Assessing Performance on Individualized Goals for Students with ASD

Lisa Ruble

University of Kentucky

John McGrew

Indiana University-Purdue University

IMFAR 2011

What is the problem we are trying to address?

- Improve the educational outcomes of children with autism
- Our 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)

Consultation

- Consultation is a method to serve a large number of students by increasing teacher skills

- Teacher consultation is effective (Busse et al., 1995; Medway & Updyke, 1985; Sheridan et al., 1996)

Challenge

- How do we conduct experimental treatment studies when
  - Treatment outcomes are often different for each child
  - Children start at different baseline levels
  - Norm referenced tests are not sensitive to measuring the goals or the effects of the intervention
  - Measurement approach needs to be able to be implemented in community settings

Primary outcome measure

- Goal attainment scaling
  - Start at different levels, different outcomes, different interventions
  - Primary outcome measure in consultation effectiveness research

- Address core symptoms of autism and pivotal skills
  - Communication
  - Social interaction
  - Learning / work behavior skills

Goal Attainment Scaling

- Developed by Kiresuk and Sherman (1968)
  - For mental health practitioners initially
  - Used by an array of disciplines today
- Can be applied as an objective outcome measurement tool
- Sensitivity to the intervention
  - Allows for measurement of outcomes that are customized to the context/individual
  - Ideal for children with different intervention outcomes and plans

Steps to Create GAS: Standard Method

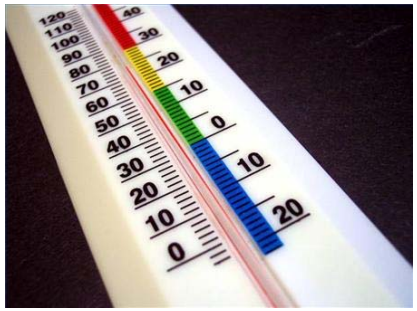
1. Identify expected goals/outcomes
2. Weight the goals according to priority (more severe, more importance = higher weighting)
3. Identify continuum of benchmarks
4. Determine baseline performance
5. Implement intervention / monitor fidelity
6. Monitor progress
7. Evaluate final goal attainment

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

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.

Assumptions of GAS



- GAS scores are comparable across individuals
  - Are the intervals between each scaled description equal?
  - Is the targeted skill measurable and objective?
  - Is the degree of difficulty equivalent between groups when starting abilities vary between children?
- Scores can be converted into standardized T scores
  - Argued as a nonstandard measure (Mackay, 1996) and assumptions of normality cannot be assumed.
  - Results based on whether data are ordinal or interval show negligible differences (Cardillo & Smith, 1994; Malec, 1999; Ottenbacher & Cusic, 1993).

Suggestions

- Conduct training on writing GAS
  - Degree of difficulty and equality of intervals
  - Verify and adjust descriptions prior to intervention
- Operationalize definitions of outcome criteria = measurability
- Collect subjective and objective data in the determination of attainment levels
- Utilize an independent observer
- Apply raw scores

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

Case Example

- Assumptions of GAS methodology and strategies to test assumptions are presented within a randomized controlled study of teacher consultation for students with autism.

Our Adaptation for Use of GAS in Consultation Outcome Research on IEP Goal Attainment

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

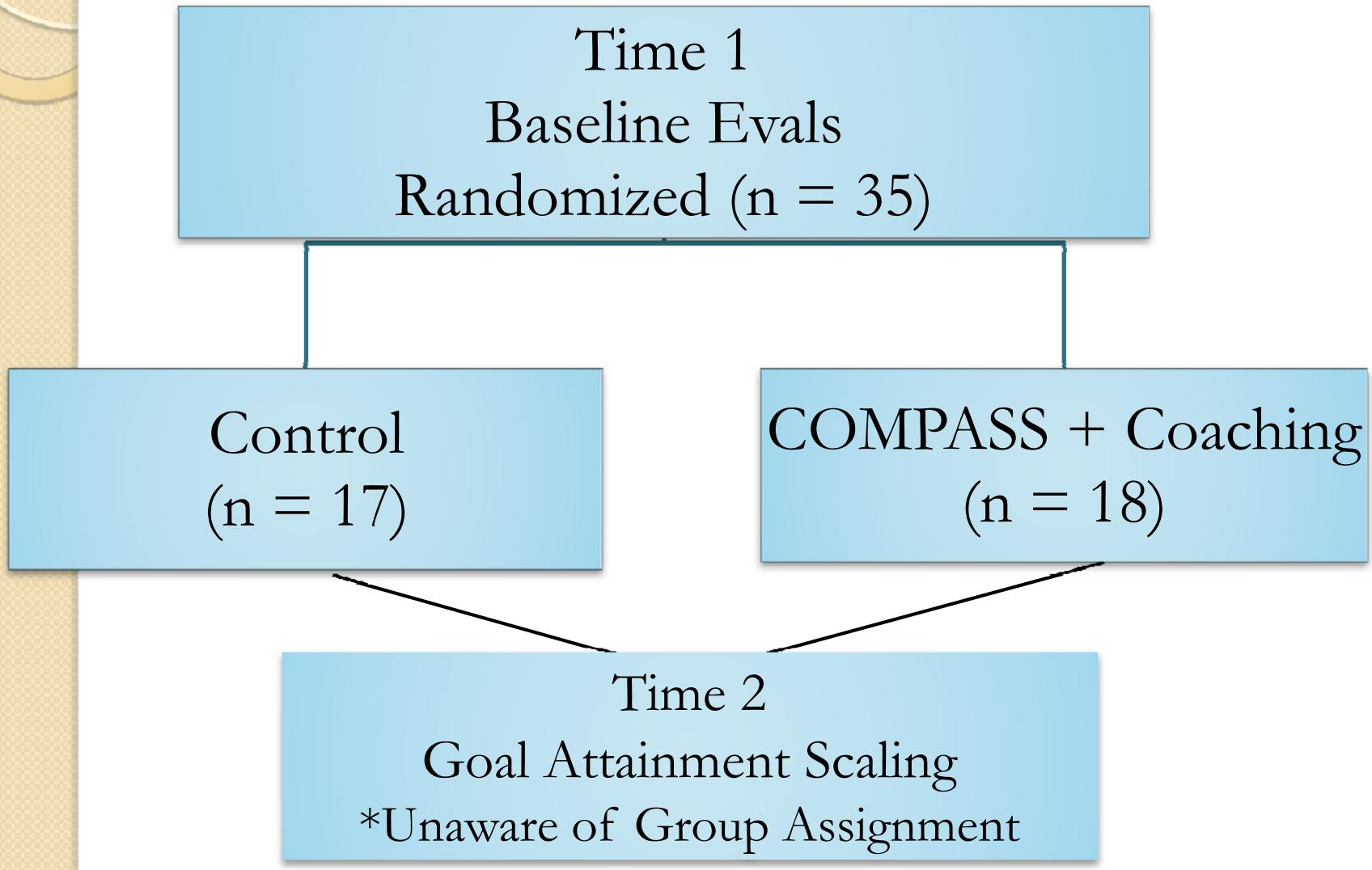
- A 5-point response scale is used (Schlosser, 2004)
- For each goal, behavioral descriptors are created

Addressing Limitations of GAS

1. Developed protocol to create GAS forms.
2. Operationalized three dimensions coded by an independent rater:
   - Measurability
   - Level of Difficulty
   - Degree of equidistance between intervals
3. Examined between group differences for the three dimensions

Randomized Single Blind Controlled

(Ruble, Dalrymple, & McGrew, 2010)



Group Comparison

- <u>Comparison Group</u>
  - Services as usual
    - Baseline GAS
  - Final GAS evaluation

- <u>Intervention Group</u>
  - 3 hour COMPASS consultation (parent & teacher)
    - Baseline GAS
  - 4 teacher coaching sessions
  - Final GAS evaluation

Primary Hypothesis

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

Anthony's IEP Objectives after COMPASS Consultation

1. When presented with a task menu, Anthony <u>will start and complete three 2-3 minute tasks</u> 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 <u>will imitate adult play activities for five actions</u> (actions with objects) with at least three different preferred objects (dinosaurs, animals, doll) each day across two weeks.

3. Anthony <u>will make 10 different requests per day</u> 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 verbalizations.

Progress Descriptions

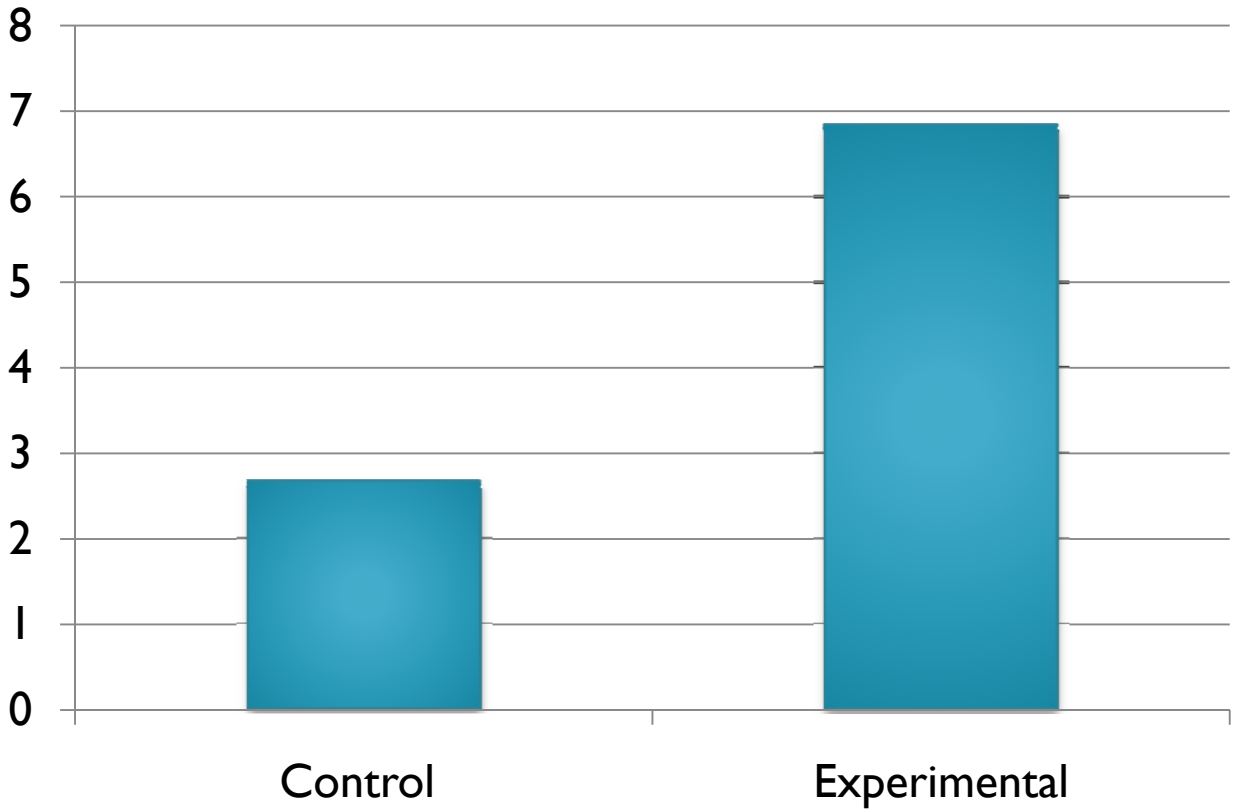
Dimension	GAS Score				
	-2	-1	0	+1	+2
Frequency of targeted skill	Lowest →				Highest
Frequency of prompting	Highest →				Lowest
Form of prompting[1]	Physical →				Independent visual supports
Context[2]	Structured One context →				Unstructured Many contexts
Person	An adult →				Many adults Many peers

children using actions with objects. Likes objects he can manipulate.	(2) minutes with at least three (1) different preferred objects (dinosaurs, animals, doll…) each day across two weeks.	for five minutes with at least three different preferred objects (dinosaurs, animals, doll…) each day across two weeks.	for five (7) minutes with at least three (4) different preferred objects (dinosaurs, animals, doll…) each day across two weeks.	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.

Interrater Reliability of GAS

- An independent rater scored the GAS form on 20% of the sample and results were compared with the primary observer's scores (both direct observation coding).
- Intraclass correlations of .73 (baseline) and .99 (final assessment) were calculated.

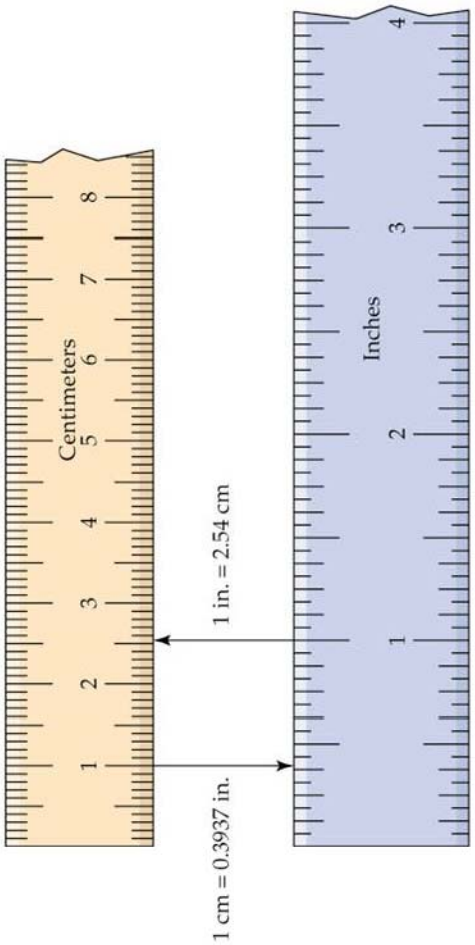
GAS Change Scores*



*$t(27) = -2.8$, $p = .01$, $d = 1.2$.

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 quality features
  - measurability
  - difficulty
  - equidistance

Definitions

Code	Description
Measurability	Number of indicators (prompt level, criterion for success; observable skill) listed for each description
Difficulty	Child's ability to perform the skill described in present levels of performance compared to what is written in the goal(number of people, prompts, or places)
Equidistance	Distance between each benchmark is equilibrated relative to the targeted objective are equilibrated and scaled appropriately (e.g., -1 = 50% less of target skill; +1 = 50% more; +2 = 100% more)

Intercorrelation Matrix

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

Results

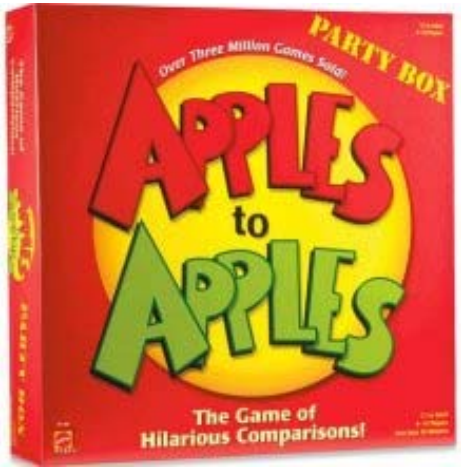
Dimension	Comparison	Consultation	t-test	Cohen's d
Measurability	2.7 (.52)	2.5 (.62)	-1.5, p = .23	.35
Difficulty	2.3 (.60)	2.1 (.66)	1.9, p = .06	.32
Equidistance	2.9 (.34)	2.8 (.42)	.29, p = .78	.26

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

Note: Based on GAS forms at baseline for the comparison group and after consultation for the experimental group; Comparison group GAS forms were masked to make them measurable;

Summary

- GAS is a promising outcome monitoring tool, but it is recommended that …
  - •Reliability be evaluated for each study
  - •Measurability, equidistance, level of difficulty be evaluated prior to intervention implementation
  - •GAS scores be based on direct observation
  - •Raw scores be used for data analysis



Acknowledgements

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