Observation & Measurement II:
Observational Methods
Interobserver Agreement
Treatment Fidelity

Class 6
SPCD 619

T A B L E 1
Quality Indicators Within Single-Subject Research

Horner et al., 2005, p. 174

Dependent Variable

- Dependent variables are described with operational precision.

- Each dependent variable is measured with a procedure that generates a quantifiable index.

- Measurement of the dependent variable is valid and described with replicable precision.

- Dependent variables are measured repeatedly over time.

- Data are collected on the reliability or interobserver agreement associated with each dependent variable, and IOA levels meet minimal standards (e.g., IOA = 80%; Kappa = 60%).

Operational Definitions

- Are objective – refer to observable components of the behavior
- Are clear and unambiguous
- Require little or no inference

- Include
  - A descriptive name
  - General definition
    - Elaboration that describes the critical parts of the behavior
  - Typical **examples** of the behavior
  - **Near nonexamples** of the behavior

SIMPLIFIED HABIT REVERSAL TREATMENT FOR CHRONIC HAIR PULLING IN THREE ADOLESCENTS: A CLINICAL REPLICATION WITH DIRECT OBSERVATION

JOHN T. RAPP, RAYMOND G. MILTENBERGER, ETHAN S. LONG, AMY J. ELLIOTT, AND VICKI A. LUMLEY

NORTH DAKOTA STATE UNIVERSITY

- *Target Behavior and Data Collection*

"Hair pulling was defined as touching the fingers to the scalp, eyebrow, or eyelashes for all 3 participants."

Coding Manuals

☐ "a set of rules, definitions, examples, and near nonexpamples that guide the observers in counting and/or indicating the duration of the behaviors of interest" (p. 36, Yoder & Symons, 2010)

☐ Ways to find these
  ■ Obtain from authors of previous studies
  ■ Write your own using definitions from the existing literature or definitions developed from experts

---

Coding Manuals should contain:

Behaviors
    Operational definitions
    3 examples and near nonexamples
    Symbols/codes used
    Log of changes that you make as you change/clarify/update
            system

Coding Procedures
    General guidelines
    What to do to prepare for observation
    "   "   during the observation
    "   "   " after the observation
    Sample data forms or screen shots of computerized systems

Reliability
    Plan for conducting IOA checks
    Plan for conducting procedural reliability checks

Data Collection

1. **Pose an experimental question**
2. **Identify and define relevant behaviors**
3. **Select behavioral dimension(s)**
4. **Select your measurement system**
    ☐ **Consider the observational settings**
    ☐ **Consider the advantages and disadvantages**
5. **Identify two or more people to conduct the observations**

T A B L E 1
Quality Indicators Within Single-Subject Research

Dependent Variable

• Dependent variables are described with operational precision.

• Each dependent variable is measured with a procedure that generates a quantifiable index.

• Measurement of the dependent variable is valid and described with replicable precision.

• Dependent variables are measured repeatedly over time.

• Data are collected on the reliability or interobserver agreement associated with each dependent variable, and IOA levels meet minimal standards (e.g., IOA = 80%; Kappa = 60%).

Selecting a Measurement System(Yoder & Symons, 2010)

- ☐ *System has*
  - ■ *Behavior sampling method*
  - ■ *Participant sampling method*
  - ■ *Coding decision recording method*

Behavior Sampling

- ☐ **Continuous**
  - ■ Code entire observation session
    - ☐ Most complete but most expensive method
  - ■ Can use a computer system to measure onset/offset {timed event} or use a simple tally system {event recording}
- ☐ Intermittent
- ☐ Interval

Event Recording

- ☐ Advantages:
  - ☐ Useful across many topographies of discrete behaviors (movement cycles)
  - ☐ Easy to use – most direct and accurate measure of the number of times a behavior occurs
- ☐ Disadvantages
  - ☐ Requires continuous observation
  - ☐ Can't be used with continuous (i.e., non-discrete) data
  - ☐ Inaccurate with high frequency behaviors
  - ☐ Confounds frequency with duration of responding

The Use of Social Stories to Promote Independent Behaviors in Novel Events for Children with PDD-NOS

Michelle L. Ivey, L. Juane Heflin, and Paul Alberto

Data Collection

Five behaviors in four categories were targeted for each of the novel routines. Data on the occurrence or nonoccurrence of specific behaviors were collected during all target sessions via event recording on a checklist. The four be-

FIGURE 2. Number of targeted skills observed during therapy activities without and with Social Story preparation. The type of activity is coded as follows: S = setting change, P = purchase activity, T = novel toy, E = novel event.

TABLE 1
Examples of Target Skills for Each Novel Event Category

Skill	Setting change	Purchase	Novel toy	Novel event
Example	Memorial fountain	Fund-raiser purchase	Kerplunk	Make a video
Attention	Remain on task	Remain on task	Remain on task	Remain on task
Key Task 1	Comment	Choose a snack	Push sticks through tube	Follow a script
Key Task 2	Read a sign	Give money	Twist the tube	Say, "cut"
Vocabulary	Memorial	Delicious	Kerplunk	Rehearsal
Request	Directions	Ask what adult wants	Marbles	Ask how to turn camera on

Behavior Sampling

☐ Continuous
☐ **Intermittent**
- ■ "periodic intervals are observed and **all** instances of key behavior occurring in that interval are recorded" (Yoder & Symons, 2010,p. 55)
  - ☐ Cheaper method; can do timed event or event sampling; not used as much as other methods

☐ Interval

Behavior Sampling

☐ Continuous
☐ Intermittent
☐ **Interval** (estimates duration/frequency)
- ■ Divides observation into discrete intervals and records whether a response occurred…
…throughout an interval (**whole** interval)
…during a fixed time period at the end of an interval (**momentary** time sampling)
…during any time in an interval (**partial** interval)

- ■ Is commonly used

JOURNAL OF APPLIED BEHAVIOR ANALYSIS   2002, **35**, 79–83   NUMBER 1 (SPRING 2002)

INCREASING SOCIAL INITIATIONS IN CHILDREN WITH AUTISM: EFFECTS OF A TACTILE PROMPT

Daniel B. Shabani, Roger C. Katz, David A. Wilder, and Kenneth Beauchamp

UNIVERSITY OF THE PACIFIC

AND

Crystal R. Taylor and Kirsten J. Fischer

CENTRAL VALLEY AUTISM PROJECT

The tactile prompt was a JTECH Series 27 pager that vibrated for 3 to 5 s when activated by a remote control. The vibrating pager was small enough to fit in the participant's pocket.

Verbal initiations. Verbal initiations were any vocal verbalizations made by the participant that were related to the play activity *and* were directed to another child. For example, if the participant showed his toy to a peer and said, "Look what I have," this was recorded as a verbal initiation. Verbal initiations that were unrelated to the play context were not scored.

PERCENTAGE OF INTE
70
60
50
40
30 Verbal initiations
20
10
0
1 2 3 4 5

Prompt   Fading
Criterion

Whole Interval Recording

Does the behavior occur **throughout** the interval? (i.e., during the entire interval)

	1	2	3	4	5	6	7	8	9	10	
Whole							X				10%
Behavior	└─┘ └─┘			└─┘		└─┘	└─┘		└─┘		

Whole Interval Recording

Student: Alice
Start time: 8:05   **End time:** 8:20   **Setting: morning recess**
Definition of behavior:
Social Interaction: within 2 feet of another child with interactions
Parallel Play within 2 feet of another child, no interactions
Alone: further than 2 feet from anyone
Organized Play: interacting with others in an organized game or activity
No Code: none of the above.
Observer: Mr. Fables (paraprofessional)

Time	Social Interaction		Parallel Play	Alone	Organized Play		No Code
	+	-			+	-	
10"	x						
20"			x				
30"			x				
40"				x			
50"			x				
60"			x				
70"					x		
80"						x	
90"						x	
100"							x

Key: 10" = 10 seconds.

Whole Interval Recording

☐ Advantages
  ☐ Provides an estimate of behavior that occurs continuously
  ☐ Can be converted to percentage
☐ Disadvantages
  ☐ Tends to **underestimate** overall occurrence of responding
  ☐ Requires continuous observation (fatigue, time)
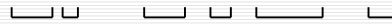
Momentary Interval Recording

Does the behavior occur **during a fixed period at the end of** the interval?

	1	2	3	4	5	6	7	8	9	10	
Momentary	X			X	X	X	X				50%
Behavior											

Momentary Interval Recording

☐ Advantages
  ☐ Very efficient for observers
  ☐ Can observe multiple students
  ☐ Does not require continuous observation

☐ Disadvantages
  ☐ Can **underestimate** responding
  ☐ Requires some type of cueing device

Partial Interval Recording

Does the behavior occur **at any time** during the interval?

	1	2	3	4	5	6	7	8	9	10	
Partial	X	X		X	X	X	X		X		80%
Behavior	⎵	⎵		⎵	⎵	⎵__		⎵			

Partial Interval Recording

Student: John
Start time: 8:05   **End time:** 8:15   **Setting:** Seventh-grade English class
Definition of behavior: On-task behavior is defined as appropriately writing, reading, talking about the assignment, or waiting to ask the teacher a question regarding the assignment**.**
Observer: Ms. Geralds (recess monitor)

minute	10"	20"	30"	40"	50"	60"
1	O	O	X	X	X	O
2	X	O	O	O	X	O
3	X	O	O	X	O	X
4	X	X	X	O	O	X
5	O	O	X	O	X	O
6	O	X	X	X	X	X
7	X	O	O	O	O	O
8	O	X	X	X	X	O
9	X	O	O	O	X	X
10	O	X	X	O	X	O

Key: O = Behavior was not observed at all during the 10-second interval.
       X = Behavior was observed at least once during the 10-second interval.
       10" = 10 seconds.

Partial Interval Recording

☐ Advantages
  ☐ Provides an estimate of behaviors that occur for a brief period of time
 ☐ Can be used with high frequency behaviors

☐ Disadvantages
  ☐ Tends to **overestimate** the occurrence of the behavior if larger intervals are used or **underestimate** if behavior occurs very frequently
  ☐ Requires some type of cueing device

Considerations

(from Yoder & Symons, 2010)

If the behavior of interest
☐ Is of short duration – consider counting number (frequency) (partial interval)
☐ "represents a state (e.g., attention)" or has a long duration, consider duration (continuous or momentary interval)

Alberto & Troutman, 2006)
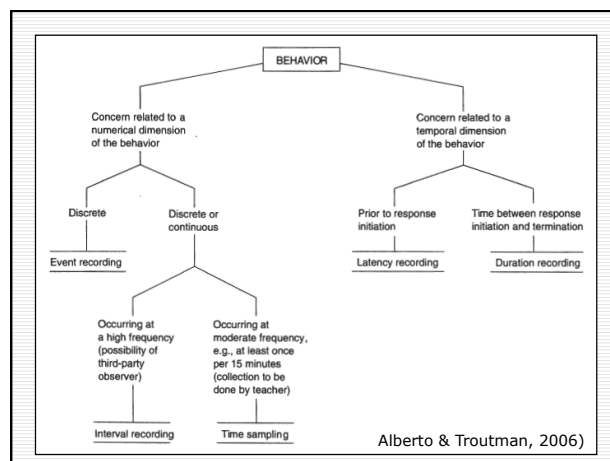
Systems

☐ **Participant Sampling** (from Yoder & Symons, 2010)
   ▪ Who do you code if you have more than one participant?
      ☐ Focal – code one participant for a designated time, then another, then another, etc.
      ☐ Multiple pass – code on only 1 participant for the entire session; frequently record session and make multiple passes, coding 1 participant at a time
      ☐ Conspicuous Sampling – observe the group and record on individuals engaged in "predefined conspicuous behaviors"

Systems

☐ **Coding Decision Recording Method**
   ▪ Direct, real-time observation
   ▪ Recording for later coding (audio or video)
   ▪ Paper/Pencil
   ▪ Computer-based program

Where & When to Observe

☐ Consider relation of setting to the research question

☐ Observations should reflect what occurs in the settings (events)
   ■ Selective sampling

TABLE 1
Quality Indicators Within Single-Subject Research   Horner et al., 2005, pl. 174

Dependent Variable

• Dependent variables are described with operational precision.

• Each dependent variable is measured with a procedure that generates a quantifiable index.

• Measurement of the dependent variable is valid and described with replicable precision.

• Dependent variables are measured repeatedly over time.

• Data are collected on the reliability or interobserver agreement associated with each dependent variable, and IOA levels meet minimal standards (e.g., IOA = 80%; Kappa = 60%).

Choosing and Training Observers

☐ Choose individuals who have sufficient time to complete training and are dependable.
☐ Provide sufficient training prior to beginning study
   ■ Overview of study (but may want naïve observers)
   ■ "how tos" of data collection
   ■ Copy of the code book
   ■ Practice, practice, practice (adjust coding rules as new considerations arise)
   ■ Establish a minimum criterion for IOA before proceeding with data collection (80 – 90%)
   ■ Retrain periodically
   ■ Show your appreciation

Reliability

☐ Consistency of measurement across conditions, regardless of who is the observer
☐ Reliability is NOT a measure of the "truth" of the behavior! It is a measure of <u>consistency</u>
   ■ Interobserver Reliability or Agreement (IOA)
   A second observer <u>independently</u> measures the target behavior(s). The two results are compared for accuracy, yielding a coefficient or percentage of agreement.

IOA: Total Percent of Agreement
(frequency-ratio approach)

$$\frac{\text{smaller number}}{\text{larger number}} \times 100\% = \text{percent of agreement}$$

Observer	Number of Times Student Raises Hand							Total Recorded
Observer 1	x		x		x	x	x	5
Observer 2		x	x	x			x	4

$$\frac{4}{5} \times 100\% = 80\% \text{ agreement}$$

IOA: Interval Agreement
(combined, point-by-point)

$$\frac{\text{agreements}}{\text{disagreements+agreements}} \times 100\% = \text{percent of agreement}$$

Observer	Number of Times Student Raises Hand						
Observer 1	x		x		x	x	x	Total = 5
Observer 2			x	x			x	Total = 3

$$\frac{3}{4+3} \times 100\% = 43\% \text{ agreement}$$

IOA: Interval Agreement
(combined, point-by-point)

$$\frac{\text{agreements}}{\text{disagreements+agreements}} \times 100\% = \text{percent of agreement}$$

Observer	Number of Times Student Raises Hand						
Observer 1	x							Total = 1
Observer 2								Total = 0

$$\frac{6}{1+6} \times 100\% = 86\% \text{ agreement}$$

IOA: Occurrence/Nonoccurrence Agreement

1. $\dfrac{\text{Agreement of Occurrence}}{\text{Agreement + disagreement}} \times 100\% = \text{percent of agreement}$

2. $\dfrac{\text{Agreement of Nonoccurrence}}{\text{Agreement + disagreement}} \times 100\% = \text{percent of agreement}$

Report as two separate statistics.

How Frequently to Conduct Reliability Checks

- ☐ Collect IOA in a minimum of 20 - 33% of observational sessions <u>per condition</u>, for each participant for each dependent variable
  - ■ Use higher percentage if situation calls for this

Procedural Reliability (sometimes also called Treatment Integrity or Fidelity)

- ☐ Measure of how the closely the intervention was implemented according to plan
- ☐ Requires
  - ■ Operational definition of the intervention (IV)
  - ■ Decision of what aspect(s) of intervention to measure
  - ■ Training of observers to conduct assessment of intervention implementation

THE TREATMENT OF SEVERE BEHAVIOR PROBLEMS IN SCHOOL SETTINGS USING ATECHNICAL ASSISTANCE MODEL
Northup et al., 1994)

<u>Treatment integrity</u>. The five teacher responses to student target behaviors were scored to assess the integrity of all assessment and intervention procedures. Each assessment and intervention procedure always specified a particular consequence that was to follow any occurrence of target behavior immediately. The percentage of student target behaviors that were followed (within two 10-s intervals) by the specified consequence served as an index of treatment integrity and provided a measure of the accuracy with which assessment and intervention procedures were implemented. Treatment integrity was scored only when the behavior for which a consequence was prescribed occurred.

Thinking Ahead on the Research Proposal: At this point you should

- ☐ Have identified a general topic
- ☐ Begun locating and reading relevant literature so that you can
  - ■ More clearly focus the specific research question you are asking
  - ■ Begin to define possible targets (DVs) and develop possible interventions (IV)

Next Week

- ☐ Discuss and practice visual analysis
  - ◼ Do the readings so you can participate in class activities

- ☐ Bring a laptop and flash drive to class