**Protocol for Determining High Quality Single Case Design that Would Allow for Causal Inference**

**Taken from WWC Standards Handbook**

Level of Intervention

What is the unit of intervention administration and data analysis?

1. Student
2. Classroom
3. Teacher
4. Course
5. School
6. Other \_\_\_\_\_\_\_\_\_\_\_\_\_

*Note*. A single case design unit need not be confined to the student, it can be a classroom or a school.

Design Criteria

Are the following criteria met?

1. The independent variable (i.e. intervention) must be systematically manipulated by the researcher who determines when and how the independent variable conditions should change
2. Is there at least one probe just prior to introducing the independent variable to establish baseline?

Additional Criteria for Multiple Baseline Design

When intervention was introduced for one participant, was there a change in data patterns for other baseline participants?

Type of Intervention

Is the intervention that has more permanent effects using a multiple baseline design rather than a ABAB design?

Number of Phases and Data Points

For the following types of singe case design, are the minimum number of phases and data points met?

1. Reversal/Withdrawal Design—at least 4 phases with at least 3 data points per phase.
2. Multiple Baseline—at least 6 phases at least 3 data points per phase within each case and there must be some degree of concurrence in timing of intervention implementation across cases.
3. Alternating Treatment—at least 4 points per condition and at most 2 data points per phase. Any phases based on more than 2 data points cannot be used to demonstrate the existence or lack of an effect because the design features rapid alternations between phases. For study design with more than 2 alternate treatments, each treatment comparison should be rated separately.

Outcome Measure

If the outcome is based on observations, the outcome must be measured systematically over time by more than one assessor. Each condition (e.g. baseline, intervention) must have at least 20% of the data points that meets inter-rater agreement. Minimum acceptable values of inter-rater agreement ranged from 0.80 to 0.90 for percentage agreement and at least 0.60 for Cohen’s kappa.

If the outcome measure is not an observation but an assessment, does the outcome measure satisfy the following:

1. Face validity?
2. Reliability (i.e. internal consistency of at least 0.5 or a test-retest reliability of at least 0.4 or inter-rater reliability of at least 0.5)?
3. Collected in a similar way across different data points?
4. For a repeated measure, was alternate forms of an assessment used?

Visual Analysis

Is there a baseline pattern of behavior established?

Was an effect established based on all the following 6 features?

1. Level (the mean score for the data within a phase)



1. Trend (best-fitting straight line for the data)



1. Variability (range or standard deviation of data)



1. Immediacy of the effect (the extent to which the level, trend, and variability of the last 3 data points in one phase are distinguishably different from the first 3 data points in the next phase)



1. Minimum overlap



1. Consistency of data in similar phases



Do the data across all phases of the study document 3 demonstrations of an effect at different times? (this criterion is analogous to finding a difference beyond chance in a two-group design)

*Note.*

If all the 6 criteria (level, trend, variability, immediacy, minimum overlap, and consistency) on establishing an effect were met and 3 effects were demonstrated at different times, then a causal effect can be established.

A non-effect can be one of the following:

* Data within the baseline phase do not provide a clearly defined pattern to establish the baseline
* Failure to establish a consistent pattern within any phase
* Long latency between introduction of the independent variable and change in outcome variable or there was an overlap between phases
* Inconsistent patterns across similar phases