6.1)	1.29(27)	.21
Total Years Autism	8.27 (8.3)	5.34 (5.5)	1.16(29)	.25

Primary Hypothesis

- Children whose teacher's and parent's participated in the consultation intervention would demonstrate better goal attainment outcomes compared to children who received their usual education program.

Goal Attainment Scale

Rating:	-2 Present level of perfor-mance	-1 Progress	0 Expected level of outcome (goal)	+1 Somewhat more than expected	+2 Much more than expected	Demon-strated for at least 2 weeks?
Skill observed? ❑ no ❑ yes Goal changed? ❑ no ❑ yes						

*Based on direct observation

Anthony's IEP Objectives after COMPASS Consultation

1. When presented with a task menu, Anthony will start and complete three 2-3 minute tasks each day without aggression with one adult verbal cue (e.g., time to work) and gestural/picture cues across two weeks.

2. During structured play, Anthony will imitate adult play activities for five actions (actions with objects) with at least three different preferred objects (dinosaurs, animals, doll) each day across two weeks.

3. Anthony will make 10 different requests per day independently (go home, eat, help, more, finished, various objects/activities) or as a response to a question ("what do you want?") using sign, pictures, or verbalization.

Goal Attainment Scale (GAS) Form

Student's Name:_____ Observer:_____ Date:_____ Coaching I II III IV

Instructions: Indicate rating, if skill was directly observed, and whether the goal has changed since the last observation.

Skill Level →	-2 Present level of performance	-1 Progress	0 Expected level of outcome (GOAL)	+1 Somewhat more than expected	+2 Much more than expected	Domain of Goal*	Demo for at least 2 weeks?**
1. Rating:_____ Was skill observed? ❑ no ❑ yes Has goal changed? ❑ no ❑ yes							
2. Rating:_____ Was skill observed? ❑ no ❑ yes Has goal changed? ❑ no ❑ yes							
3. Rating:_____ Was skill observed? ❑ no ❑ yes Has goal changed? ❑ no ❑ yes							

* Note: "Domain" refers to social, communication, learning skill, adaptive behavior, motor/sensory or academic goals.
**Has student been able to perform at criterion level for at least two weeks?

Page 123 | Chapter 8: COMPASS: Coaching Sessions: Implementing Plans and Monitoring Progress

COMPASS: Providing Direction. By Lisa A. Rohde, Nancy J. Dalrymple, and John H. McGrew 2010

Progress Descriptions

Dimension	GAS Score				
	-2	-1	0	+1	+2
Frequency of targeted skill	Lowest →				Highest
Frequency of prompting	Highest →				Lowest
Form of prompting[1]	Physical →				Independent visual supports
Context[2]	Structured One context →				Unstructured Many contexts
Person	An adult →				Many adults Many peers

Goal Attainment Scale (GAS) Form for Anthony

-2 Present level of performance	-1 Progress	0 Expected level of outcome (GOAL)	+1 Somewhat more than expected	+2 Much more than expected
Aggresses when given a task he does not want to do. Is difficult to motivate. Does not have a more appropriate way to communicate refusals or to negotiate.	When presented with a task menu, Anthony will start and complete three (1) 2-3 minute tasks each day without aggression with one (2) adult verbal cue (e.g., time to work) and gestural/picture cues across two weeks.	When presented with a task menu, Anthony will start and complete three 2-3 minute tasks each day without aggression with one adult verbal cue (e.g., time to work) and gestural/picture cues across two weeks.	When presented with a task menu, Anthony will start and complete three (4) 2-3 minute tasks each day without aggression with one (0) adult verbal cue (e.g., time to work) and gestural/picture cues across two weeks.	When presented with a task menu, Anthony will start and complete three (6) 2-3 minute tasks each day without aggression with one (0) adult verbal cue (e.g., time to work) and gestural/picture cues across two weeks.
Has difficulty imitating others, especially children using actions with objects. Likes objects he can manipulate.	Anthony will imitate play activities for five (2) minutes with at least three (1) different preferred objects (dinosaurs, animals, doll…) each day across two weeks.	Anthony will imitate adult play activities for five minutes with at least three different preferred objects (dinosaurs, animals, doll…) each day across two weeks.	Anthony will imitate adult play activities for five (7) minutes with at least three (4) different preferred objects (dinosaurs, animals, doll…) each day across two weeks.	Anthony will imitate adult (peer) play activities for five (10) minutes with at least three (6) different preferred objects (dinosaurs, animals, doll…) each day across two weeks.
May use aggression as a way to request. Relies on adult prompts to make requests.	Anthony will make 10 (5) different requests per day independently (with verbal cues) or as a response to a question (go home, eat, help, more, finished, various objects/activities) using sign, pictures, or verbal on a daily basis.	Anthony will make 10 different requests per day independently (go home, eat, help, more, finished, various objects/activities) or as a response to a question ("what do you want?") using sign, pictures, or verbalization on a daily basis.	Anthony will make 10 (15) different requests per day independently (go home, eat, help, more, finished, various objects/activities) or as a response to a question ("what do you want?") using sign, pictures, or verbalization on a daily basis.	Anthony will make 10 (20) different requests per day independently (go home, eat, help, more, finished, various objects/activities) or as a response to a question ("what do you want?") using sign, pictures, or verbalization on a daily basis.

Interobserver Reliability of GAS:

- An observer scored the GAS form on 20% of the sample and results were compared with the primary observer's scores (both direct observation coding).
- Weighted kappa coefficient of .65, indicating moderate to good agreement.

Fidelity

- 96% (teacher)
- 83% (parent)

COMPASS Consultation Fidelity Checklist

1. The COMPASS consultation is comprised of a multidisciplinary team defined by:

- ❑ teacher and parents both attend meeting
- ❑ child's other school and community based therapists attend meeting

2. COMPASS is collaborative as defined by:

- ❑ goals include those suggested from home and family
- ❑ planning for child's program is based on input from all participants
- ❑ each member contributes ideas for teaching the goals

3. The COMPASS consultation process incorporates:

- ❑ checklists that are used to help organize information, identify child's needs, and solicit input from all members
- ❑ facilitated guidance and structure from the consultant
- ❑ a total picture of child at home, in the community, and at school

4. Treatment goals that came from the COMPASS consultation are:

- ❑ described in clear behavioral terms
- ❑ measurable and observable

5. COMPASS consultation results in a teaching plan that:

- ❑ identifies at least three priority concerns
- ❑ prioritizes concerns that relate to home, community and school
- ❑ identifies specific skills that the child must learn in order to address to accomplish each of the priority concerns
- ❑ links the specific teaching strategies to each identified skill

6. The teaching strategies described in the COMPASS plan:

- ❑ are developed AFTER goals are generated
- ❑ are individualized for the child and the goal
- ❑ Aare described in behavioral terms

7. Team believes the child's ability to learn is based on environmental as well as child factors:

- ❑ there was a discussion of specific environmental factors for each goal
- ❑ the philosophy of the environment as a factor as important as the child in determining child progress is discussed
- ❑ team completes and discusses COMPASS forms on child's strengths/challenges and environment's strengths/challenges

8. COMPASS results in members having a broader understanding of the child:

- ❑ Family members report that they have a better perspective on school issues
- ❑ teachers report that they have a better perspective on home and community issues

9. COMPASS consultation results in proactive problem solving:

- ❑ interactive problem solving is implemented by team members providing input and ideas
- ❑ specific barriers are identified and plans to overcome barriers are identified
- ❑ members learn a framework for problem solving that can be used again by individual team members when needed

Page 106 | Chapter 7: COMPASS Consultation Action Plan—Step B

COMPASS: Providing Direction. By Lisa A. Ruble, Nancy J. Dalrymple, and John H. McGrew. 2010.

Fidelity

- Teacher adherence: (5 point scale)
- 1 "0% implementation" to 5 "100% implementation"

Coaching Session	Fidelity Score Mean (SD)
1	3.0 (2.0)
2	3.5 (1.3)
3	4.0 (1.2)
4	4.1 (1.2)

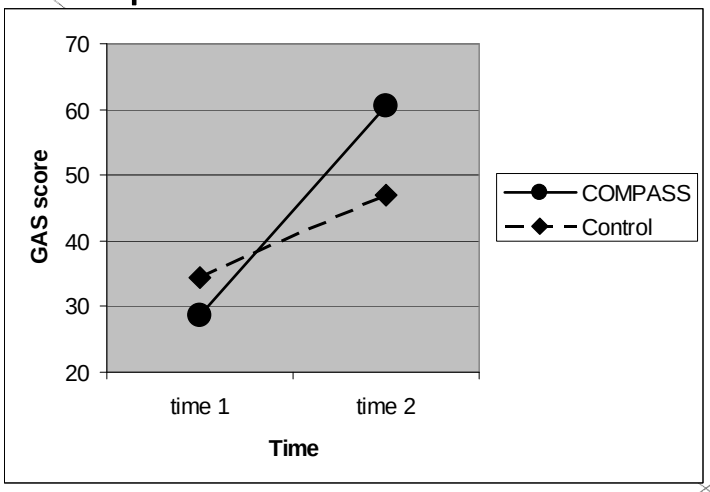
Interventions Outside School

- No significant difference between groups for the number (t (27) = -.83, n.s.) or hours (t (33) = -1.00, n.s.) of services received outside of school.

Results

- Do children whose teachers receive COMPASS have better IEP goal attainment outcomes?

Group x Time Interaction



**Raw GAS change scores: $t(27) = -2.6$, $p = .02$, $d = 1.0$*

Alternative explanation of results…

- The level of difficulty in the GAS forms was lower for the consultation group vs comparison so they can't be compared
- The intervals between the scales were not equivalent so group differences can't be compared
- The measurability of the objectives was poorer for the comparison group so they can't be compared

Group Comparisons

- Operationalized measurability, difficulty, equidistance
- Two raters coded GAS forms separately until 80% agreement
- Rater unaware of group assignment did final coding
- All coding was based on observation instead of parent/teacher report

Definitions: Measurability

1 – None or only one indicator (prompt level, criterion for success; observable skill) is listed 2 - Two of the three indicators (prompt level, criterion for success; observable skill) are provided 3 – Describes all three indicators (prompt level, criterion for success; observable skill)

Definitions: Difficulty

1 – Skill is very close to what the child is already described as able to perform in the present levels of performance

2 – The present levels of performance indicates that the child is able to perform the skill in limited ways compared to what is written in the objective (limited people, prompts, or places); if PLEP says child has difficulty doing it, score a "2"

3 – The present levels of performance indicates that the child is unable to perform skill with anyone, anywhere, or with any prompts compared to what is written in the objective

Definitions: Equidistance

1 – None or only 1 of the three descriptions are equilibrated appropriately in reference to the targeted objective (which is zero).
[Prompts described as going from environmental to verbal, and /or the skill frequency increases or reduces by less or more than 50% (do not include the present levels of performance -2 description].

2- Two of the three descriptions are equilibrated appropriately in reference to the targeted objective (which is zero).
[Descriptions are scaled accordingly from verbal to physical to visual/environmental prompts or the frequency of skill increases or reduces by at least 50% relative to the targeted objective (do not include the present levels of performance -2 description)].

3 – All of the three descriptions relative to the targeted objective are equilibrated and scaled appropriately.
[Hierarchy of skills are dropped from verbal to physical to visual/environmental or frequency of skill increases or reduces by at least 50% for the majority].

Intercorrelations between dimensions

	Measurability	Difficulty	Equidistance
Measurability	--		
Difficulty	.06	--	
Equidistance	-.06	.03	--

T-test Analysis of GAS Scores*

Dimension	Comparison (n=34)	Consultation (n=48)	t-test	Cohen's *d*
Measurability	2.7 (.52)	2.5 (.62)	-1.5, p = .23	.35
Difficulty	2.3 (.60)	2.1 (.66)	1.9, p = .06	.32
Equidistance	2.9 (.34)	2.8 (.42)	.29, p = .78	.26

Original results the same, after controlling for level of difficulty
ANCOVA, F (1,70) = 21.0, p .000, d = 1.0

Note: Comparison group GAS forms were masked to make them measurable

*Based on GAS forms at baseline for the comparison group and after consultation for the experimental group

Current Study

American Recovery & Reinvestment Act NIH Challenge Grant 5RC1MH89760

- Same experimental design and participant characteristics, but…
- Three groups
  1. COMPASS + FF Coaching (3 hours + 4 sessions, about 1 - 1.5 hours each)
  2. COMPASS plus WEB-based coaching
  3. Online teacher training (instead of SAU)
- All teachers did own taping for coaching sessions (we still did direct observation for final determination of progress at end of year)

Additional GAS issues when videotaping is used…

- Parent and teacher report is subjective
  - When teachers make and provide their own videotapes of instruction, do they take the "best" example?
  - We collected extra data point (live) that corresponded to a teacher coaching session
- Is coding based on live observation and from taped observation reliable?
  - Each has own advantages/disadvantages
    - Live allows more information of the context
    - Taped allows for observation to be played back

Opportunity to explore additional GAS issues when using videotape analysis

- Scores from teacher supplied taped instruction (n=11) was reliable with researcher supplied taped instruction
  - ICC = .65
- Scores based on coding done from live observation and coding from the same observation but taped were reliable
  - ICC = .66

But what about the three dimensions of GAS?*

Dimension	Comparison (n=21)	Consultation (n=31)	t-test
Measurability	2.9 (.30)	2.6 (.49)	2.4, p = .02
Difficulty	2.2 (.44)	2.4 (.56)	-1.3, p = .20
Equidistance	2.8 (.40)	2.7 (.44)	.56, p = .58

*Based on GAS forms at baseline for the comparison group and after consultation for the experimental group; coded at end of the school year during final evaluation by independent observer

Bottom Line

- Do teacher present the base case scenarios when they provide the taped observation?
  - No – time is of the essence
    - We are only using researcher supplied tapes for GAS scores for research purposes
- Are scores based on tapes reliable with scores based on live coding? Yes
- Is it necessary to monitor the dimensions of measurability, level of difficulty, and equidistance? Yes and early on

Acknowledgements

- Teachers, Indiana & Kentucky
- Families and Children
- Nancy Dalrymple, Co-investigator
- Michael Toland, UK
- Research team, UK
  - RAs: Ryan Johnson & Lauren Feltner
  - GRAs: Rachel Aiello, Jessie Birdwhistell, Jennifer Hoffman
  - Co-investigators: Jennifer Grisham Brown & Lee Ann Jung
- Research was supported by Grants No. R34MH073071 and 1RC1MH089760 from the National Institute of Mental Health.