Experimental Reasoning: Multiple Baseline Designs

Philip N. Chase
Simmons University

Problems with Reversals

- We described the following problems of reversal designs:

  - Ethics of Stability
  - Ethics of Replication
  - Irreversible effects
  - Multiple treatments or conditions and order effects

Problems with Alternating Treatments

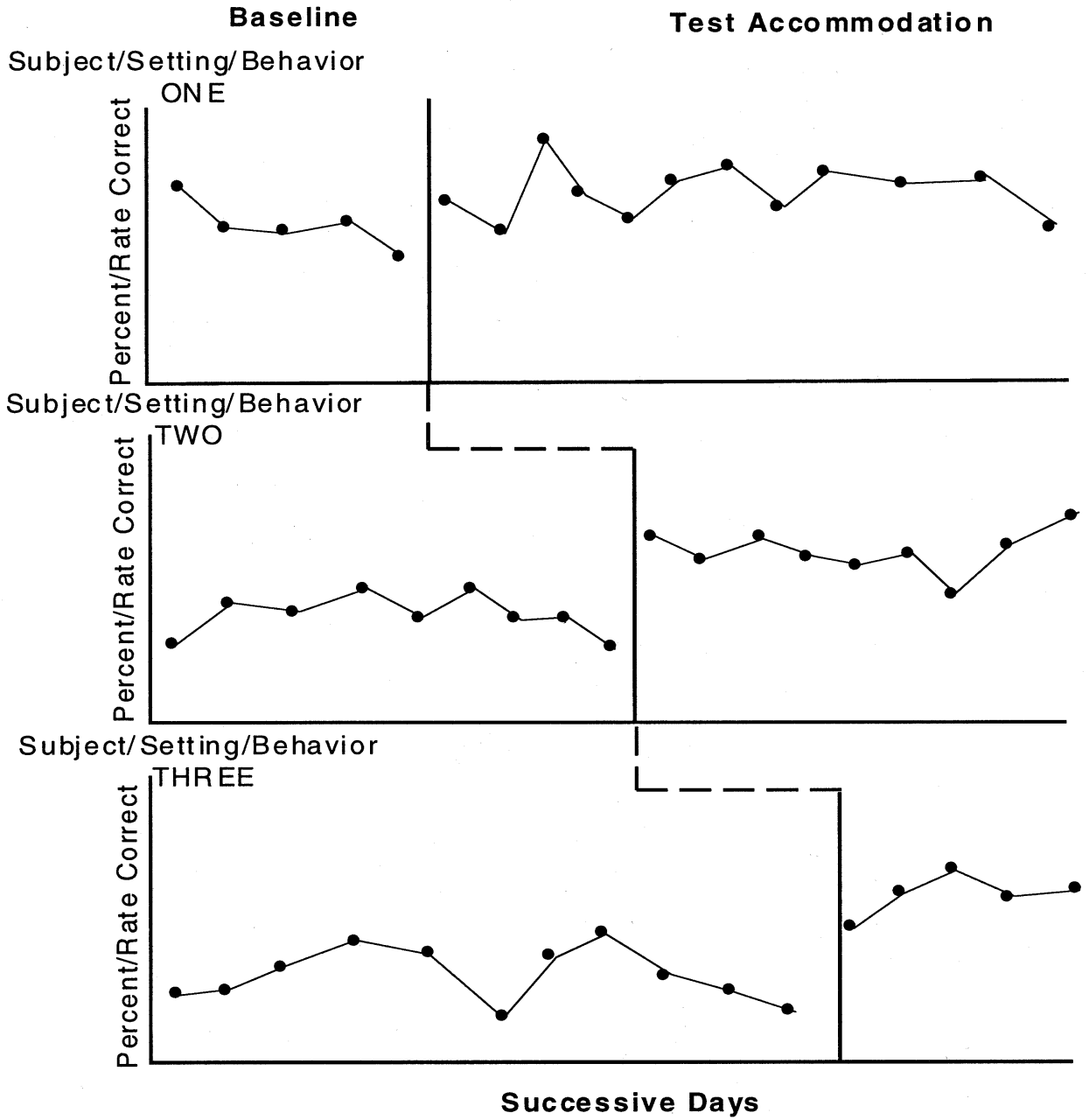
- Only partially minimizes ethics of replication
- Doesn’t manage problems of irreversibility

Multiple Baselines to the Rescue

- The introduction of a treatment in **staggered** fashion across behaviors, settings, or subjects so that the condition can be repeatedly compared to baseline conditions.

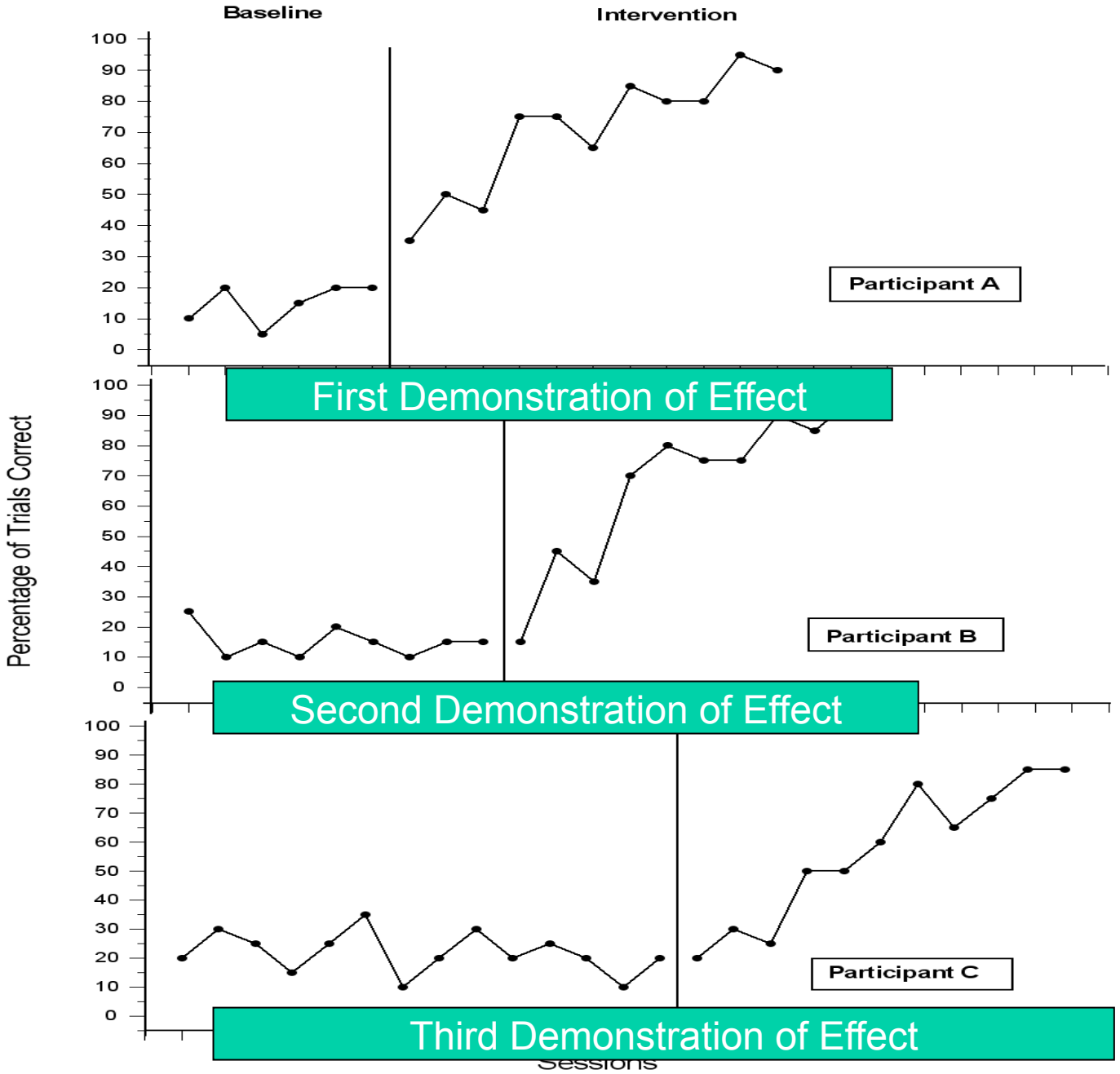
- Not this

Staggered →


Baseline
Test Accommodation
Subject/Setting/Behavior
ONE
Percent/Rate Correct
Subject/Setting/Behavior
TWO
Percent/Rate Correct
Subject/Setting/Behavior
THREE
Percent/Rate Correct
Successive Days

Logic of Multiple Baseline

- Baseline to treatment change assessed repeatedly
- Baselines and treatments of different lengths of time minimize maturation and some other threats
- Replication across settings, behaviors, or subjects
- Difficult to examine multiple treatments and order effects


Baseline
Intervention
Participant A
First Demonstration of Effect
Participant B
Second Demonstration of Effect
Participant C
Third Demonstration of Effect
Percentage of Trials Correct
Sessions

Multiple Baseline Designs

- Most widely used for evaluating treatment effects in ABA
- Highly flexible
- Main advantage: You do not have to withdraw, reverse or alternate a treatment

Teams

- State two reasons for using a multiple baseline design (MBD) instead of a reversal design.

Five Kinds

- Multiple baseline across behaviors
- Multiple baseline across participants
- Multiple baseline across settings
- Multiple probe
- Delayed multiple baseline

Multiple Baseline Across Behaviors

- 2 or more different behaviors of same subject
- Subject serves as his/her own control
- After steady state baseline responding, independent variable is applied to 1st behavior, while other behaviors are kept in baseline conditions
- When steady state responding is reached for 1st behavior, then IV is applied to next behavior

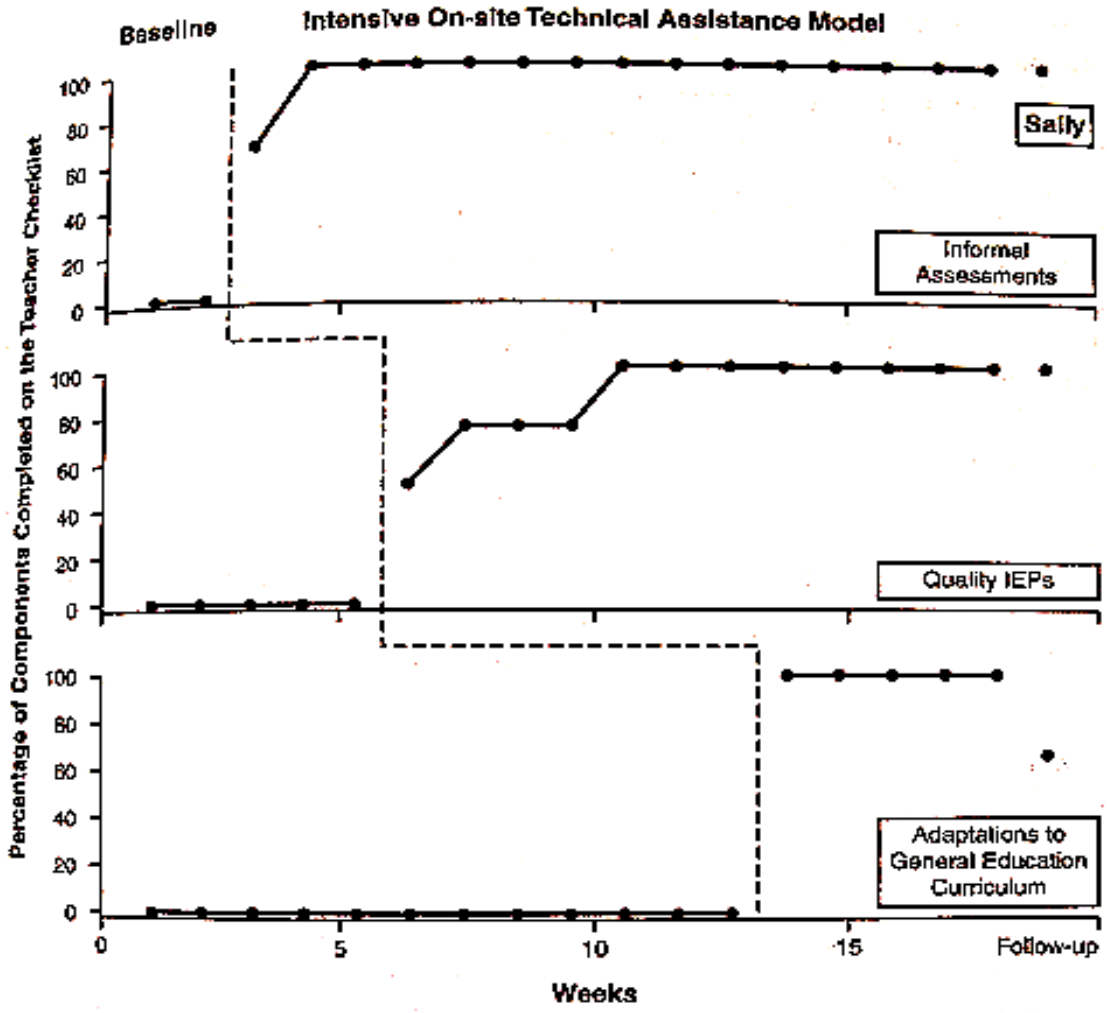
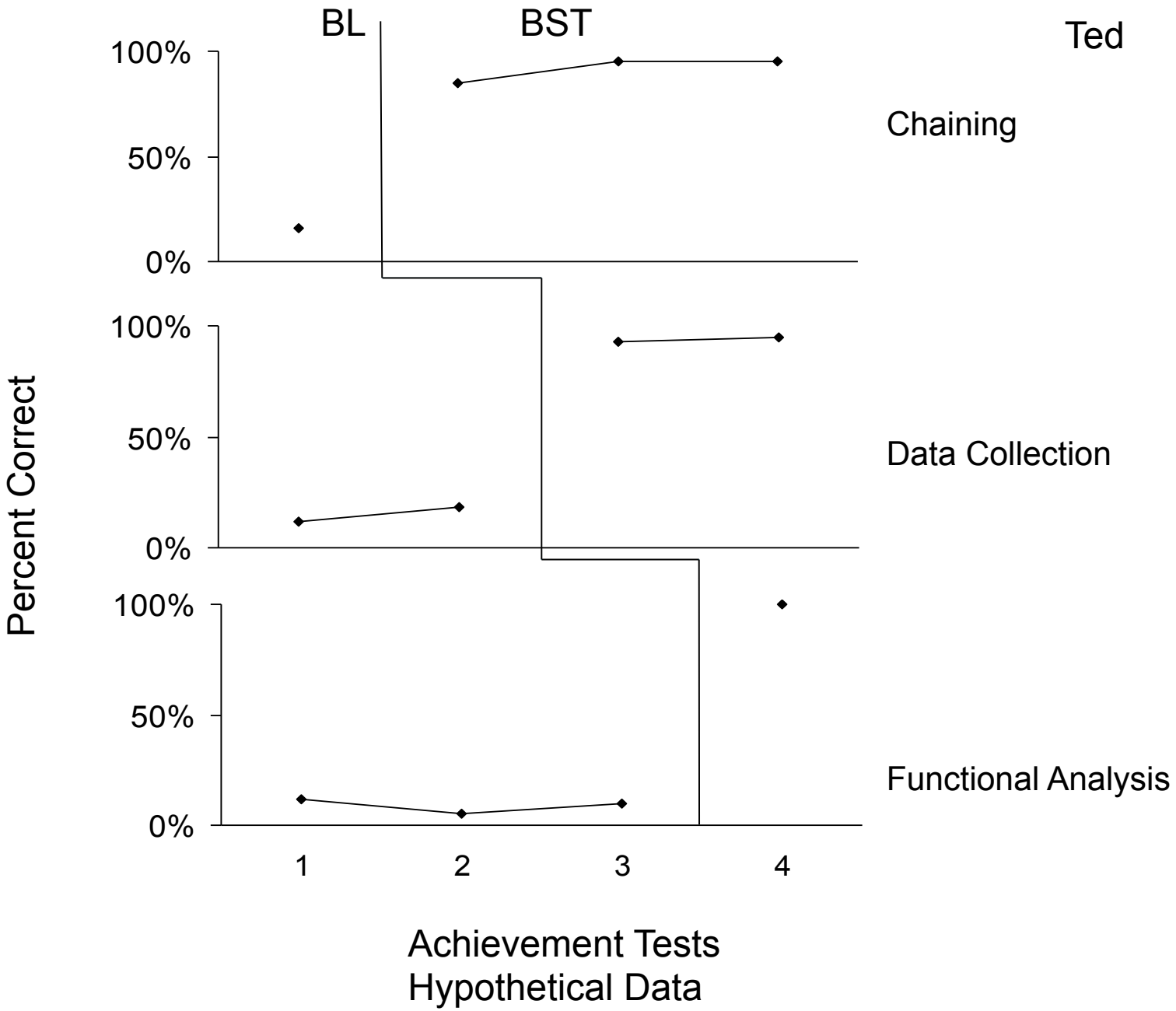


FIGURE 11.1 ***Example of a multiple-baseline-across-behaviors design.*** The dependent variable was the percentage of components correctly completed for each of three teaching skills: (1) informal assessments, (2) quality IEPs, and (3) adaptations to the general education curriculum. The intervention was an intensive on-site technical assistance intervention to improve the teaching practices of a special educator, Sally.

Source: From N. M. Clark, L. S. Cushing, and C. H. Kennedy, "An Intensive On-Site Technical Assistance Model to Promote Inclusive Practices for Students with Severe Disabilities," manuscript submitted for publication.


BL
BST
Ted
100%
50%
0%
Chaining
Data Collection
Functional Analysis
Percent Correct
1
2
3
4
Achievement Tests
Hypothetical Data

Multiple Baseline Across Settings

- A single behavior is targeted in two or more different settings or conditions
- After steady state baseline responding, independent variable is applied to 1st setting, while other settings are kept in baseline conditions
- When steady state responding is reached for 1st setting, then IV is applied to next setting

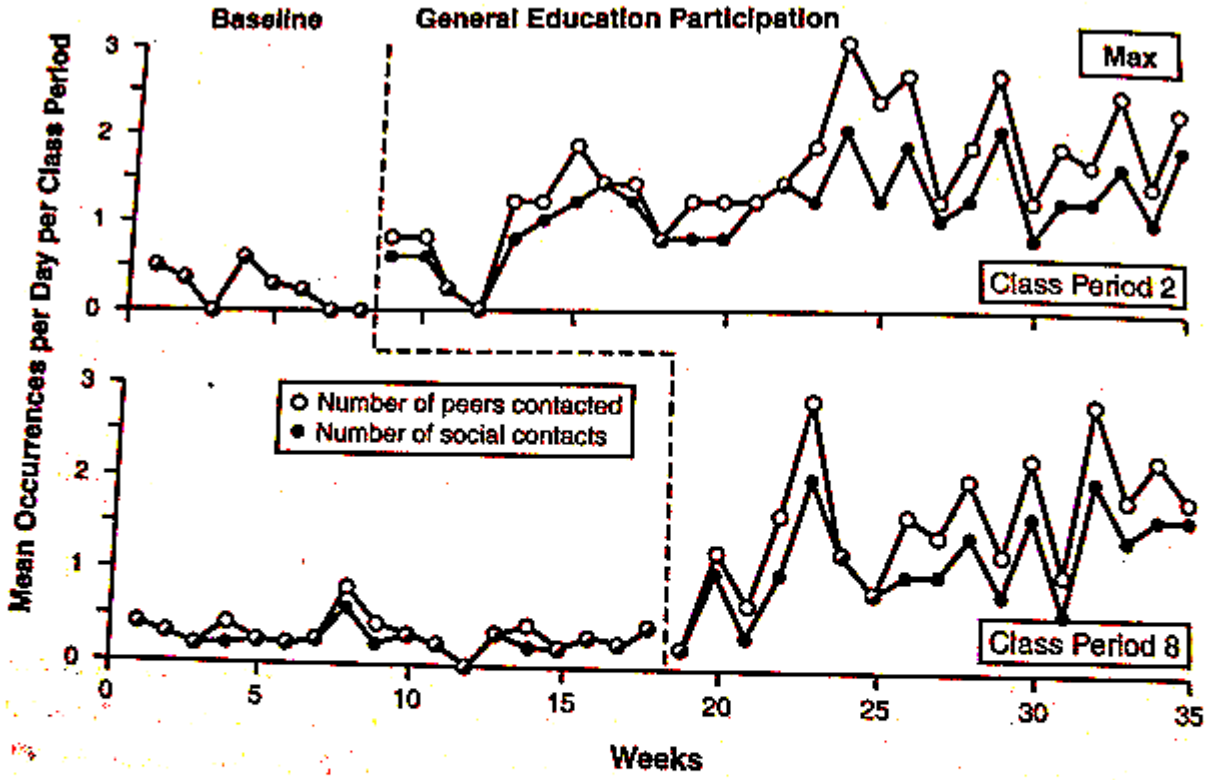
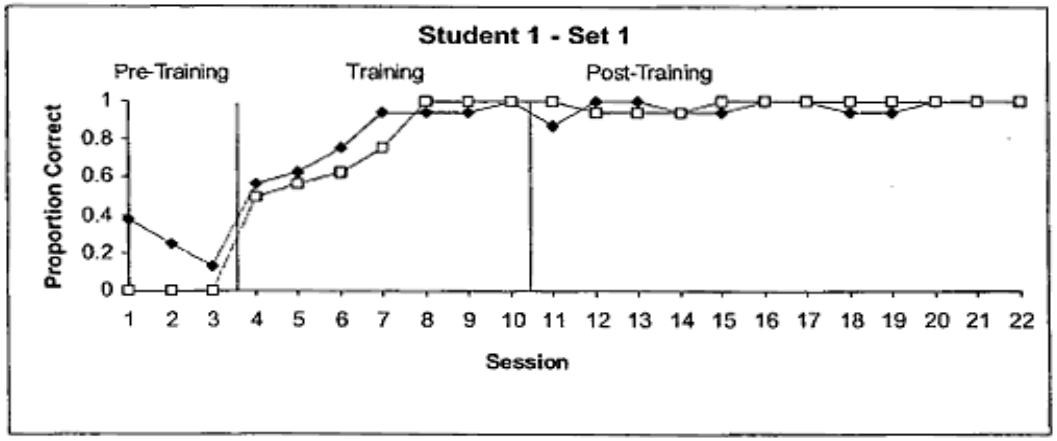
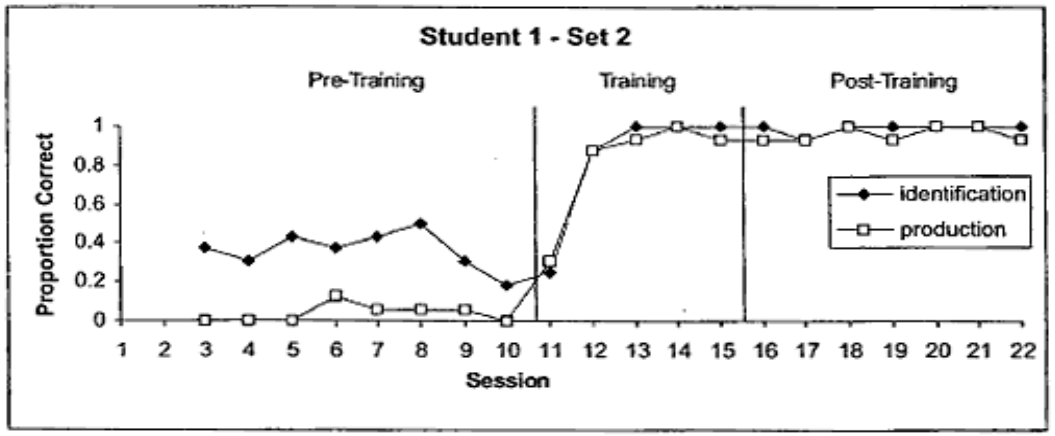
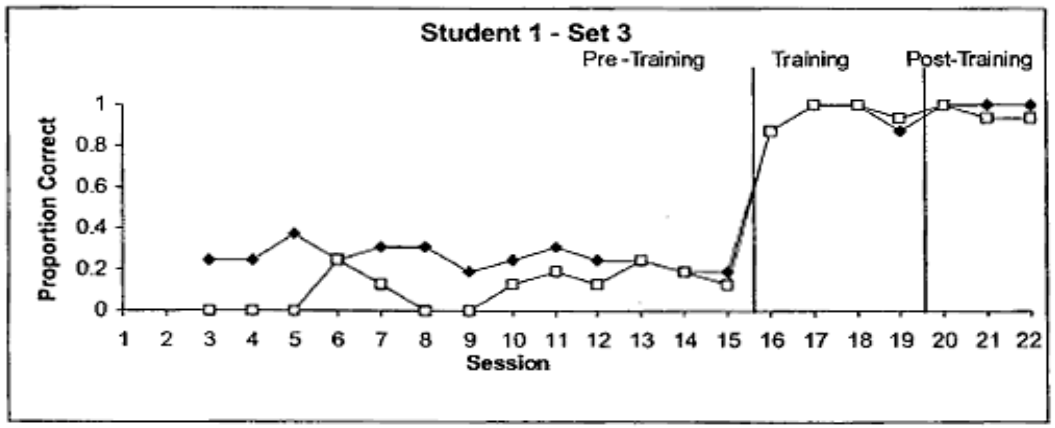


FIGURE 11.3 ***Example of multiple-baseline-across-settings design.*** The study measured effects of a peer support program on the social contacts between a student with severe disabilities (Max) and his peers without disabilities in two classes. This experimental design represents the minimal number of tiers a multiple baseline can have and still show a replicated effect.

Source: From C. H. Kennedy, L. Cushing, and T. Itkonen, "General Education Participation Increases the Social Contacts and Friendship Networks of Students with Severe Disabilities." *Journal of Behavioral Education*, 1997, 7. Copyright 1997 by Human Sciences Press. Reproduced with permission.



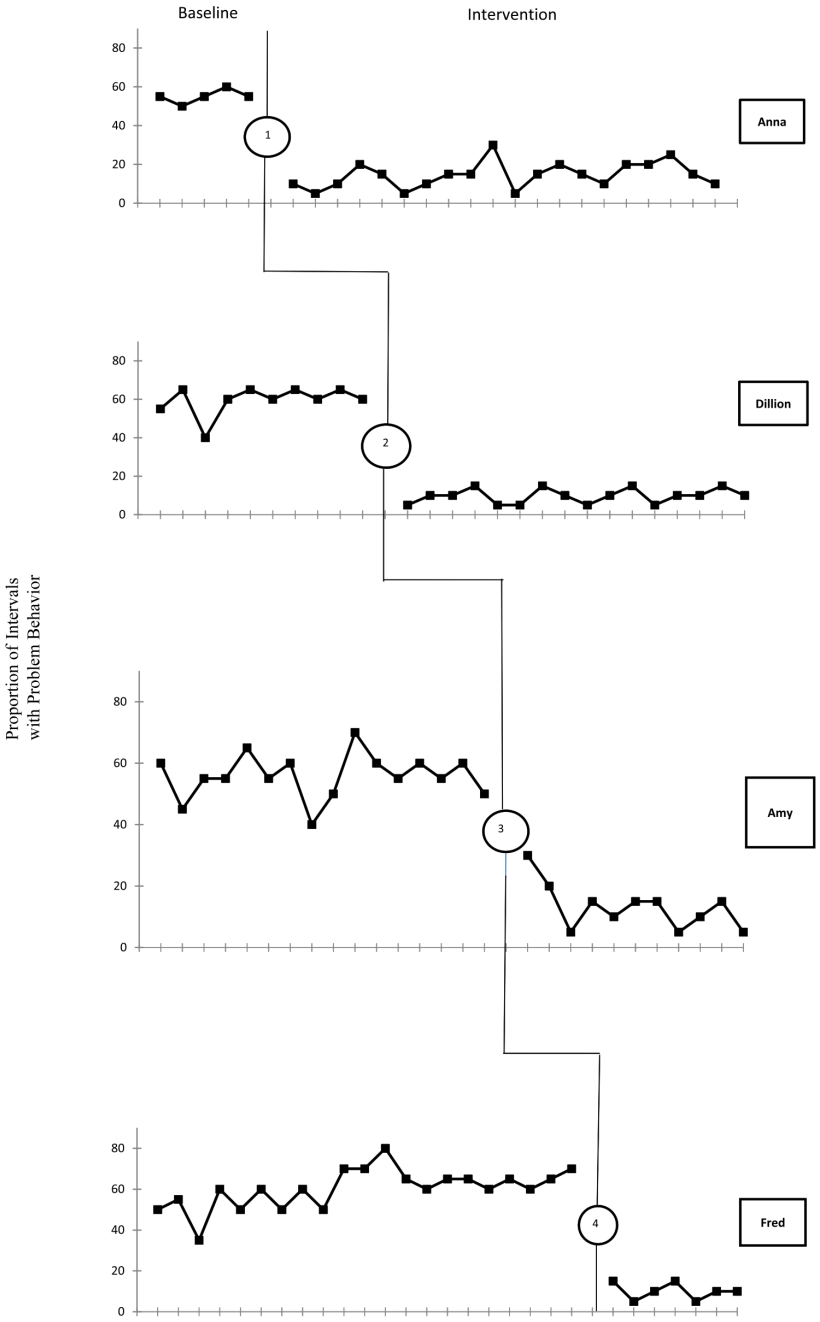




http://
www.animatedspeech.com

Multiple Baseline Across Subjects

- One target behavior for 2 or more subjects in the same setting
- After steady state baseline responding, independent variable is applied to 1st subject, while other subjects are kept in baseline conditions
- When steady state responding is reached for 1st subject, then IV is applied to next subject
- Most widely used multiple baseline design

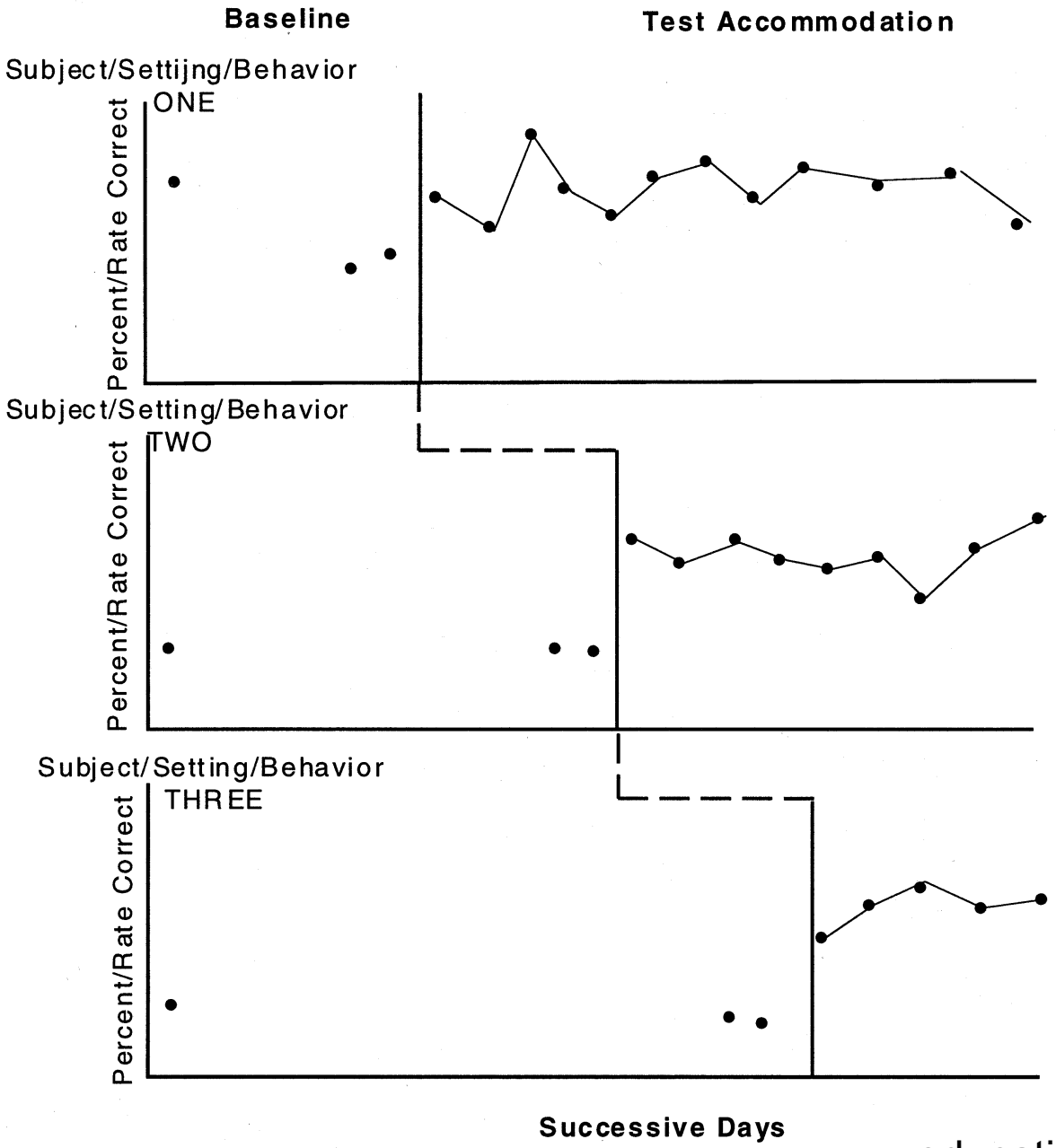

Baseline
Intervention
Proportion of Intervals with Problem Behavior
80
60
40
20
0
1
Anna
Dillion
2
Amy
3
Fred
4

Multiple Baseline Across Subjects

- Not a true single-subject design because verifications and predictions for each subject have to be inferred from replications with other subjects
- Weakness and advantage (Johnston & Pennypacker, 1993a)
  - Weak for single subject internal validity
  - Advantage for across subject external validity

Multiple Probe Design

- Same logic as multiple baseline design
- Intermittent measures or probes are made during baseline and after treatment
- An alternative when concurrent data collection is not feasible
- Requires a reasonable assumption of stability in the data
- Often used with step like program like chaining



education.umn.edu/

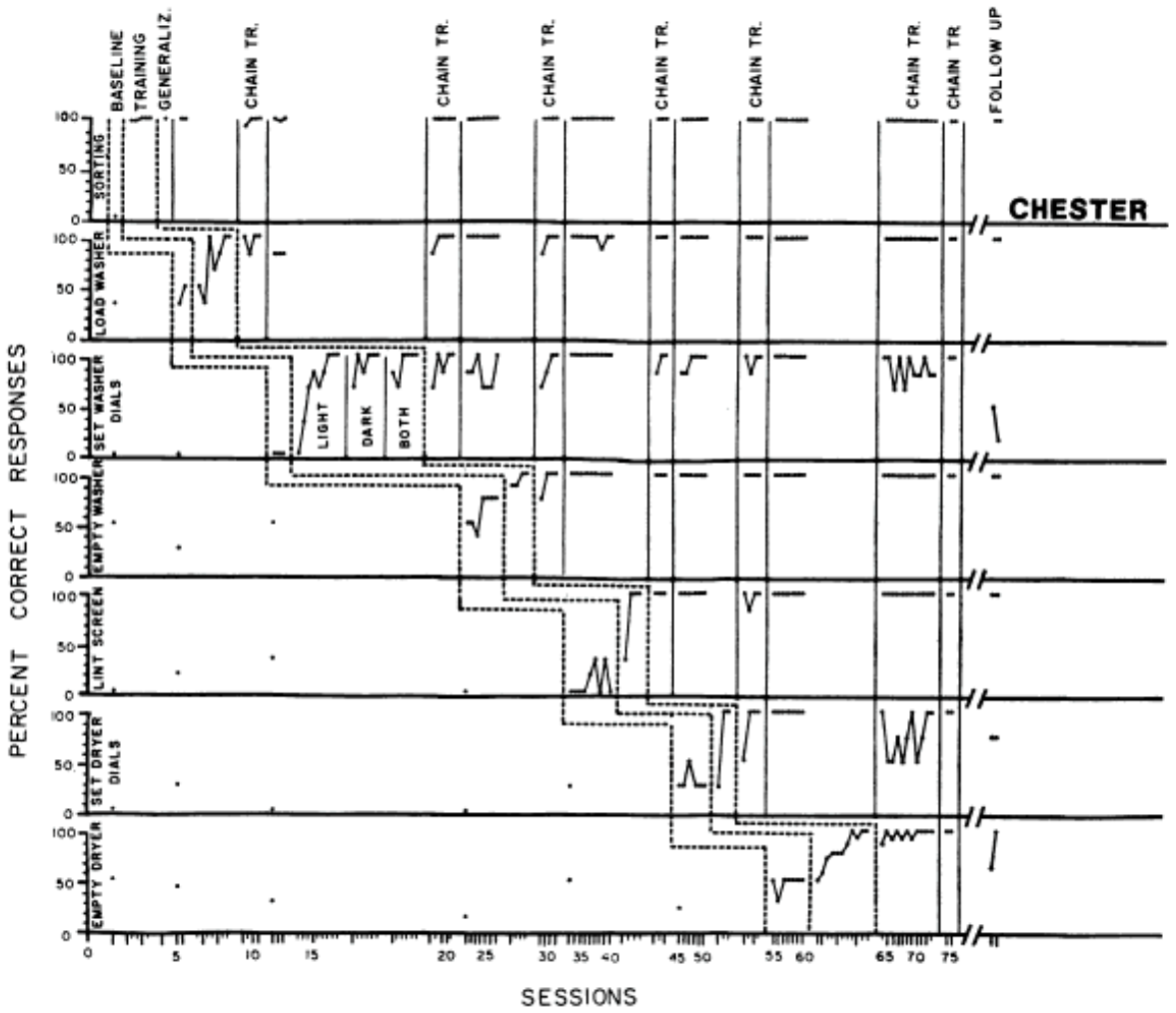


Fig. 1. Percent correct responses for each trial on each component of the laundry chain. Heavy vertical lines on the horizontal axis represent successive training sessions. Lighter vertical lines along the horizontal axis indicate trials within a session.

Delayed Multiple-Baseline

- Same logic as multiple baseline
- Second, third, and fourth baselines are introduced later in time (see pg. 212)
- Good for when new behaviors, settings, or subjects become available after the initial baseline has started.
- Good when practical problems prohibit measuring all from the beginning of the study.

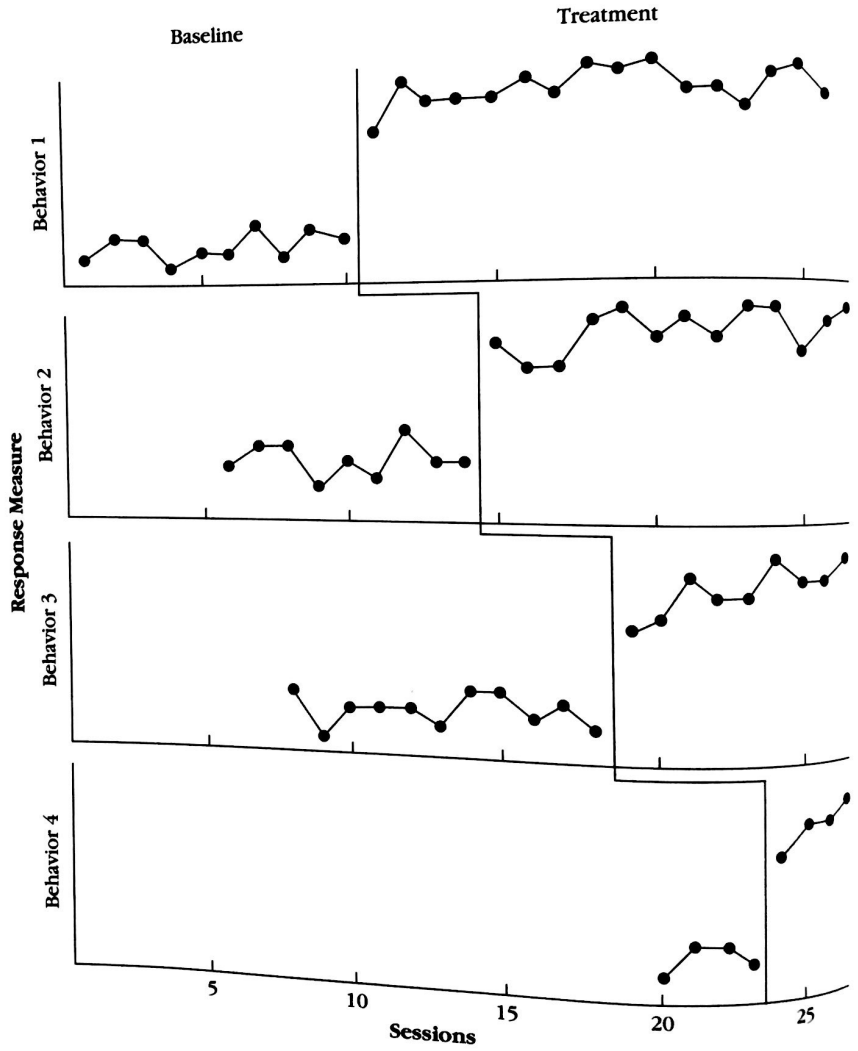


Figure 9.8 Graphic prototype of a delayed multiple baseline design.

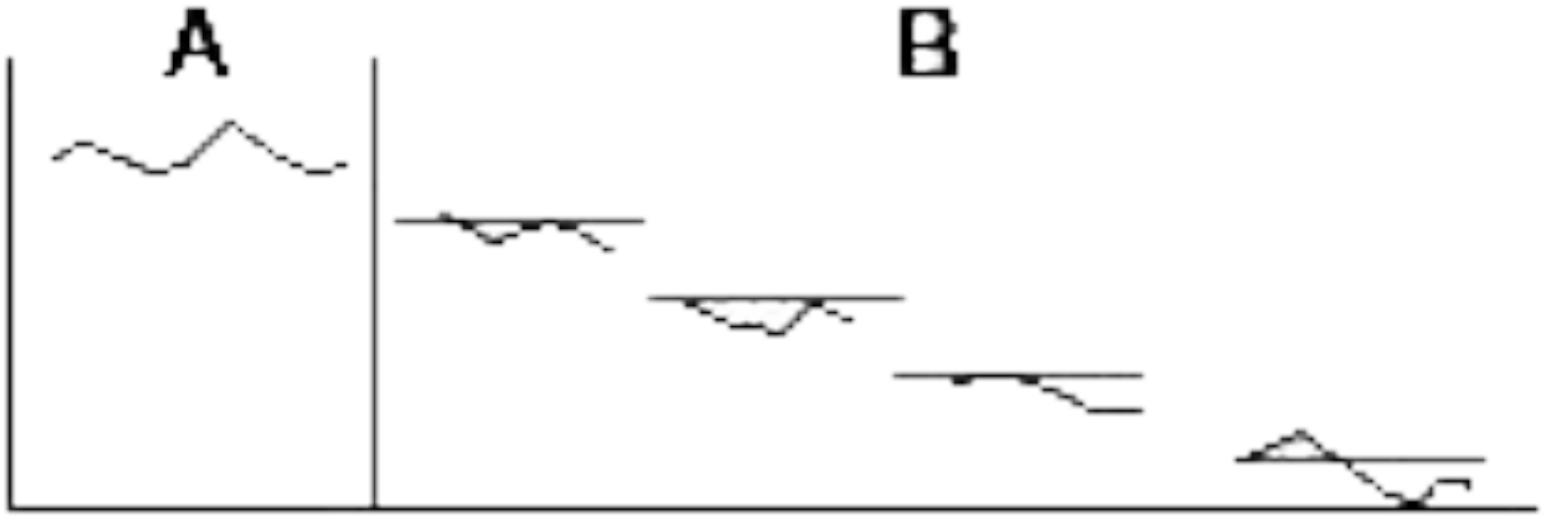
Teams

- Provide an example of an experimental question where you would use a multiple baseline design across settings.
- Provide another example of an experimental question where you would use a multiple baseline across behaviors.

Changing Criterion Designs

- Involves an initial baseline
- Set a criterion for behavior change
  - E.g., Ten requests for the bathroom per hour
- Once criterion met, set a new criterion
  - E.g., Six requests for the bathroom per hour
- Once criterion met continue to set new
- Good for treatments that can become more stringent over time

Changing Criterion Designs



www.baam.emich.edu

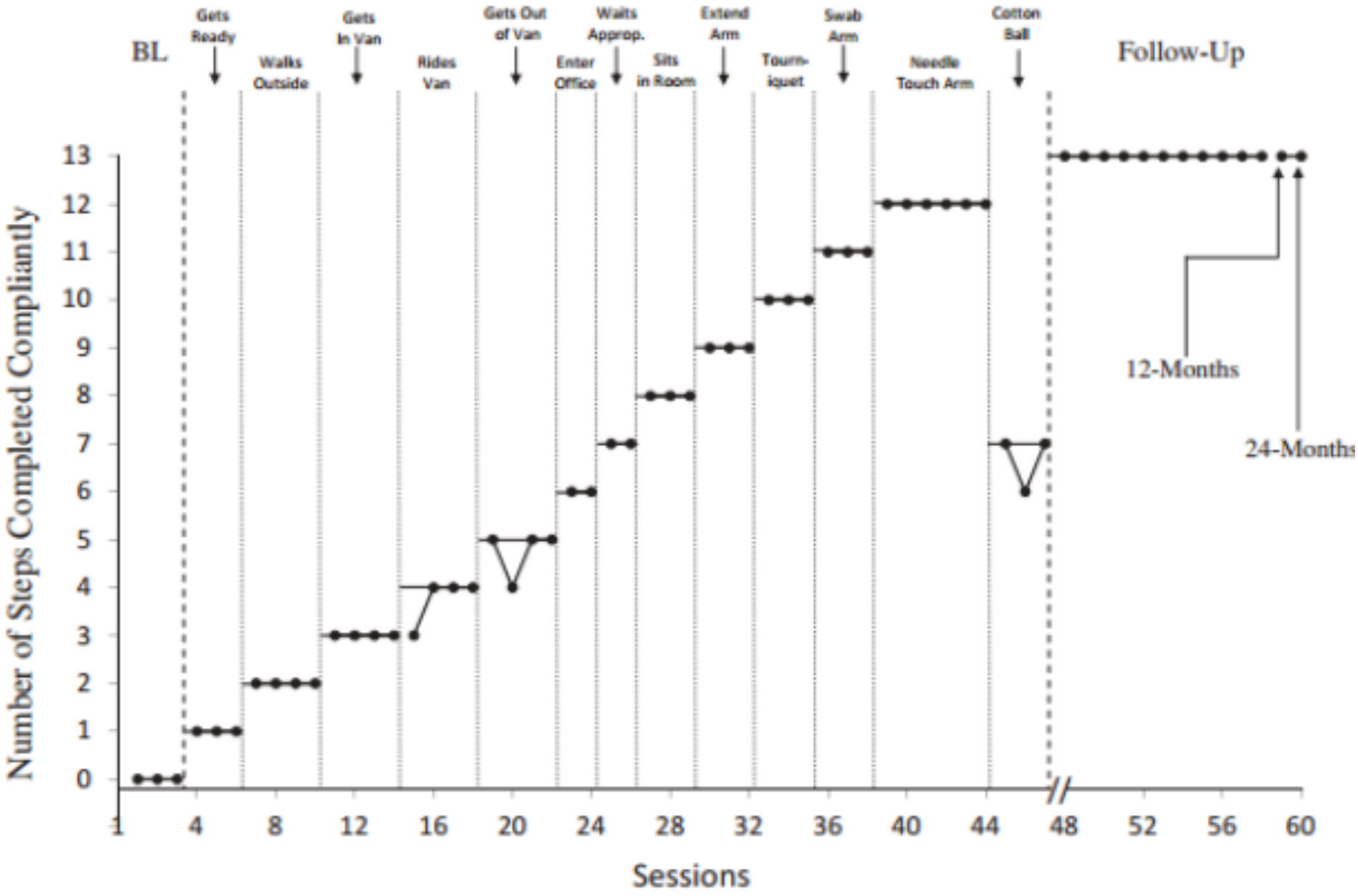


FIGURE 1 Number of steps in the task analysis Amber completed compliantly.

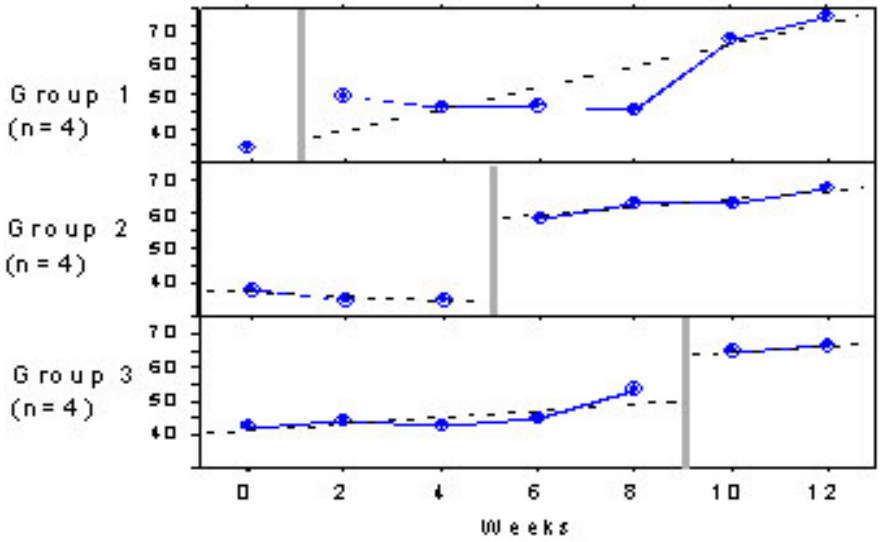
Teams

- Provide an original example of an experimental situation where the changing criterion design would be appropriate.

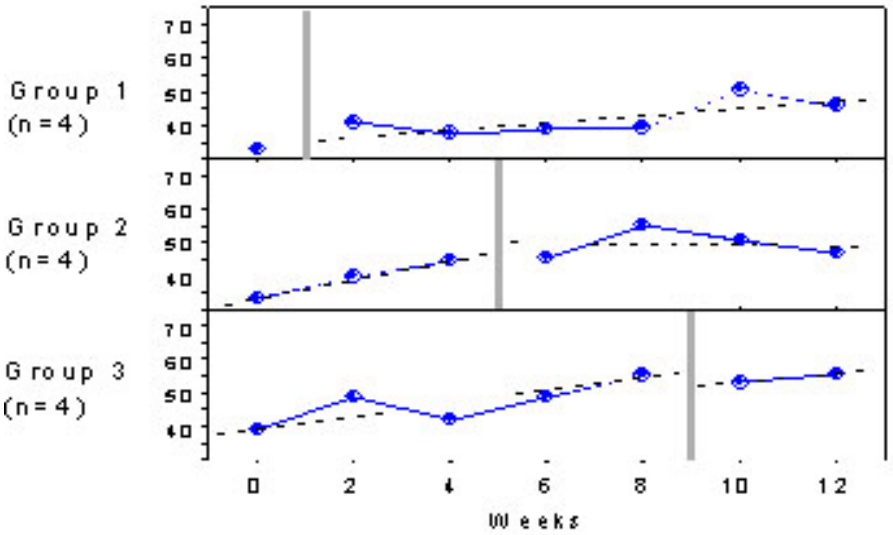
A few last points about research designs…

- Select the elements of design that best match your question, the subject, ethics, and the logistics of the situation.
- Many good research designs combine elements of reversals, multiple-baselines, and changing criterion designs.

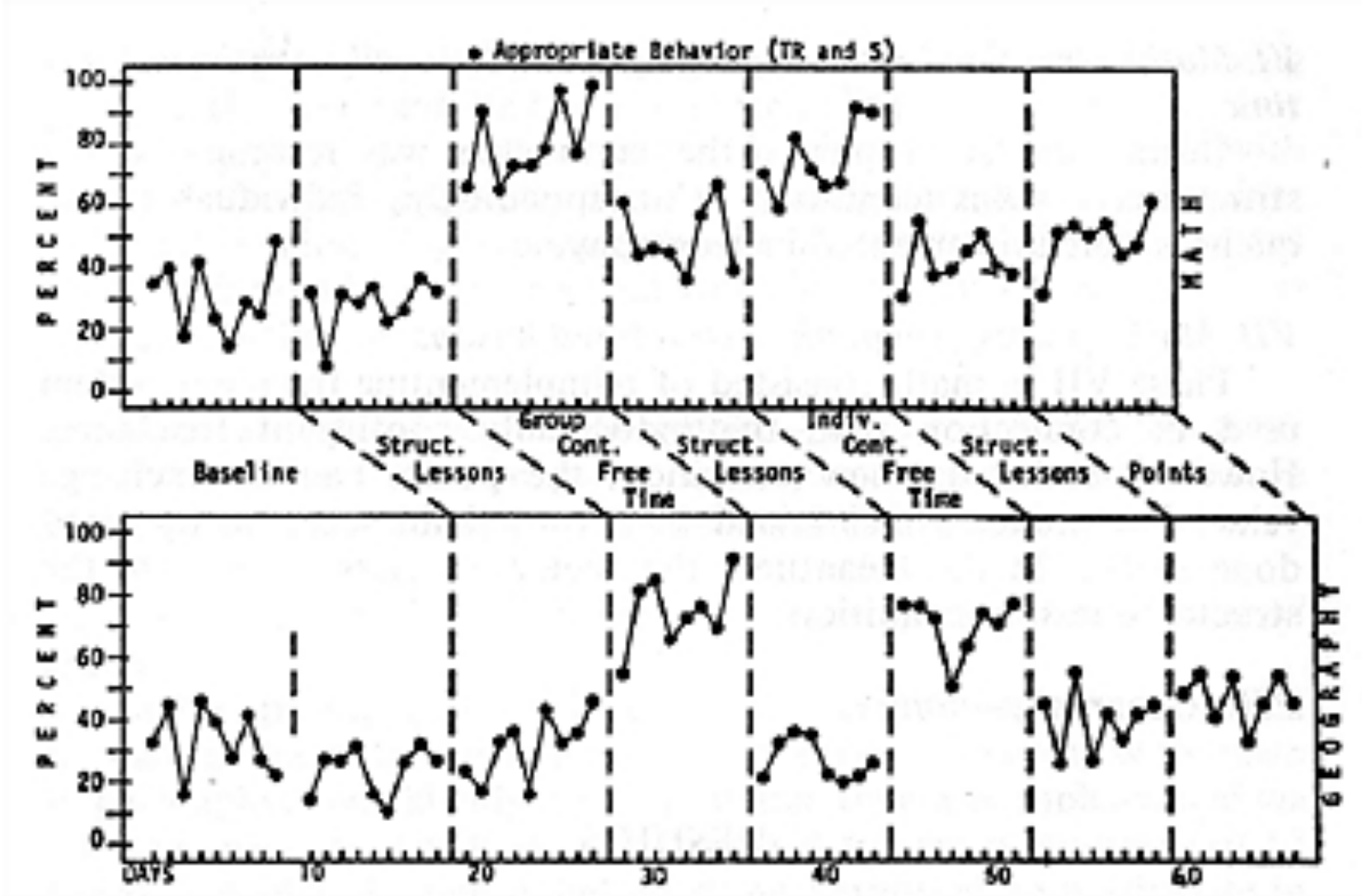
Improvement Trends for Oral Reading Fluency related to 'Read Naturally' Treatment with Grade 1 Bilingual Students - High Engaged Students (N=12)



Improvement Trends for Oral Reading Fluency related to 'Read Naturally' Treatment with Grade 1 Bilingual Students - Low Engaged Students (N=12)



Reversal within a Multiple-baseline

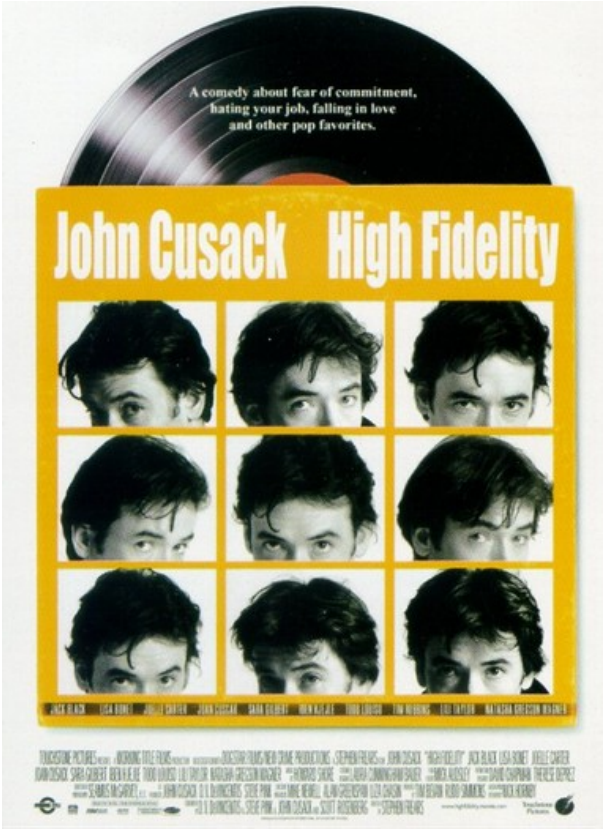


- Treatment integrity (procedural fidelity): Measure how well the treatment was carried out.

- Treatment integrity refers to the extent to which the treatment was carried out as defined by the investigator.
- If the treatment is not carried out as planned and described by the experimenter two problems of validity could arise.

- If the treatment is not carried out as planned, the result may not be due to the treatment as defined.
- Alternatively, the treatment could fail to produce a result, but if done properly it would have.

- How do you minimize problems of treatment integrity or fidelity?



- Tightly defined, operational definitions of the treatment
- KISS:
  - Keep
  - It
  - Simple
  - but **S**ophisticated.
- Training of staff
- Management of Staff
- Collect data on agreement between definition of treatment and practice

Data on Integrity

- List steps of treatment as in bullet statements
- Create a check list or rubric
  - Check list-simple yes or no decision
  - Rubric-more complex rating or weighting
- Use checklist/rubric to record whether carried out
- Report as percent accuracy or integrity

Graphing Multiple Baseline Data

- Graph the data from the next slide.
- Make up names for behavior 1, 2, and 3.
- Make up Treatment
- Make sure you include:
  - labels for the x and y axis, labels for the conditions, the condition lines, data points, data paths, and a figure caption.

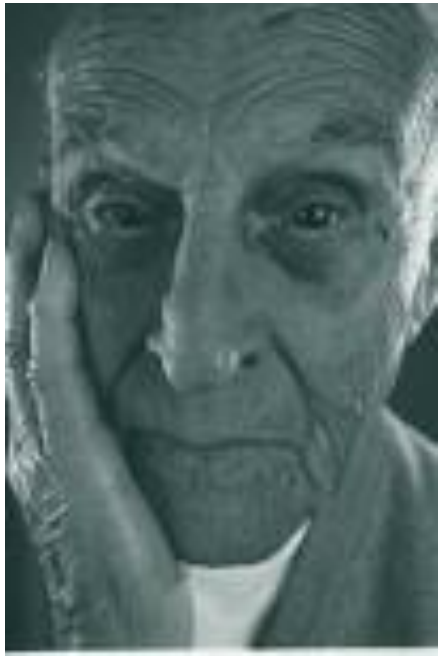
	Behavior 1		Behavior 2		Behavior 3	
Weeks	Baseline 1	Treatment 1	Baseline 2	Treatment 2	Baseline 3	Treatment 3
1	10		0		33	
2	0		0		100	
3	0		0		0	
4		10	0		33	
5		33	10		0	
6		100	10		0	
7		33		100	0	
8		100		100	0	
9		100		100	0	
10		100		100		100
11		100		100		100
12		100		100		100

Results

- Write a brief results section for the data you graphed above. Describe the results in terms of:
- level
- trend
- variability
- comparisons between conditions

Final Considerations

- State the fundmental reason that behavior analysts use single-subject or intrasubject research designs.



- State how group designs and single subject designs vary in terms of replication.

- List the kinds of confounding variables (threats) you should examine in all applied behavior analysis studies.

Sample Multiple-Choice Test Questions

- Take out a piece of paper
- Write down the letter that is the best answer for each question that I present to you.
- We'll go through these quickly as practice
- I'll present what I think is the best answer and then we can discuss.

1. Which of the following is an example of an overt behavior?
   A. Expecting something to happen based on a dream
   B. Feeling good about a dream
   C. Reporting a dream
   D. Dreaming of something

2. Which of the following is the best example of a defined behavior?

A. Monica is a manic depressive with severe psychosocial tendencies

B. Monica eats too much food between meals and before bed.

C. Monica daydreams for 10 out of every 30 15 second interval when reading his math text

D. Monica looks at the window for 50 percent of her reading time

8. Nina bites her mother. Which of the following is the best way to measure her behavior?

A. Duration of each biting episode
B. Frequency with which she bites her mother
C. Frequency with which her mother says, “No biting.”
D. Number of teeth marks on mother’s arm.

10. Francesca was asked to cut a tomato into eighths. Which of the following would be a convenient and reliable measure of her behavior?

A. Permanent products
B. Duration
C. Frequency
D. Latency

11. Karen's mother complains that Karen laugh is too long. Which of the following would be the most convenient and reliable measure of laughing for Karen?

A. Frequency
B. Permanent product
C. Latency
D. Duration

14. **Circle all** of the following that involve continuous recording.

A. Nina records the number of self-injurious behaviors that Karen engages in during playtime at the preschool.

B. Monica records the length of time that Luca is seated at his seat.

C. Paolo records whether self-injurious behavior is occurring when his beeper goes off every twenty seconds.

D. Luca records whether the bass guitar can be heard distinctly on the digital recording.

15. Marco wishes to measure 5 different behaviors every 30 seconds. Which recording procedure is best suited for such a task?

A. Whole interval
B. Partial interval
C. Momentary time sampling
D. Frequency

19. Nina has too many things to do to use a chart and pencil in order to keep track of the number of times Luca uses bad words in the clinic. Which of the following is likely to help the most?

A. A stopwatch to record the duration of his use of bad words

B. Marking on a checklist of bad words that Luca has used before whenever she hears a bad word

C. Using whole interval recording for twenty minutes each day during lunch to estimate the number of bad words

D. Transferring coins from one pocket to another whenever she hears a bad word and counting them at the end of the session.

- Filippo has eaten a lot of bread with his meals for years. Recently he has developed problems with his weight and he wishes to reduce the amount of bread he consumes each day. What is the best design for assessing whether gradually changing the amount of bread per day that he targets will be effective in reducing bread eating?

A. Reversal

B. Multiple-baseline across settings

C. Alternating treatments

D. Changing criterion

- Which of the following is the **best** reason to use a multiple-baseline design over a reversal design?

A. The client's family think that experiments with their child are unethical.

B. You read that a multiple-baseline design provides better experimental control than a reversal design.

C. Your supervisor argues that the behavior you are measuring is not likely to be reversible.

D. The client's family thinks a multiple-baseline will fit better into their day-to-day life.

- Which of the following is the best example of a **measurement** confound for a study concerning a new curriculum?

A. The students who received the new curriculum were assigned to the study by their teachers and those in traditional class were not.

B. The students who received the new curriculum received instruction by computer and those in a traditional class received lectures.

C. The students who received the new curriculum were observed by a one research assistant and those in a traditional class were observed by a different research assistant.

D. The students who received the new curriculum were in a computer lab, whereas the traditional class occurred in a regular classroom.

5. Which of the following involves the introduction of a condition in staggered fashion over time so that the condition can be compared to baseline conditions?

A. Multiple-baseline

B. Alternating treatment

C. Changing criterion

D. Withdrawal

- Francesca wishes to investigate whether a good nutritional breakfast will affect Marco's behavior. She is interested in examining this for performance in math class, on the football field, and violin lessons.

  Which of the following designs would be best for this situation?

  A. Delayed multiple baseline
  B. Multiple-baseline across behaviors
  C. Multiple-baseline across settings
  D. Multiple-baseline across students

Luca wishes to examine whether a procedure for reinforcing incompatible behaviors will reduce the severity of self-injurious behavior on the part of a 16 year old client. He already has good records of the frequency and intensity of these self-injurious behaviors as a baseline. What is the best design he should use based on ethical, practical, and experimental considerations?

A. An ABAB Reversal

B. Reversal with brief baseline probes

C. Alternating treatments

D. Multiple-baseline across settings

You will Also see multiple choice questions that ask:
Select All that Apply

- You will also have few short answer questions, like the following.
- Try to answer them quickly today.

5. Provide a sound behavioral definition for the following:
joking

6. Say whether your example is a behavior or a permanent product.

7. Say whether your behavioral definition is a topographical definition or a functional definition and justify your answer.

That’s It! Your Questions!
Submit your group answers by 19:00