

Single-Case Experimental Designs

Ed Daly

Educational Psychology Department
University of Nebraska-Lincoln

Science's replication crisis

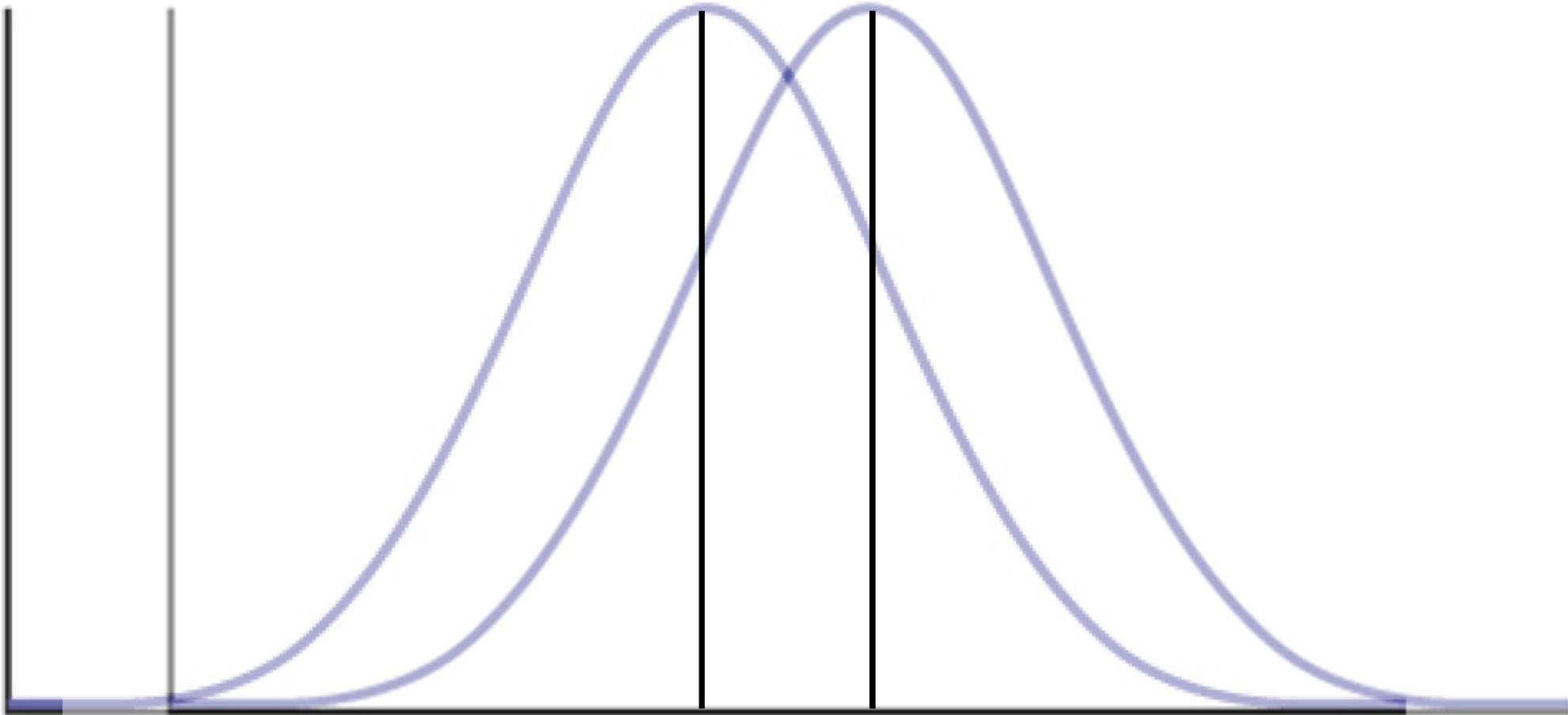
- Bayer Pharmaceutical study begun in 2011
 - Analysis of 67 in-house research projects on target identification for new drugs in oncology, gynecology, and cardiovascular disease
 - Reproduced only 21% of the results entirely
- Biotechnology company Amgen
 - Similar analysis of 53 studies launching Amgen's preclinical research projects on cancer
 - Only 11% of their findings were reproducible
(Chevassus-au-Louis, 2019)
- Open Science Collaboration (2012) Estimating the reproducibility of psychological science: replicated 36% of the original 100 studies

Why are things so bleak?

- Too few replications?
- Lack of common units of measurement and over-reliance on relative units of measurement rather than absolute units?
- An incomplete picture of people's behavior and controlling variables over time?
 - Might it have something to do with the way we handle variability in research subjects, and...
 - How our treatment of variability relates to the clients to whom we want to generalize in natural settings?

What is the typical unit of analysis
in applied research?

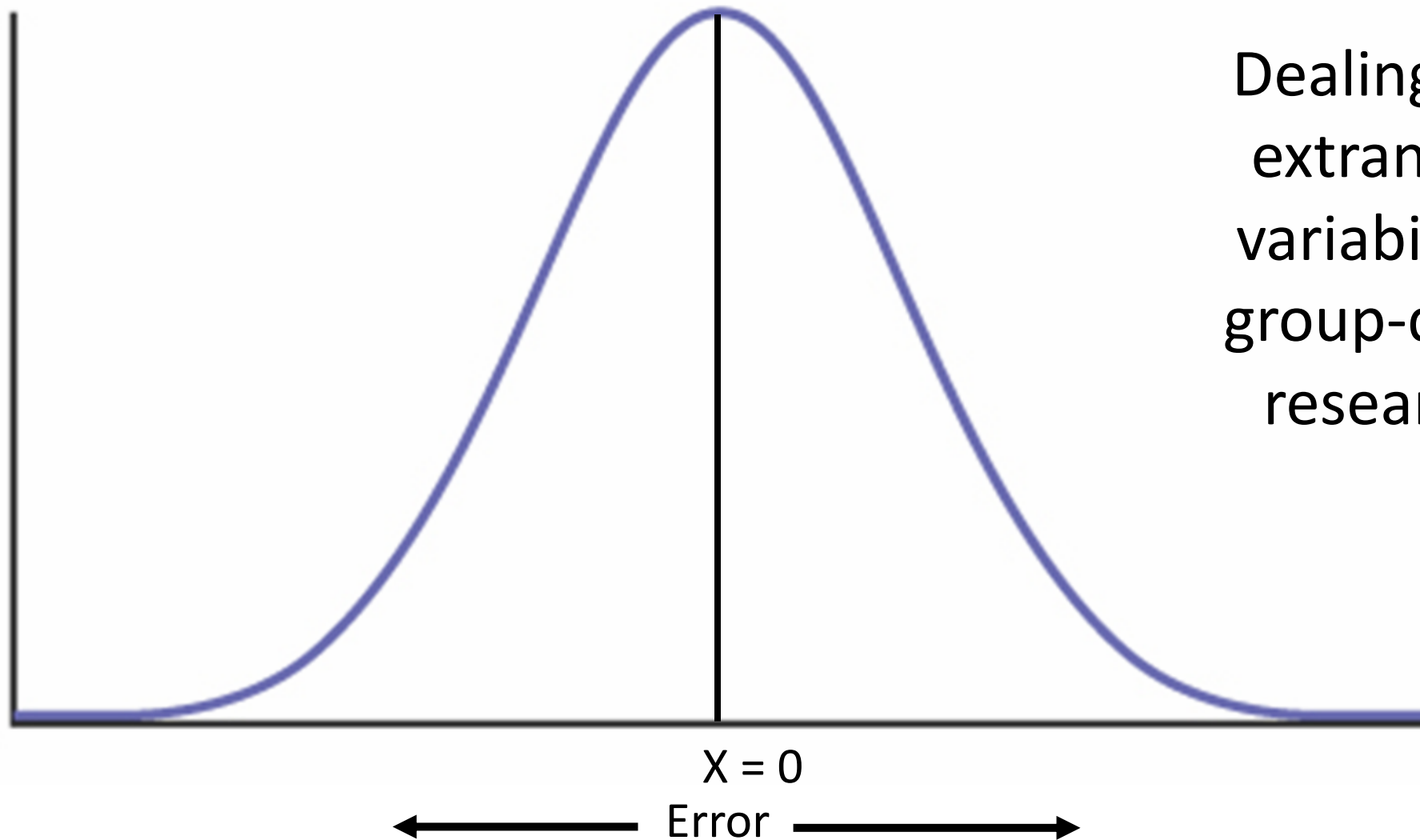
And why do we do applied research?



$X_C \longrightarrow X_T$

+ ES

Variability in the Treatment Group



Dealing with
extraneous
variability in
group-design
research...

Perhaps a more important
question:
How are they doing over time?

Has anyone figured out how to?...

- Directly control environmental events...
- While making strategic comparisons to rule out threats to interpretation,
- While repeatedly measuring responding over time,
- While studying a small number of individuals,
- While conducting observations over long periods of time,
- While still achieving successful replications with other individuals?

Is it possible to conduct a treatment study as a true experiment...

- When a group design is not possible,
- And establish a functional relationship between treatment and outcomes (specifying the mechanisms of change),
- While seeing how behavior changes over time,
- While gathering a lot of information about participants and their circumstances,
- With a methodology that is generalizable to applied settings?
 - And thus can account for individual differences?
 - And also allow me to make adjustments if things are not working out?

Historical Precedents for Studying the Individual

- The *Experimental Analysis of Behavior*
 - Skinner (1966): “... instead of studying a thousand rats for one hour each, or a hundred rats for ten hours each, the investigator is likely to study one rat for a thousand hours” (p. 21)
 - Murray Sidman’s (1960) publication of *Tactics of Scientific Research*
- Applied Behavior Analysis officially launches its flagship journal, the *Journal of Applied Behavior Analysis*, in 1968
- Single-case studies begin to appear in increasing numbers in other major psychological and psychiatric journals

Is anyone using them?

Professional fields in which SCED research is published

- Special Education
- School Psychology
- Clinical Psychology
- Organizational Psychology
- Developmental Disabilities
- Medicine
- Nursing
- Physical and language therapy
- Sports psychology

A sampling of journals Publishing SCED studies

- Journal of Applied Behavior Analysis
- Journal of Special Education
- Journal of School Psychology
- Behavior Research and Therapy
- Exceptional children
- American Journal of Speech, Language, Hearing
- Topics in Early Childhood Special Education
- Remedial Education
- Exceptional children
- Research in Developmental Disabilities
- Journal of Emotional and Behavior Disorders
- Child and Family Behavior Therapy, etc.

A contemporary call for more *idiographic* research

- Barlow, Nock, & Hersen (2009): “It is significant that it was the rediscovery of the study of the single-case in basic research, coupled with a new approach to problems in the applied area, that marked the beginnings of a new emphasis on the experimental study of the single-case in applied research.” (p. 26)
- Barlow & Nock (2009) call for the intensive analysis of the individual using strategies that allow one to establish ***causal relations*** among variables and therefore discover functional relations

Misconceptions about Single-Case Designs

1. Single-case designs are merely case studies
2. Single-case designs lack objectivity because of their reliance on visual analysis
3. Single-case designs are only useful in the field of Applied Behavior Analysis

SCEDs's Unique Perspective on Reliability, Generality, and Variability...

- Sidman (1960): “Generality and variability are basically antithetical concepts”
- Precision of control is the key!
 - **Intra-subject replication** establishes reliability which makes replication across individuals (**inter-subject replication**) more likely, thereby establishing the generality of the results
 - A strong model for replication *within units of analysis*
 - Pursuit of robust treatment effects by...
 - Tracking down sources of variability, and
 - Using strong treatments that produce good variability

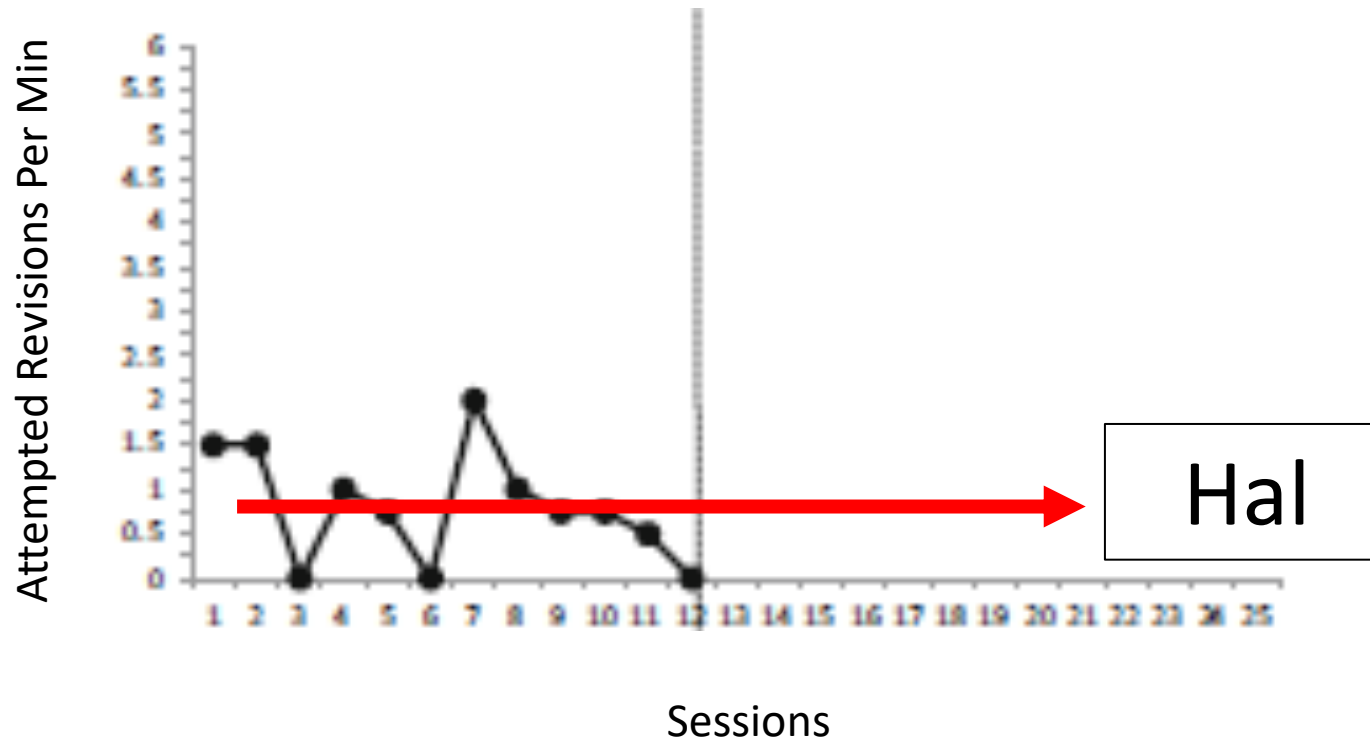
Kazdin (2011)

“In applied work, the interventions (e.g., variations of reinforcement) have had very robust effects. Investigators who use single-case designs have emphasized the need to seek interventions that produce dramatic changes in performance. Interventions that produce dramatic effects are likely to be more generalizable across individuals than are effects that meet the relatively weaker criterion of statistical significance.”

What's the best way to gain greater precision?

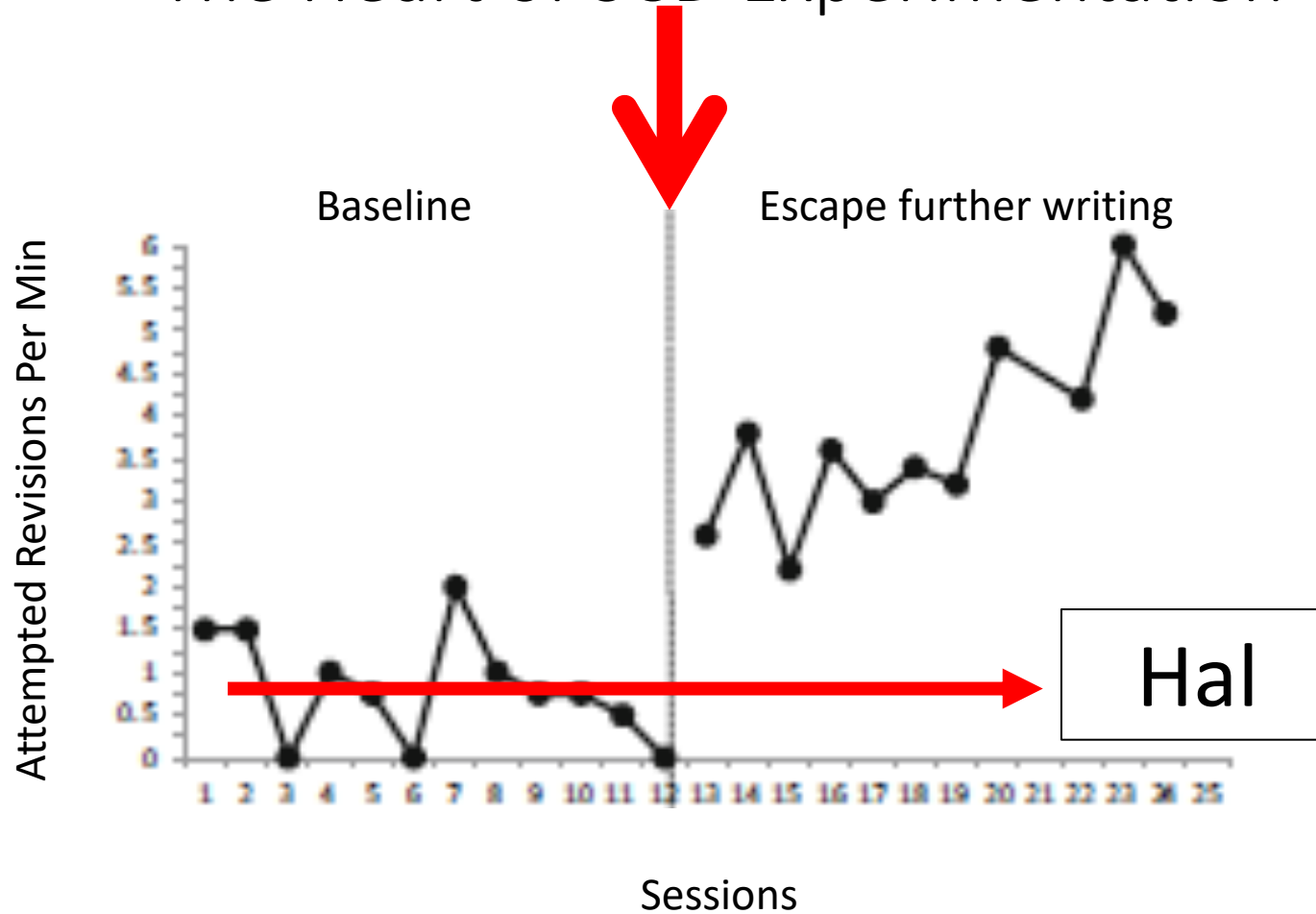
- Establish a steady state, and
- Conduct strategic comparisons
 - Directly manipulate treatment variables to control variability while ruling out threats to validity (internal validity)

Steady State and Baseline Logic



The Phase Change

The Heart of SCD Experimentation

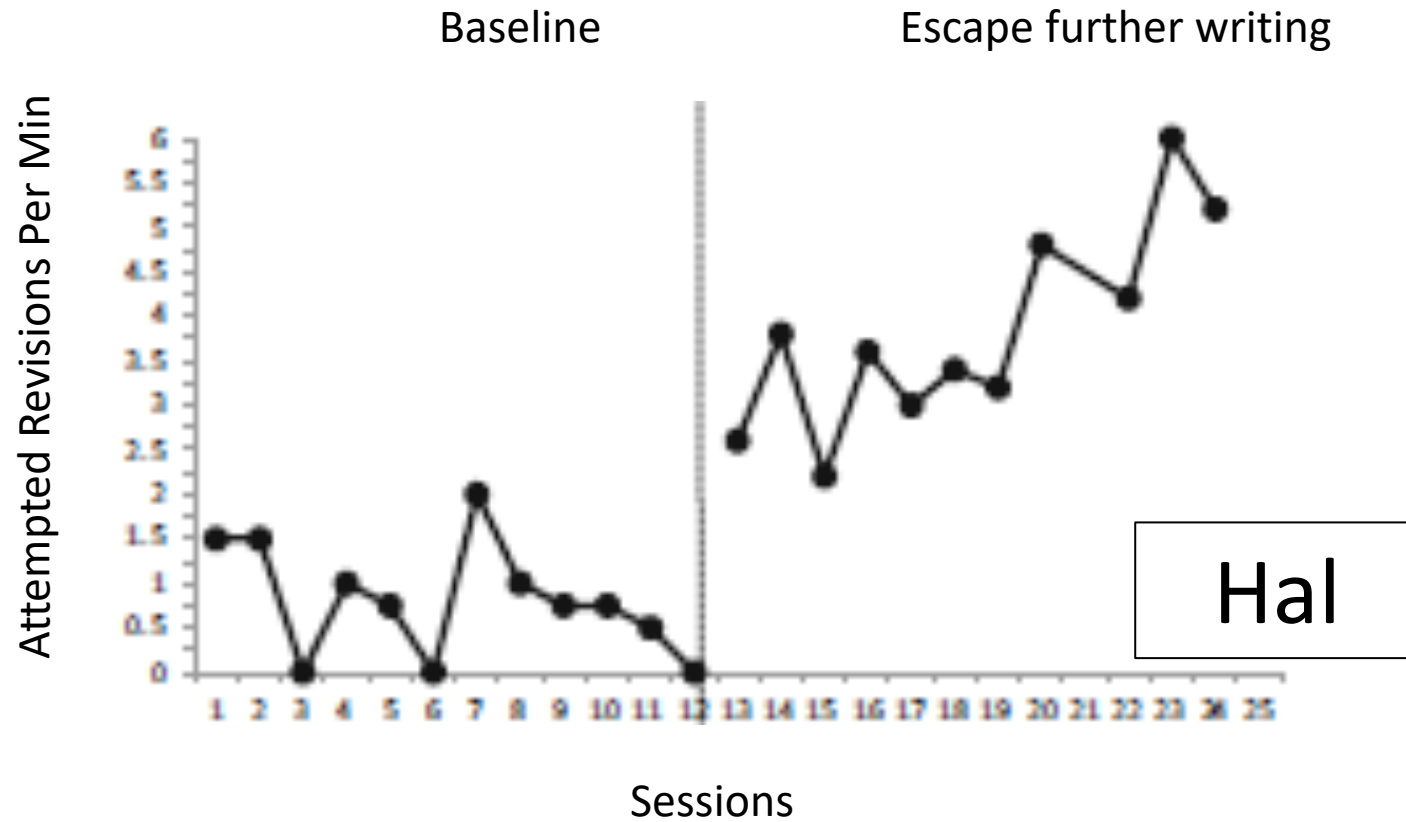


Visual Analysis

The Data Interpretation Method

1. Collect enough data points to establish stability and reliability
2. Visual analysis within conditions
 - Considering level, variability, and trend
3. Visual analysis across conditions
 - Comparing phases using level and trend

Is subjectivity a problem?



Different functions of replication...

Within-design replications



Experimental control

Replications across units of
analysis

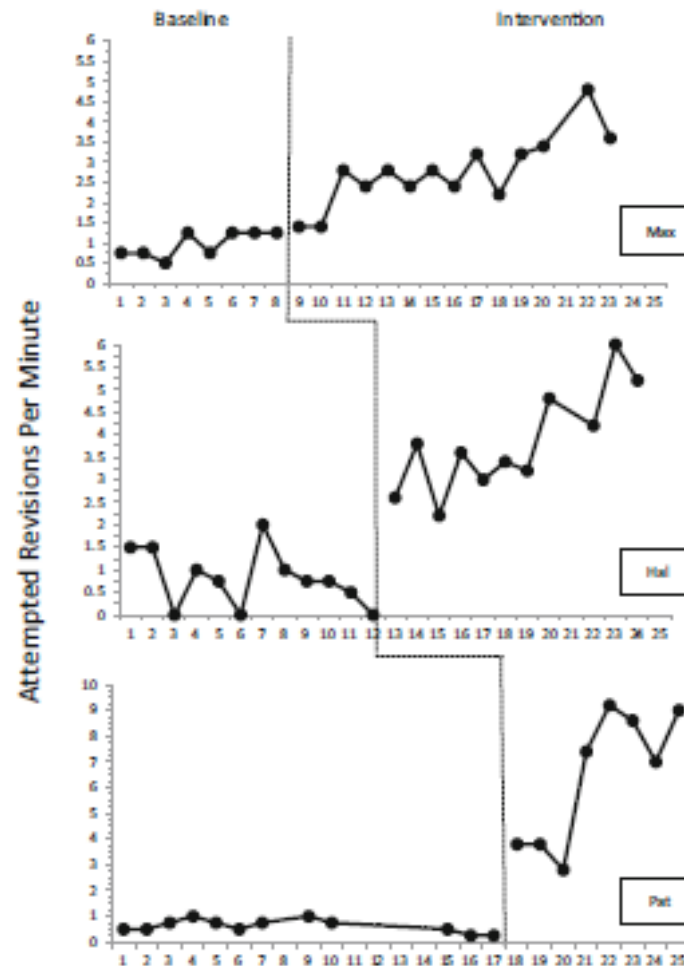


Generality

Experimental control established through replication

Repeatedly making and testing predictions about performance under each condition

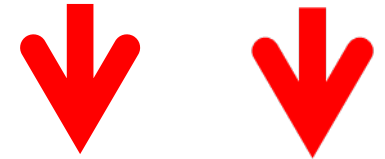
Holtz & Daly (2019)



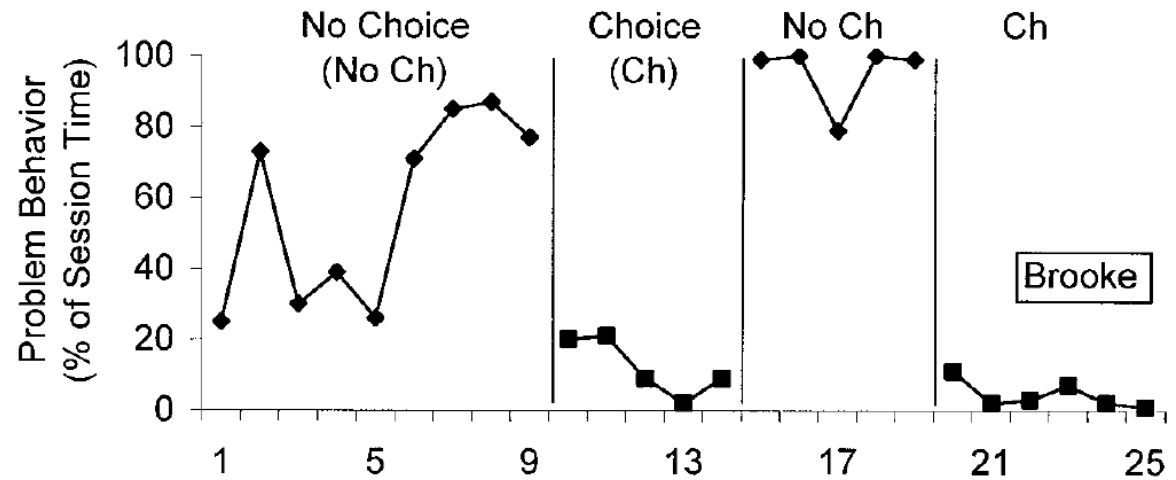
Strategic Comparisons

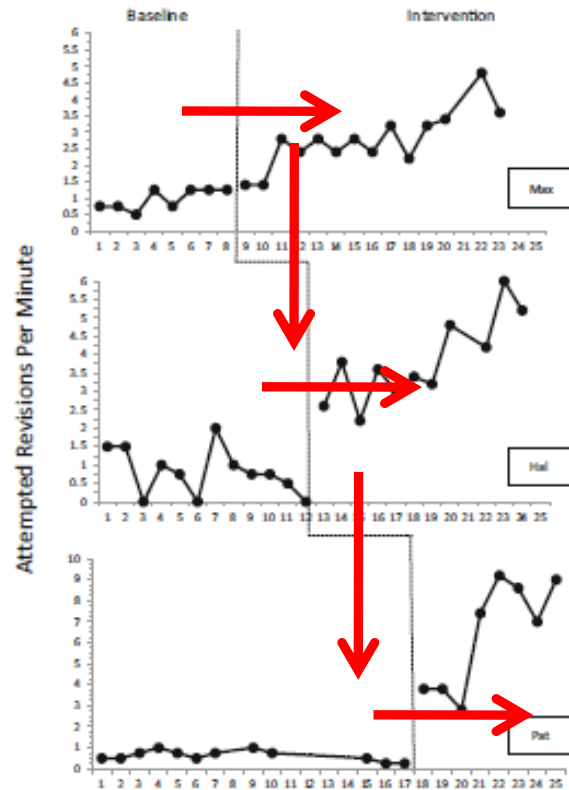
Experimental control established through replication

Rules out history,
maturation, testing effects...



Using choice of task to
improve disruptive
behavior





Decision Making with a Multiple-Baseline Design

Different functions of replication...

Within-design replications



Experimental control



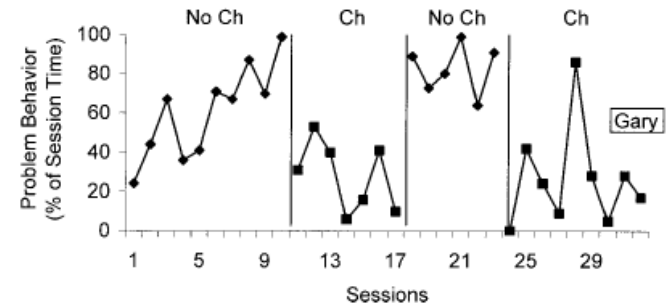
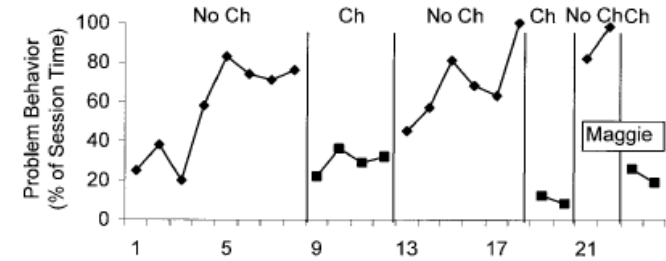
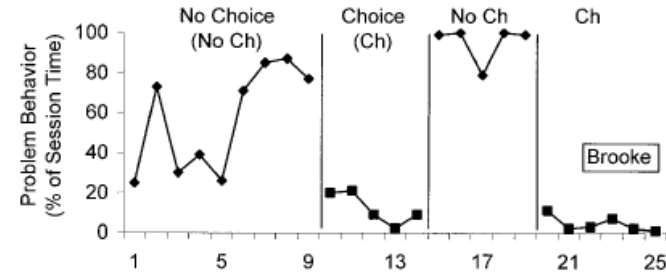
Replications across units of analysis



Generality

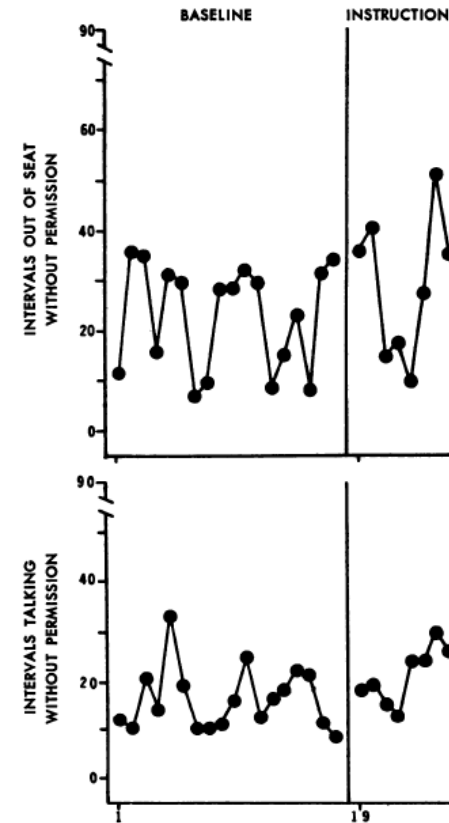
Generality established through replication

Using choice of task to improve disruptive behavior



What if things are not working out?...

- Ramp et al. (1971) first examined whether instructions would reduce out-of-seat behavior



- Ramp et al. (1971) then tried a delayed timeout procedure
 - Red light illuminated briefly when out of his seat or talking
 - Each violation led to 5-min detention later during gym or recess

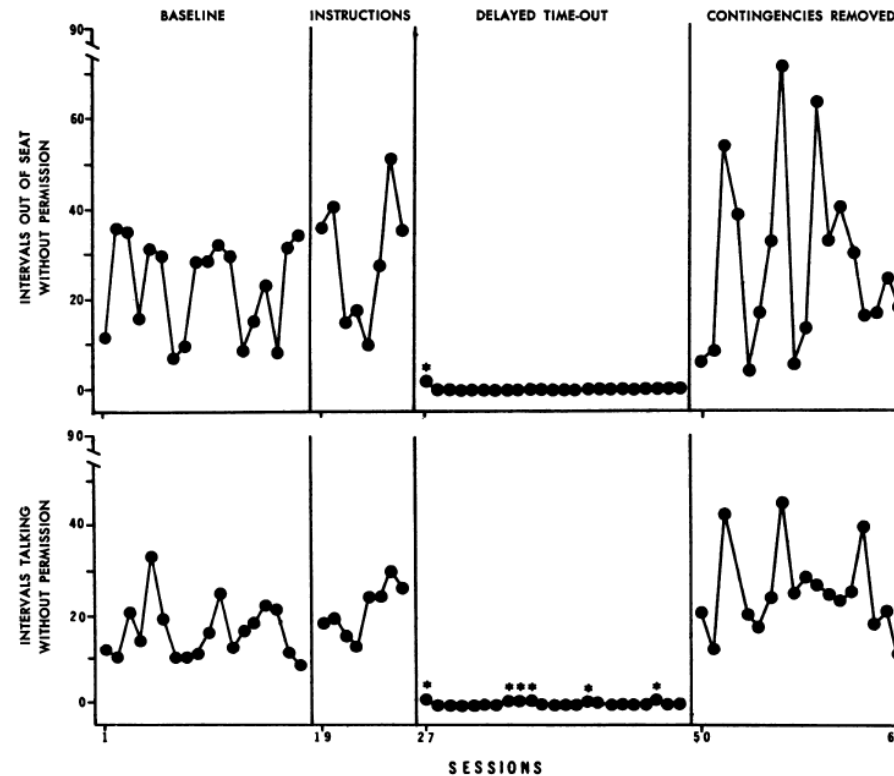
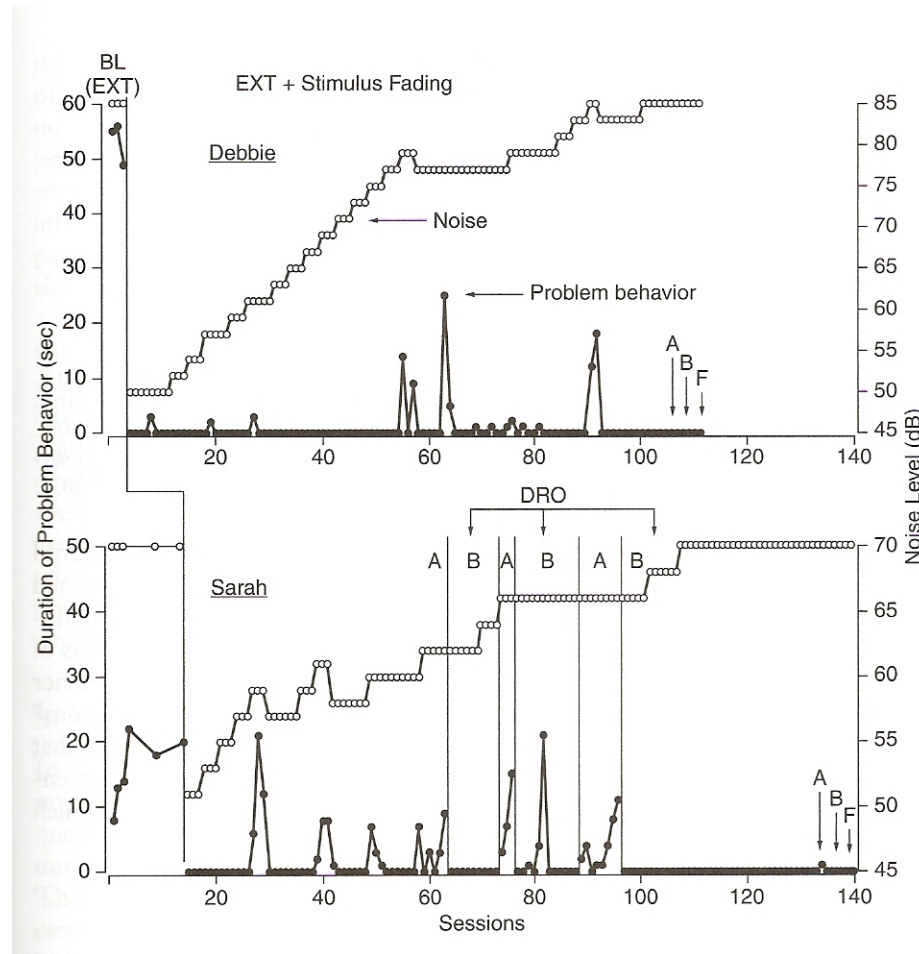


Fig. 1. Each point represents one session and indicates the number of intervals in which the subject was out of his seat (top) or talking without permission (bottom). A total of 90 such intervals were possible within a 15-min session. Asterisks over points indicate sessions that resulted in time being spent in the booth.

Response-Guided Experimentation

Try something else!

An evaluation of extinction + stimulus fading in reducing SIB duration (sec)



McCord et al. (2001)

Accounting for individual
differences with SCD...

A Word about Effect Sizes...

Significant Limitations

- None are comparable to those used for between-group approaches
- Parametric Approaches: Regression and Multilevel Models
 - Meeting statistical assumptions is extremely difficult
 - Autocorrelation causes violations of the assumption of independence made with large- n group comparison designs
 - Model misspecification can cause autocorrelation among the residuals, an autocorrelation of a type different from what statisticians mean
 - Rate and proportion data are not normally distributed, but there are currently no adequate methods for transforming the data that do not distort results
 - The effects of trends on data are not well understood
 - Methods for synthesizing results across (and not within) studies are “primitive”
 - Different studies using different outcomes

Nonparametric Methods with Promise...

- PAND (Percentage of All Non-Overlapping Data)
(Parker, Vannest, & Davis, 2011)
- Robust Improvement Rate Difference (R-IRD)
(Parker, Vannest, & Brown, 2009)
- Baseline Corrected Tau
 - Calculator online: <http://ktarlow.com/stats/tau/>
(Tarlow, 2017)
- Fisher, Kelley, & Lomas (2003) “Structured Criteria for Visual Inspection”
- For other online tools: **[HTTP://WWW.SINGLECASERESEARCH.ORG/](http://WWW.SINGLECASERESEARCH.ORG/)**

Resources:

- The Single Case Reporting Guideline in Reporting Behavioural Interventions (the SCRIBE Statement)
(Tate et al., 2016)
- IES National Center for Special Education Research
<https://ies.ed.gov/funding/researchDesign.asp>
- What Works Clearinghouse Single-Case Technical Documentation
<https://ies.ed.gov/ncee/wwc/Document/229>

Summing up SCDs

- The individual as the unit of analysis, just like in practice
- Establishing level of responding over time to establish (a) normal course of behavior, and (b) reliability
- Within-unit replication and strategic comparisons to rule out threats to validity
- Pursuit of precision of control
- Design flexibility
 - Response-guided experimentation
 - Examining interaction effects for functional variables
 - Potential for use in practice

Kazdin (2011)

“If behavior can be turned on and off, so to speak, demonstrating a causal relation and ruling out threats to validity reflect extraordinary clarity. Indeed, such clarity, in my mind, arguably exceeds the clarity of an RCT that shows some statistically significant differences between groups.”

References

See handout