

# Treatment Integrity, Reliability, and other Evidence-Based Challenges in Implementing Ethical ABA Programs

National Autism Conference  
Pennsylvania State College  
College Town, PA  
5 August, 2014

John W. Esch, Ph.D., BCBA-D

1

1

## Philosophy 101

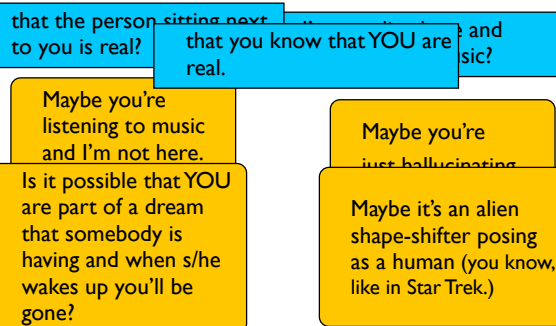
How Do We Know About Things?

Some Thought Experiments (and  
music) Follow

2

2

How Do You Know...



3

3

## How Do You Know...

that your students learned last year?

that student grades are an accurate measure of what was learned?

What was teaching was or a student's

Maybe s/he already knew those things or "mastered" or finally the concept."

that your student's tests evaluations are reliable?

An observation?

Would s/he get the same evaluation if you gave the test again

4

4

---

---

---

---

---

---

---

---

---

---

## Teaching and Knowing

What's teaching?

It's what the teacher does.

It's an art.

It's a science.

5

5

---

---

---

---

---

---

---

---

---

---

## Teaching and Knowing

What's teaching?

A functional definition:

Doing something that changes behavior  
(that which is learned)

6

6

---

---

---

---

---

---

---

---

---

---

## Teaching and Knowing

Knowing involves...

Truth v. Belief

Fact v. Fiction

pix: unicorn and  
elephant

pix: flat earth v. planet

7

7

## Knowing and Teaching

Science is a method of knowing

Behaviorism is the science of  
knowing how behavior changes

8

8

## *Interesting Causes*

World view:

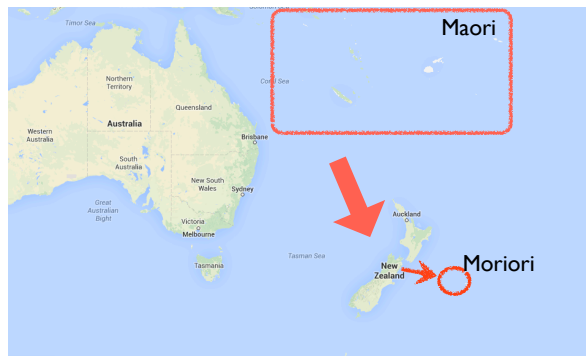
- Behavior is a product an organism's capacity and the environment.
- Environmental stimuli change and organism's behavior (within its capacity)
- Selectionism and survival of a species/group
  - Insects v. insecticides and genetically modified seeds (GM)
  - Bacteria v. antibiotics
  - Moriori cultural practices v. Maori cultural

Diamond, Jared (1997), *Guns, Germs, and Steel*

9

9

## Interesting Causes



Diamond, Jared (1997), *Guns, Germs, and Steel*

10

10

## Applied Behavior Analysis: the methodology of behaviorism

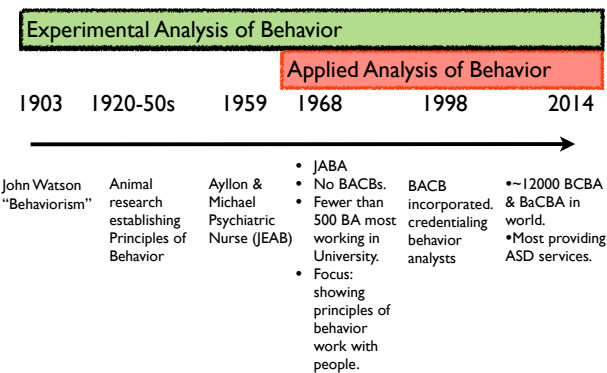
Discriminate “truth” from “belief,” “fact” from “fiction”

- Guidelines for determining important behaviors and determining *who* decides what's important (i.e., *ethics, social validity and generality*)
- Identifying environmental principles that change behavior
- *Using principles* to develop procedures to change behavior
- Guidelines for determining *what* responses to teach and take data on (i.e., *dependent variables, DV*)
- Methods to determine accuracy of data (i.e., reliability of DV )
- Methods to determine if a *teaching procedure* was correctly done (i.e., *fidelity/integrity of procedures*)
- Guidelines: graphic displays, interpretation, evaluation of data and teaching decisions (i.e., *continue, modify, or find a different procedures*)

11

11

## A brief history of Applied Behavior Analysis (ABA)



12

12



## Behavior Analysis: 1920s- 1940s

Not Applied:

- Scientists interested in behavior changes
  - rats running mazes
    - what environmental events makes them run faster
  - how teach a pigeon to peck a red light v. a white light
    - what environmental events are useful to teach a selection response (discrimination training)
  - environmental events that make an animal "work" longer (intermittent schedules of reinforcement)

13

13

## Pigeons Playing Ping Pong



14

14

Baer, Wolf, & Risley (1968)

Defined the Field of Behavior Analysis

(changing behavior a.k.a., "teaching" and "learning")

7 Dimensions of Applied Behavior Analysis

Applied

Behavioral

Analytical

Conceptual Systems

Technological

Effective

Generality

15

15

How are today's professional  
Behavior Analysts doing  
with respect to the 7 dimensions?

What would ASD treatment look like  
when 7 dimensions are applied?

16

16

My job: Teach ASD children in a classroom in a  
public school using ABA  
(the 7 dimensions-whatever they mean).

"How do I know my teaching is?"

Applied

17

17

Applied: Interested in Important behaviors



A person



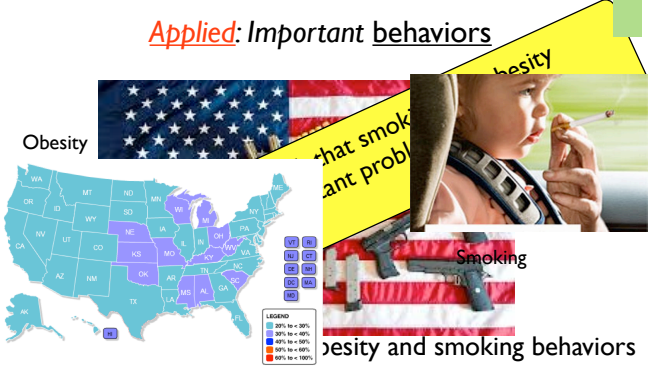
A culture

What behaviors are important?

18

18

## Applied: Important behaviors

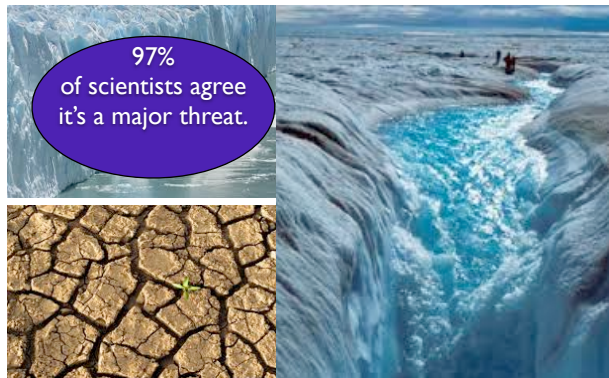


Cultural discussion still going on

19

19

Climate change a major problem?



20

20

## Applied

What are important behaviors in education?

How can we identify important behaviors?

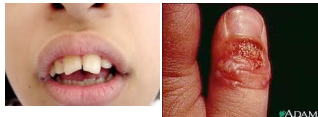
21

21

## Is “important” defined by a topography?



Finger sucking  
in an infant



Finger sucking  
in an older child

Can't define “important behaviors”  
a topography

Context is important: age, location, audience

Adverse affect on person or culture: truancy, literacy, computer  
and social competence

22

22

## *Applied: Important behaviors*

What behaviors are important?



making a bed



stereotypy

Not all bothersome behaviors are important

23

23

## *Applied: Important behaviors*

### • What About: Stereotypy

- Motor:
  - hand flapping
  - rocking
- Vocal:
  - scripting
  - humming
- Visual:
  - waving fingers
  - tracking lines
  - spinning object

### Considerations:

- Context: age, location, audience
- Treatment is often time consuming and ineffective
- What other reinforcers?
  - will reducing object spinning eliminate the only Sr?
  - can other stimuli be conditioned as competing reinforcers?
  - can access to stereotypy be used as reinforcer?
- Will treatment interfere with skill acquisition programs?
  - is reducing stereotypy more important than learning to mand?

24

24

## Ethical issues:

What behaviors are important and when they are important?

Choosing between deceleration and acceleration targets

What's more important?

- Reduce injurious head hitting or teaching child to write his name?
- Teach attending to teacher or reducing ear playing?
- Reduce hand flapping or teach motor imitation/mands to a non-verbal child?
- Reduce vocal stereotypy or increase the number of reinforcing activities for a person?

25

25

## How do I know if my teaching is (ABA) “Applied?”

- Education is dealing with this: Core curriculum v. local control
- Cultural value v. value of local control
- Burden and cost to society
  - skilled workers, scientists, doctors,
- ASD children
  - cost to culture v personal value to child
- Outcomes: immediate v. deferred outcomes (behavioral cusps)
  - behavioral cusps: looking at person, mand training
  - language skills v. skills of daily living
  - academic v. custodial programs

26

26

Ok, I'm teaching a behavior (skill) that's important to the culture and person, but....

How do I know my teaching is

Behavioral

27

27

Behavioral means: A response is

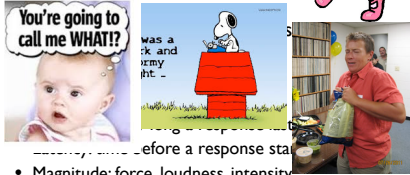
Observable

Some Response Examples

- walking, typing, reading aloud
- opening a peanut bag
- naming things, conversing



Countable



- Magnitude: force, loudness, intensity

28

28

Behavioral means: A response is

Observable

Precise o

Countable



How many tricks does the rat do? 3? 6?

How many goals does the black rat make?

29

29

Applied Behavioral is Observable and Countable

Culturally value

Personal value



How many unassisted steps taken?

30

30

## Behavioral

Defining  
what to  
count

**Operational definitions**

Defining  
when to  
count it

31

31

## Behavioral

Defining  
what to  
count

**Operational definitions**

Defining  
when to  
count it

what is knowing  
understanding,  
expressing, and  
confident

Knows animals, numbers  
Understands class rules  
Expresses feeling appropriately  
Is socially confident in groups

when do you count  
knows,  
understands,  
expresses, confident

32

32

## Academic operational definitions

Knows animals:

count (re

l: say, nam

Write

Sign

Knows numbers:

c: point, giv

r motor re

Knows letter sounds:

Knows the classroom  
rules:

Knows how to play:

(function):

ing

Tact: name something

Echoic: repeat an audible sound

Intraverbal: talk about something in its absence

33

33

## Academic operational definitions

What does “*know*” mean?

“Knowing” as doing

- say names of numbers when asked
- responding to “count to 100?”
- reads numbers to 100 on number line
- selects numbers to 100 when asked
- follows classroom rules in classroom

“Knowing” as saying

- state classroom rules when asked
- says numbers in a song, e.g. “5 little monkeys”

34

34

## Operational definitions of problem behavior

*Self-injurious behavior*: Any forceful hand hit to the body

*Property destruction*: Any destruction of materials

*Aggression*: Any forceful and unwanted contact or attempted contact with another person

Operational definitions should include response examples and non-examples

When possible CALL A RESPONSE by its topography, not a presumed class, e.g., head hitting v. SIB.

35

35

## Behavioral

Defining what to count

- mand (things, information)
- tact (feature, function, class)
- echoic
- intraverbal
- Vocal, written, signed responses
- Other motor Rs

Operational definitions

Defining how to count it

- cumulative frequency
- duration/latency
- intensity/force
- rate (frequency/time)
- # of different topographies
- percentage correct

36

36



## Teaching practices: Assessing progress

### Common teaching practices (letter names and sounds)

- Pre-test
  - Treatment: general teaching on all names and sounds*
- Post-test
  - percent (80%) mastery

### Behavioral teaching practices (letter names and sounds)

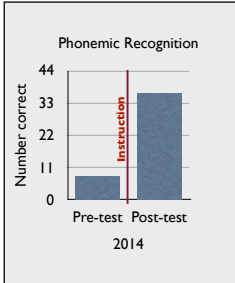
- Baseline (# correct)
  - Treatment: specific teaching on sounds/names missed*
- Repeated measures
  - graphic displays of progress
  - goal: 100% correct

37

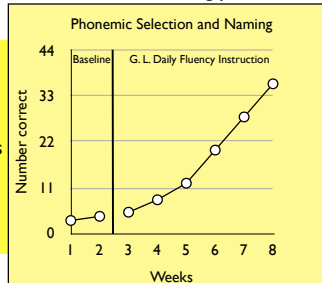
37

## Graphic Displays

Common teaching practice



Behavioral teaching practice



38

38

## Commonality of Applied Behavior Analysis and Teaching

- Changing important behavior
  - Personally and culturally value
  - the dependent variable (DV) of our teaching procedures
- Response (DV) is observed and counted in some way
  - \*changes in the DV may our validate teaching (procedure)

\* Or the results may show that our (teaching) procedure was ineffective.

39

39

Behavioral: A response is observed and counted

Whose responses are counted?

The Teacher?



The Student?



40

40

---

---

---

---

---

---

---

---

---

---

Things that affect our observing and counting of our DV

Who's behavior is counted?

41

41

---

---

---

---

---

---

---

---

---

---

Who's behavior was counted?

FAR  
FAR  
FAR

42

42

---

---

---

---

---

---

---

---

---

---

# Who's behavior was counted?

McGurk Effect

Tritone Illusion

Shepard  
Tone Illusion

Weak  
operational  
definitions

Poorly trained  
observers

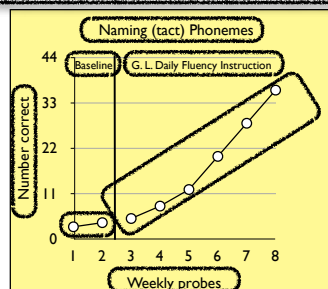
Observer  
Motivation

43

43

So What? Data are data?

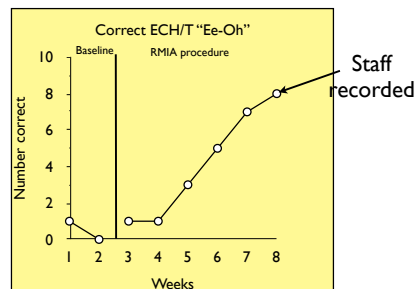
**What could go wrong?**



44

44

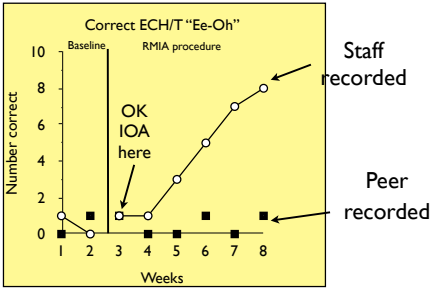
## Data Reliability



45

45

Data Reliability



46

46

Reliability of DV (IOA): A Case Study

Teaching Challenge: Non-vocal child

- Sign language: not evoking vocal responding
- Echoic training / not effective
- Stimulus-stimulus pairing / not effective

Rapid Motor Imitation Antecedent (RMIA)

- Procedure:
- “do this”: 6 rapid motor imitation responses
- show picture, give an echoic model

“Eeyore”



47

47

RELIABILITY (IOA) PRACTICE

Target: vocal Tact/Echoic “Eeyore”

Instructions:  
Session: Circle “C” correct or “X” if incorrect.  
IOA % (Percent Interobserver Agreement): Circle “Y” if match and “N” if no match. Calculate IOA = #Ys / #Ys + #Ns

Trial#	Session 1	IOA 1	Session 2	IOA 2
1	C X	Y N	C X	Y N
2	C X	Y N	C X	Y N
3	C X	Y N	C X	Y N
4	C X	Y N	C X	Y N
5	C X	Y N	C X	Y N
6	C X	Y N	C X	Y N
7	C X	Y N	C X	Y N
8	C X	Y N	C X	Y N
9	C X	Y N	C X	Y N
10	C X	Y N	C X	Y N
# correct		% IOA	# correct	% IOA

48

48

[illegible]
$$Y_s =$$
$$N_s =$$

number is our session data. Goes on a graph.

percentage tells how reliable the data on that session is.

50

[illegible]

“Eeyore”: too difficult and not defined well.  
We decided to shape “Eeyore”

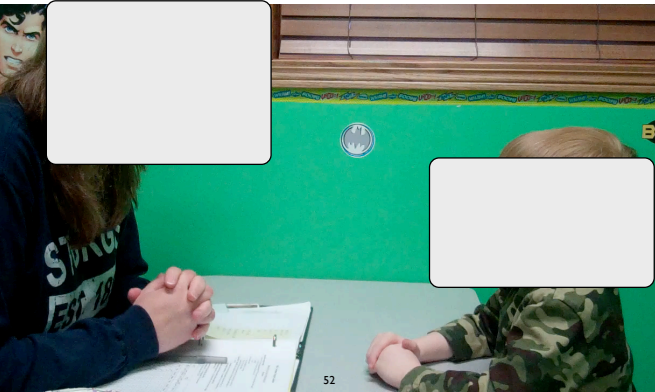
1. **Dependent VARIABLE:** first vocalization within 5s following either teacher model and lasting 1-2s,
  - a. EE-OH
    - i. **Examples.** ee-oh, ee-uh must glide together without break and last 1-2s
    - ii. **Non-examples. separation** ee...oh; **repeats:** ee-ee, oh-oh; **wrong vowel** ee-ai

[illegible]

- 51

51

Target: glided “Ee-Oh”



52

---

---

---

---

---

---

---

---

---

---



---

---

---

---

---

---

---

---

---

---

Reliability (IOA) Practice

Target: vocal Tact/Echoic: **Ee-Oh glide**

Instructions:  
 Session: Circle "C" correct or "X" if incorrect.  
 IOA % (Percent Interobserver Agreement): Circle "Y" if match and "N" if no match. Calculate: IOA = #Ys / #Ys + #Ns

Trial	Session 1	IOA 1	Session 2	IOA 2
1	C X	Y N	C X	Y N
2	C X	Y N	C X	Y N
3	C X	Y N	C X	Y N
4	C X	Y N	C X	Y N
5	C X	Y N	C X	Y N
6	C X	Y N	C X	Y N
7	C X	Y N	C X	Y N
8	C X	Y N	C X	Y N
9	C X	Y N	C X	Y N
10	C X	Y N	C X	Y N
# correct	2	% IOA	# correct	% IOA

Ys =  
Ns =

Ys / Ys + Ns =

number is our session data. Goes on a graph.

percentage tells how reliable the data on that session is.

53

---

---

---

---

---

---

---

---

---

---

## Not waiting to Respond:

“Ee-Oh” as we were saying it.

Possible Solutions

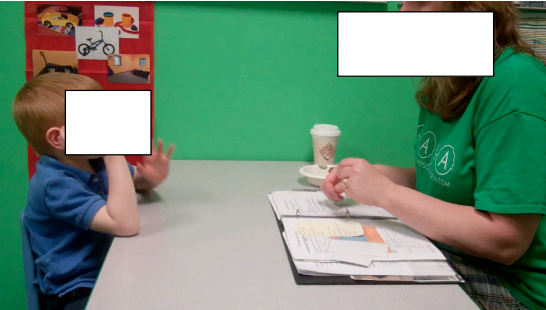
- Implement “wait” program
- Use concurrent R to strengthen accuracy

1. **Dependent VARIABLE:** Concurrent vocalization with either teacher model and lasting 1-2s,

- Ee-Oh
  - Examples.** ee-oh, ee-uh must glide together without break and last 1-2s
  - Non-examples.** separation ee...oh; repeats: ee-ee, oh-oh; wrong vowel ee-ai

54

Concurrent ECH/Tact Model: Ee-Oh glide



55

55

Reliability (IOA) Practice

Target: vocal Tact/Echoic: **Concurrent Ee-Oh glide**

Instructions:  
Session: Circle "C" correct or "X" if incorrect.  
IOA % (Percent Interobserver Agreement): Circle "Y" if match and "N" if no match. Calculate: IOA = #Ys / #Ys + #Ns

Trial	Session 1	IOA 1	Session 2	IOA 2
1	C X	Y N	C X	Y N
2	C X	Y N	C X	Y N
3	C X	Y N	C X	Y N
4	C X	Y N	C X	Y N
5	C X	Y N	C X	Y N
6	C X	Y N	C X	Y N
7	C X	Y N	C X	Y N
8	C X	Y N	C X	Y N
9	C X	Y N	C X	Y N
10	C X	Y N	C X	Y N
# correct	8	% IOA	# correct	% IOA

Ys =  
Ns =

Ys / Ys + Ns =

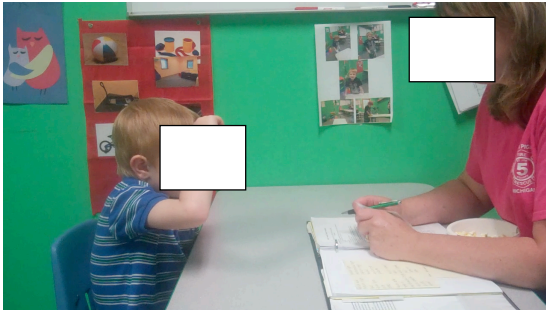
number is our session data. Goes on a graph.

percentage tells how reliable the data on that session is.

56

56

Ee-Oh Tact Follow up



57

57

Target: vocal Tact/Echoic

## Concurrent Ee-Oh glide

Instructions:  
 Session: Circle "C" correct or "X" if incorrect.  
 IOA % (Percent Interobserver Agreement). Circle "Y" if match and "N" if no match. Calculate IOA = #Ys / #Ys + #Ns

Trial	Session 1	IOA 1	Session 2	IOA 2
1	C X	Y N	C X	Y N
2	X	Y N	C X	Y N
3	C X	Y N	C X	Y N
4	C X	Y N	C X	Y N
5	C X	Y N	C X	Y N
6	C X	Y N	C X	Y N
7	C X	Y N	C X	Y N
8	C X	Y N	C X	Y N
9	C X	Y N	C X	Y N
10	C X	Y N	C X	Y N
# correct	8	% IOA	# correct	% IOA

Ys =  
Ns =

Ys / Ys + Ns =

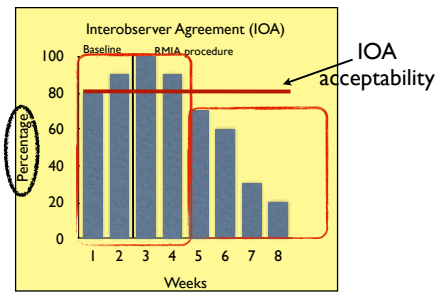
this number  
is our session data.  
Goes on graph.

58

this percentage  
tells how reliable  
that session is.

58

## Reliability (IOA)



Few correct  
responses

More correct  
responses

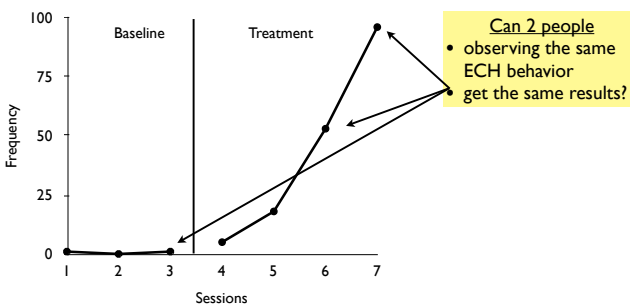
59

59

Is the child's ECH responding improving?

Have we asked: Are the data reliable?

Correct Echoics



60

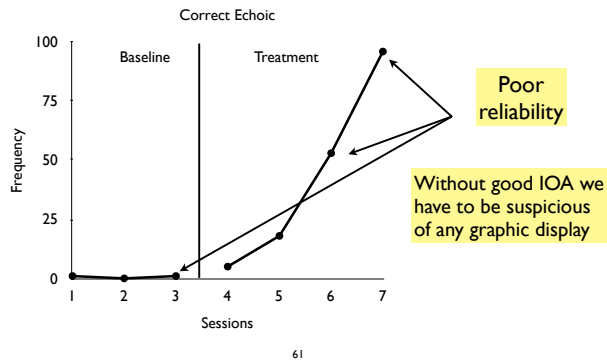
60



## Is the child's ECH responding improving?

IOA = 50%:

2 observers agree only half the time that an echoic response is correct

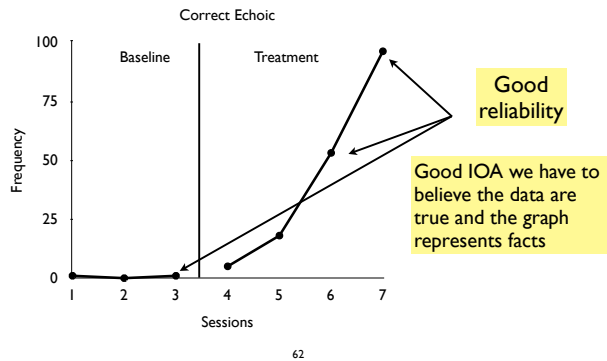


61

## Is the child's ECH responding improving?

IOA = 90%:

2 observers agree only half the time that an echoic response is correct



62

## Reliability of Observed Behavior

### Why is Reliability Important?

Can't interpret graphic displays without it.

Without IOA

- Believability of data and graph is uncertain
  - Graphic display may be false
- What was observed?
  - Is operational definition clear?
  - Were "expectation" influencing IOA
- Can't make program changes without good IOA

63

## Threats to dependent variable (DV) reliability

- Ambiguous operational definition
  - staff saying, "looks like SIB," "I think that's a correct ee"
- Poorly designed data sheet
  - too complicated, incomplete (e.g., no instructions)
- Poor staff training
  - can't identify a target response
  - inaccurate take data
- Poor staff oversight
  - don't take data
  - make up data to fill in data sheet at end of day
- Organizational values, unawareness, or indifference
  - it's not important to the organization

64

64

## Effects on Teaching....

Rule: Can't evaluate graphic displays without evidence of the believability of the data, i.e., IOA.

Can't make informed program decisions without believable data

65

65

Did they say we have to collect IOA on every session!  
No Way!



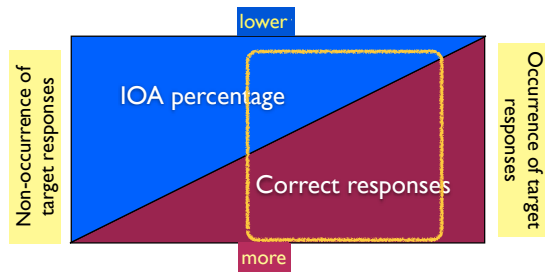
How much IOA?

- research standard: 33%
- no APBA practice standards

66

66

## Possible IOA Outcomes



67

67

## Collecting IOA suggestions

Acquisition Targets  
(# of mands and tacts, play duration)

Deceleration Targets  
(SIB, Pica, disruptive behavior, PA, tantrum)

Difficult without:

- Dedicated trained staff
- BCBA Supervisor

possible with

Instructional data collection

- 10 trial data
- daily probes

when

Target most likely

- later in acquisition training
- earlier with problem behavior

and

Brief samples

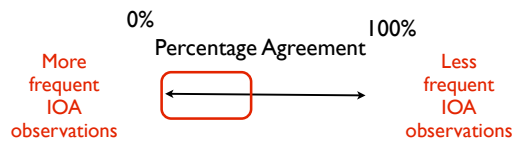
- Lead teacher/consult
- short video tape

68

68

## Behavioral: IOA Tip

Look at *IOA percentage*.

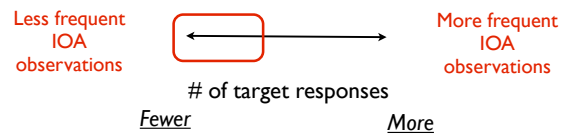


69

69

## Behavioral: IOA Tip

Look at *Occurrence of target behavior*



70

70

Believable data/graphic displays:

It's a professional ethics issue!

71

71