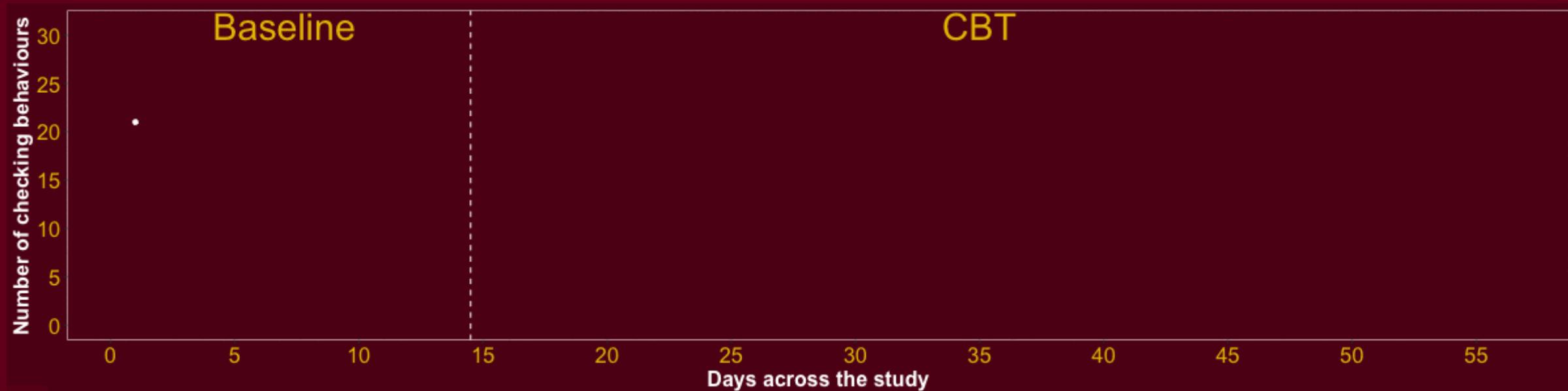




Single Case Experimental Design #2

A workshop for routine clinicians - Day 2: Analysing time series data

University of Sheffield - Clinical and Applied Psychology Unit





Steve Kellett

Clinical Psychologist

Sheffied Health & Social Care NHS
Foundation Trust



Mel Simmonds- Buckley

Researcher

University of Sheffield



Chris Gaskell

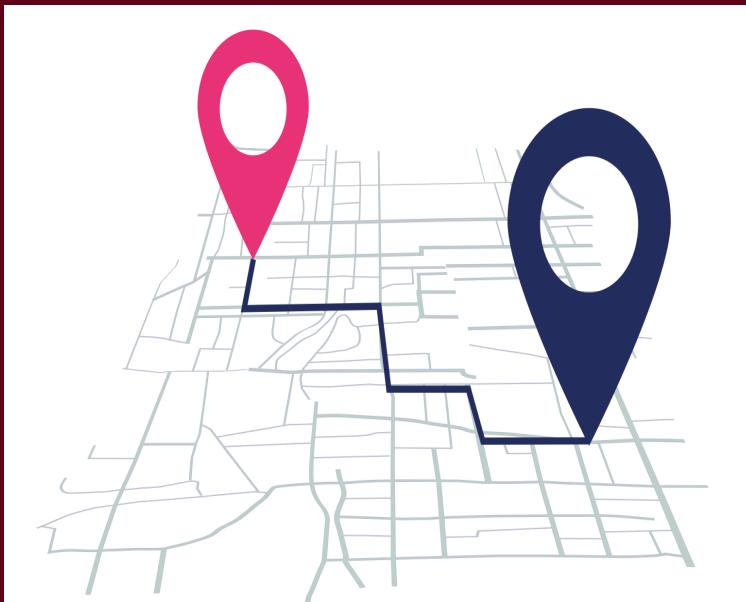
Clinical Psychologist

Salford Royal NHS Foundation Trust





① Navigating these slides



Day 1

Aims + Intro

Why

Measures

Designs

Setting it up

Day 2

Analysis

Practice

Resources

Note: all underlined white and red text are hyperlinks!



🧭 Navigating these slides

⌨️ Keyboard shortcuts - Press 'h'.

📍 Overview slides - Press 'o'.

✍️ Scribble on slides - Press 's'.

📄 This presentation is also available as a [pdf](#) and [powerpoint](#)

✉️ If you find any problems with these slides then e-mail [Chris](#).



Slides made using Xaringan and R Markdown.



Plan for Day 2

- Visual analysis and interpretation
- Statistical analysis and caveats
- Assessing nomothetic outcomes
- How it fits with your assignment
- SCED analysis software and resources
- Practical exercise: Have a go at some analysis!



Key features of SCED analysis

SCEDs are fundamentally different to group-based research.

- Group research -> averages/combined findings/overall effects
- SCED -> raw data/stability, variability, trends/individual effects

Visual analysis – this is the main analysis! Plot and analyse time series data visually

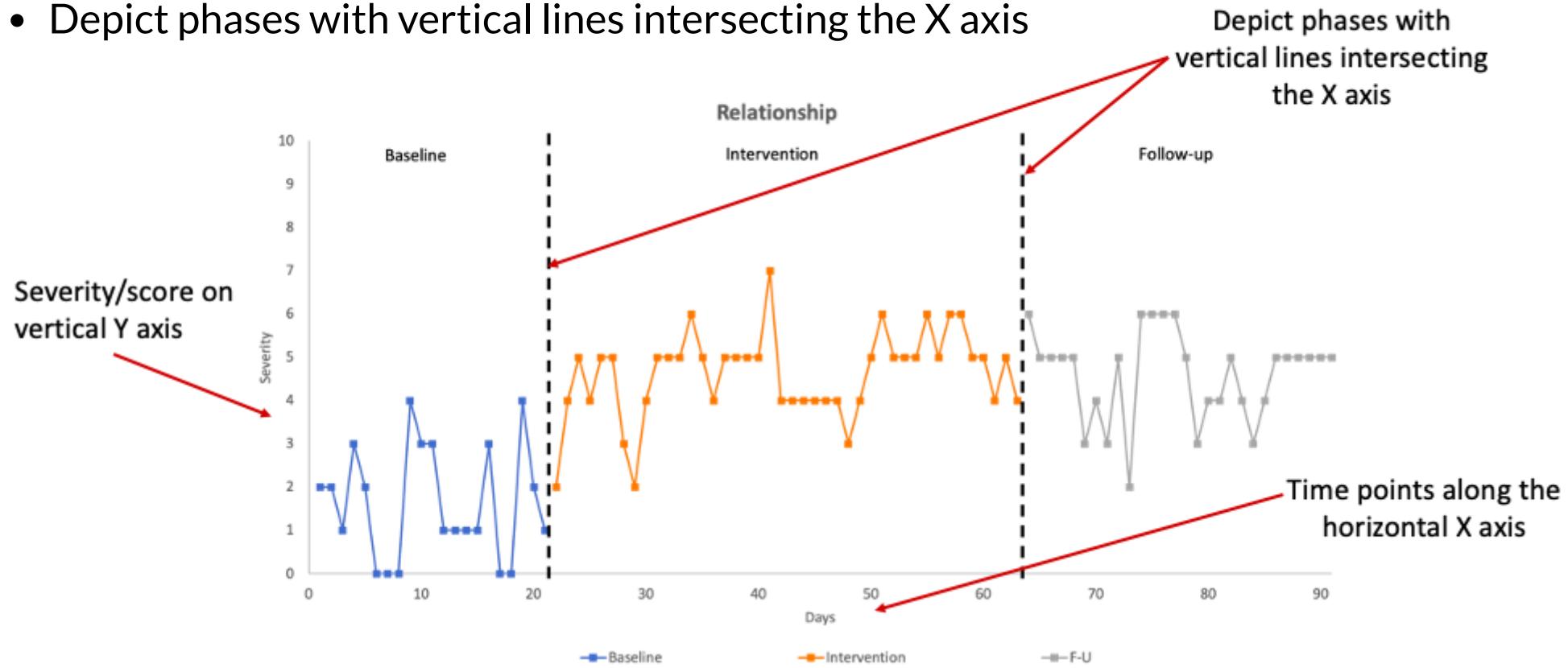
Statistical analysis:

- Use statistics to confirm visual analysis
- Never interpret SCED statistics in isolation – viewed alongside visual plots
- Be aware of the limitations of statistical analysis in SCED



Visual analysis

- Plot each idiographic measure in a time series graph.
 - Time points along the horizontal X axis.
 - Severity/score on vertical Y axis
- Depict phases with vertical lines intersecting the X axis





Visual

Central Tendency

Trend Lines

Trend Bands

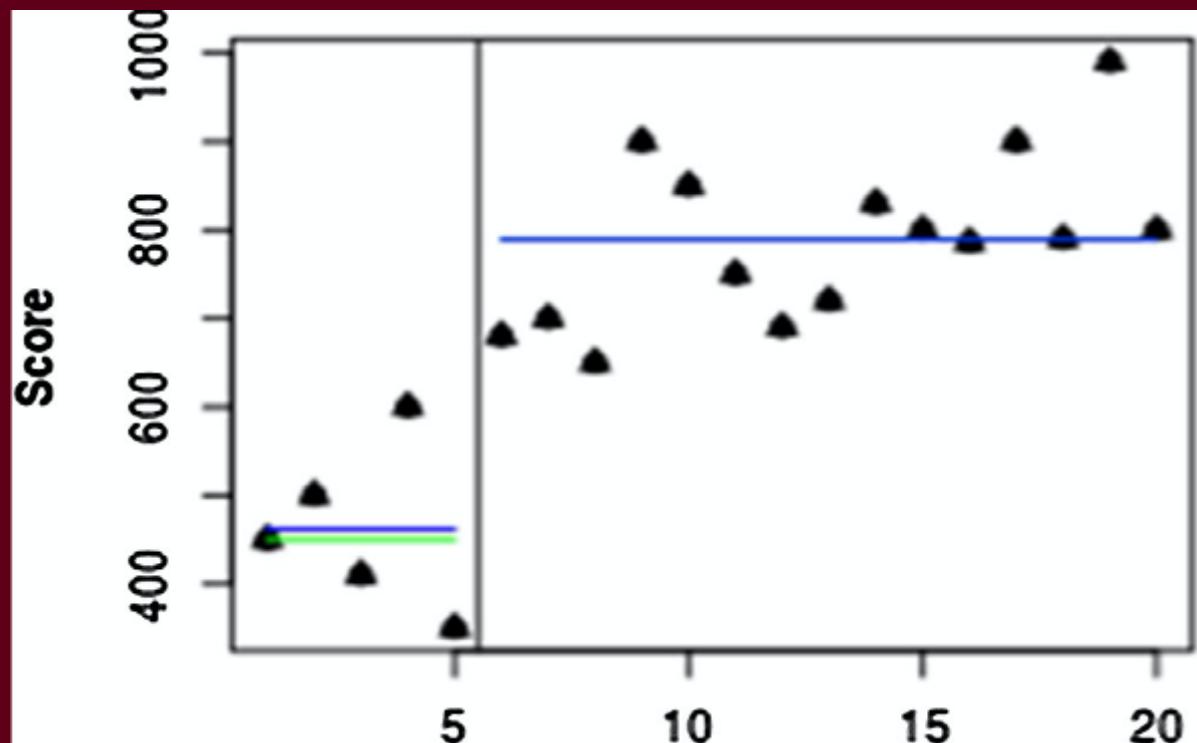
Stability Lines

Supplementing Visual
Analysis

[Visual](#)[Central Tendency](#)[Trend Lines](#)[Trend Bands](#)[Stability Lines](#)

Draw a horizontal line at the level of the measure of central tendency (e.g. median, mean) of data points in separate phases.

Central Tendency Lines.

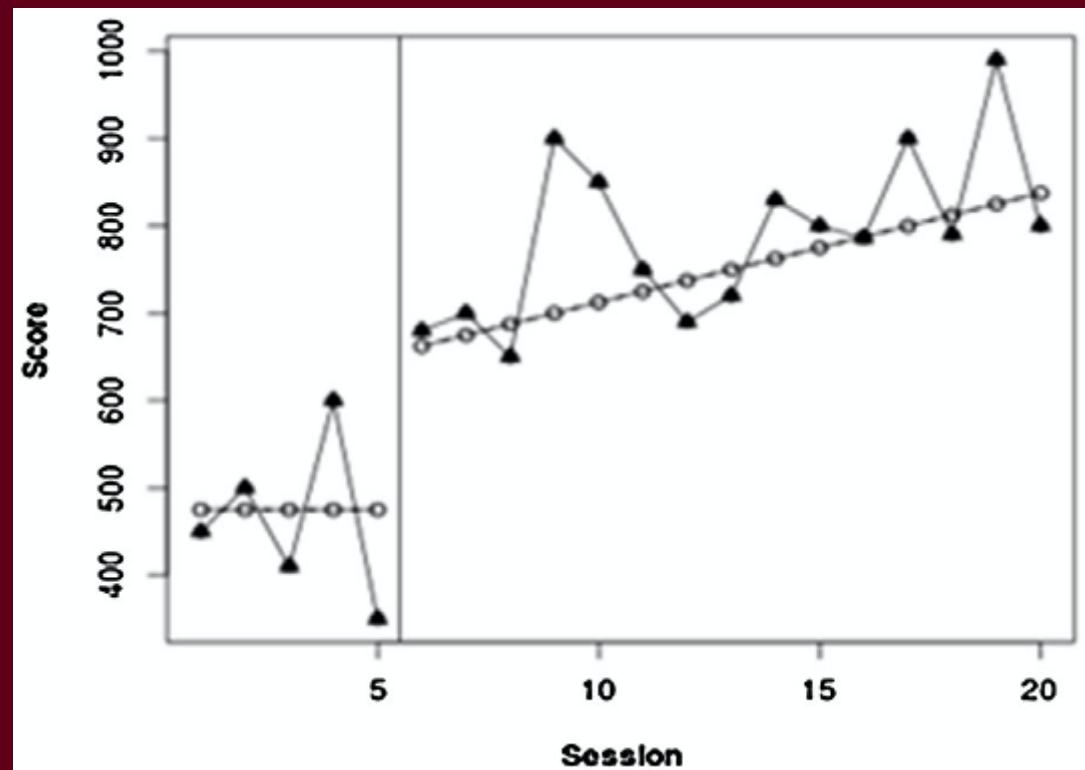


[Visual](#)[Central Tendency](#)[Trend Lines](#)[Trend Bands](#)[Stability Lines](#)

There are various methods for plotting trends:

- Split-middle trend line: Draw line connecting the median of each half phase.
- OLS regression trend line: Line drawn based on regression equation for each phase.

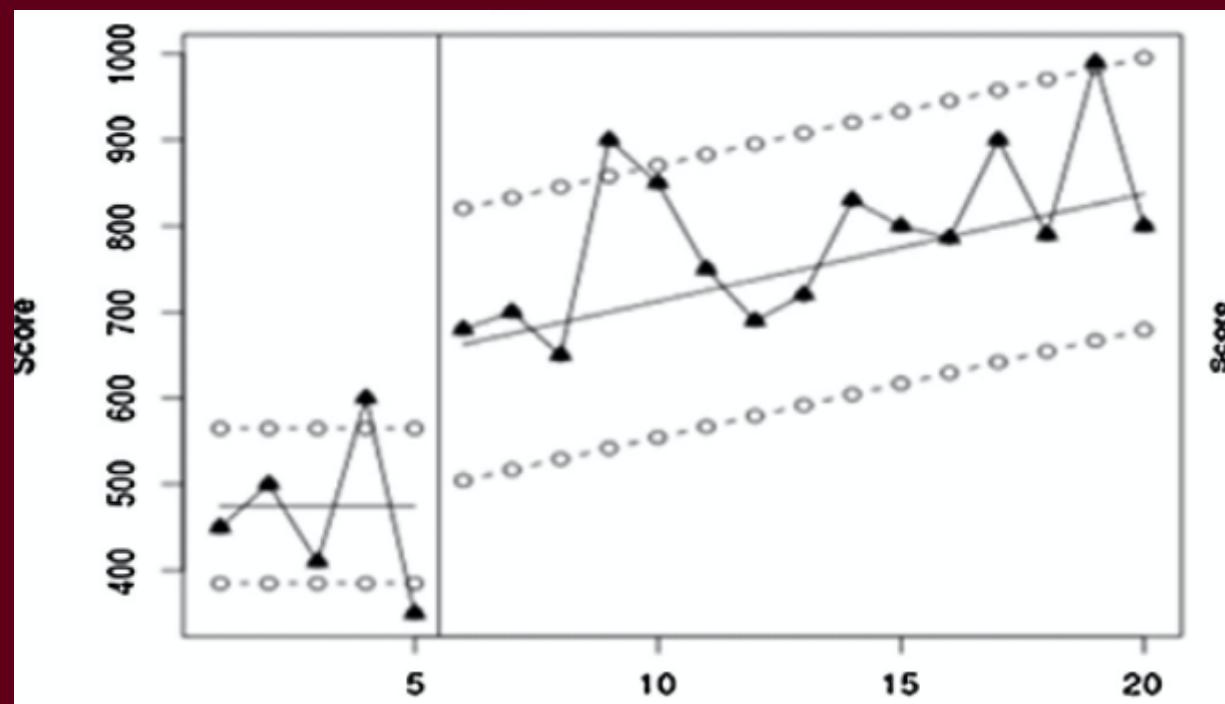
Trend Lines.



[Visual](#)[Central Tendency](#)[Trend Lines](#)[Trend Bands](#)[Stability Lines](#)

Suggested that trend lines only represent the data if 80% of the data points fall within envelope bands around the trend line.

Trend Line Stability Bands.

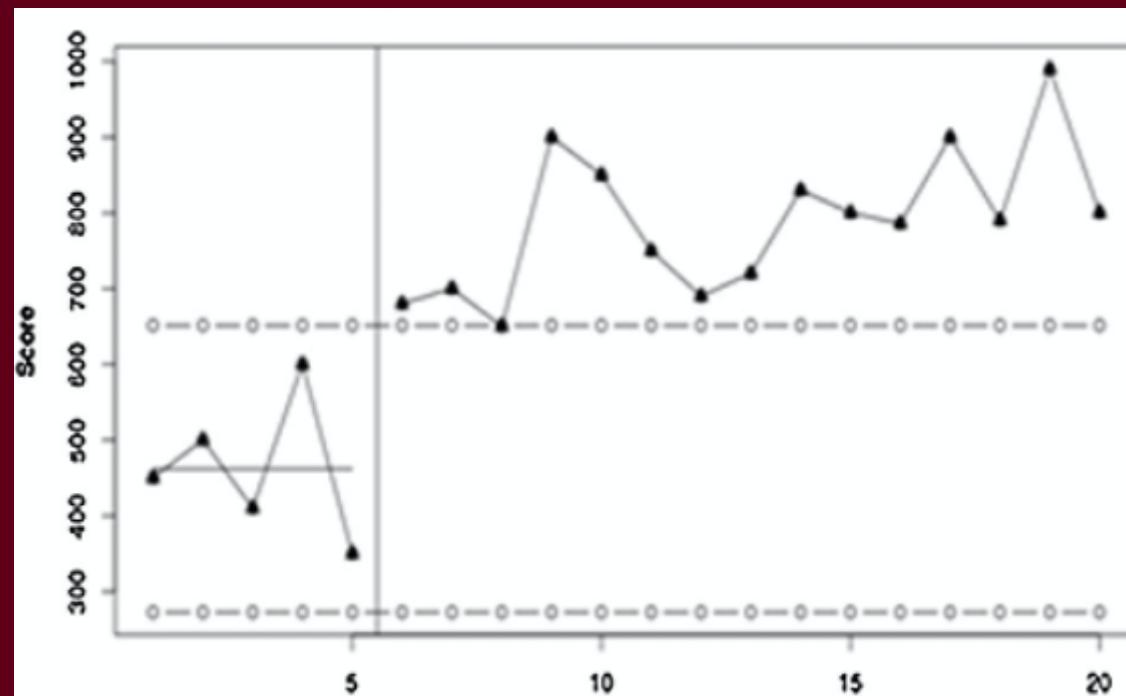


For source and further information on effective visual plots see [\(Krasny-Pancini & Evans, 2018\)](#).

[Visual](#)[Central Tendency](#)[Trend Lines](#)[Trend Bands](#)[Stability Lines](#)

Calculate the mean and SD of each phase and then draw bands 2 SD above and below the mean. Expand the bands into the comparison phase.

Stability Band Lines



For source and further information on effective visual plots see [\(Krasny-Pancini & Evans, 2018\)](#).



Interpreting visual plots

- How much overlap between data points in each phase?
- Less overlap = more likely to be differences between phases i.e. an effect!
- Smaller variability within phases = easier to detect effects
- Few overlapping data points between phases = intervention has had an effect.
- Lots of overlap between data points in the two phases = minimal effect of intervention



Table III. Visual Analysis Criteria to Evaluate Data in Single-Case Research

-
1. Changes in means—Visually detectable differences in the average rate of the data on the outcome measure as phases are changed (e.g., Baseline to Intervention).
 2. Changes in level—Differences in the pattern of the data from the very end of one phase (e.g., Baseline) to the very beginning of the next phase (e.g., Intervention).
 3. Changes in trend or slope—Differences in the trend lines (e.g., upward or downward movement) on the outcome measure across phases.
 4. Changes in variability—Differences in the variability (range or standard deviation) of the data across phases.
 5. Changes in latency—The elapsed time between the beginning of a phase (e.g., Intervention) and the subsequent change in the data.
 6. Consistency in patterns—Similarities in patterns of data from phases with the same conditions (e.g., Baseline and Return to Baseline).
-

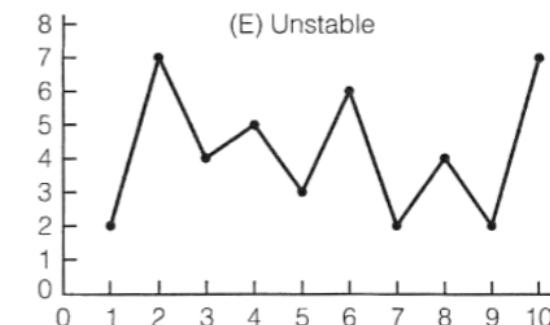
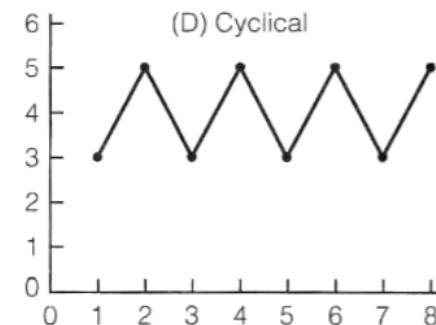
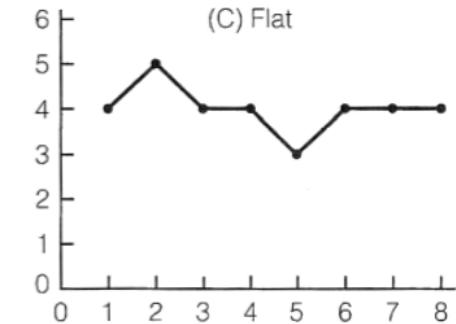
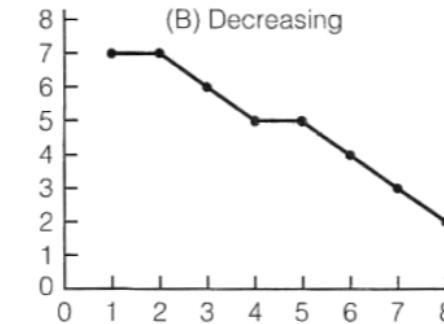
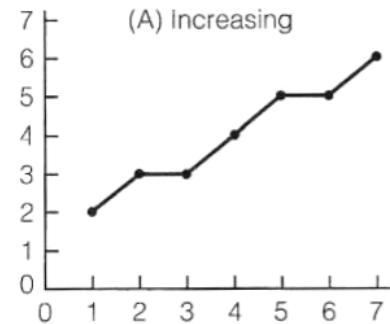
Note. These criteria should be assessed both individually and collectively.

Source: [Cohen et al \(2013\)](#).

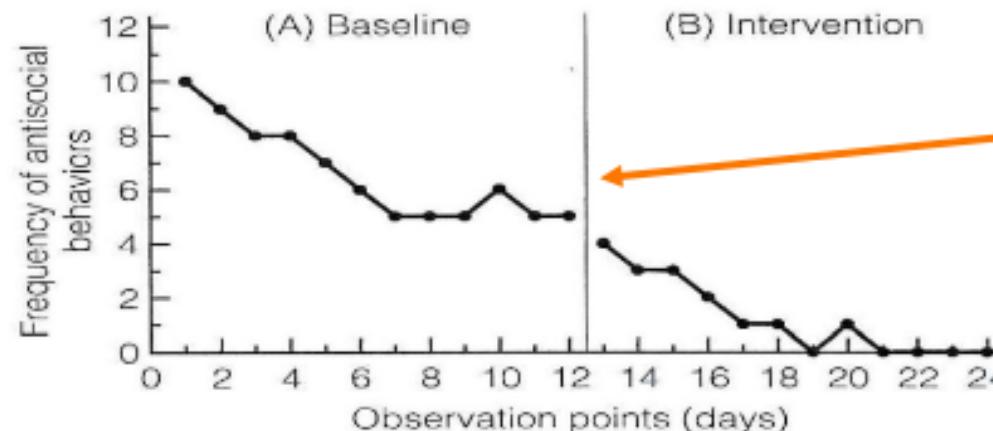


Interpreting baseline trends

- Baseline may not be stable. Always need to interpret intervention phase in the context of baseline trend.
- Statistics allows us to assess and adjust for baseline trends (more on Tau-U later)



For source click [here](#).



(B) Intervention

Improvement in the baseline – less of an intervention effect and continued spontaneous recovery?

Need to consider what might be a successful outcome when the baseline has variability or shows improvement

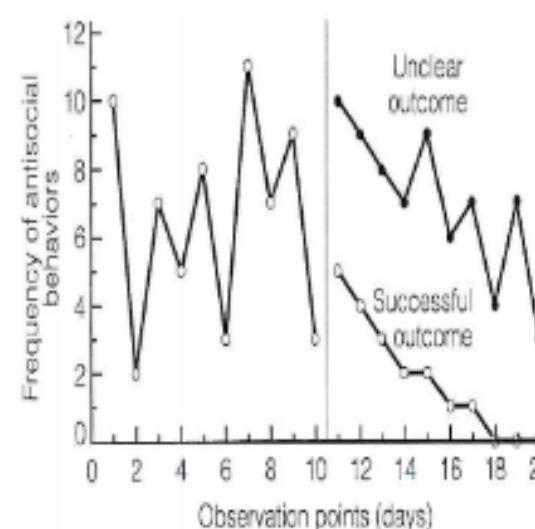


Figure 14-6 Graph of Two Hypothetical Outcomes with an Unstable Baseline (AB Design)

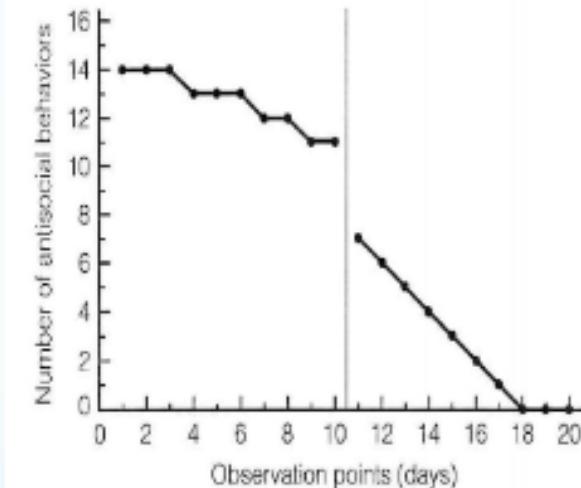


Figure 14-7 Graph of a Hypothetical Outcome Supporting Intervention Efficacy with an Improving Baseline (AB Design)



Statistical Analysis

.



Idiographic outcomes

Visual interpretation can be subjective so accompanying statistical analysis can aid interpretation.

However need to be aware;

- Can be influenced by number of time points/length of baseline
- Can be misleading if there is baseline trend
- Often difficult to compare
- No consensus on best statistical approach/effect size metrics
- Statistically significant changes may not have clinical significance
- Assumptions of many statistical tests are violated in time-series data (lack of independence caused by autocorrelation)



Autocorrelation

Autocorrelation (also known as serial dependence) is:

The degree of similarity between one data point in a time series and a previous lagged time point i.e. how correlated are the data points.

E.g., the weather is autocorrelated: the noon temperature on a Wednesday is predicted by what the noon temperature was on Tuesday, and to a lesser extent what the noon temperature was on Monday, or Sunday.

Time series data is often autocorrelated. A persons behaviour/thoughts/feelings are not random from day to day so the data is not strictly independent

- Statistical tests often based on the assumption that data are independent.
- When autocorrelation is present this assumption is violated (increases risk of Type 1 error [false positive])



SCED effect sizes

Metrics for capturing between phase differences

- Nonparametric (overlap) and parametric (SMD, LOR etc.)
- Historically - simple, yet commonly used overlap statistics highly criticised (e.g., PND)
- More recently developed overlap effect sizes have improved (e.g. NAP, Tau)

Good practice to calculate more than one overlap metric

Remember they are relative to whether the desired direction of improvement is shown by an increase or decrease in the measure

- Desired outcome is **increase**: intervention data points that are higher than the baseline.
- Desired outcome is **decrease**: intervention data points that are lower than the baseline.



Non-overlap effect sizes

Common effect sizes	Description *(if an increase is the direction of improvement)	By hand
Percentage exceeding the median (PEM)	Proportion of observations in phase B that improve upon the median of phase A.	Yes
Percentage of non-overlapping data (PND)	*Proportion of observations in the B phase that exceed the highest observation from the A phase. Severe limitations as influenced by no. of time points in the baseline.	Yes
Percentage of all non-overlapping data (PAND)	*Proportion of observations remaining after removing the fewest possible number of observations from either phase so that the highest remaining point from the baseline phase is less than the lowest remaining point from the treatment phase.	Yes
Nonoverlap of all pairs (NAP)	*Proportion of all pair-wise comparisons where the treatment phase observation exceeds the baseline phase observation, with pairs that are exactly tied getting a weight of 1/2.	No
Improvement rate difference (IRD)	The robust phi coefficient corresponding to a certain $2 \times 2 \times 2$ table that is a function of the degree of overlap between the observations each phase	No
Tau/Tau-U	Tau provides an effect estimate that does not make any adjustments for baseline trend. Tau-U is similar to Tau but includes an adjustment for baseline trends.	No



Non-overlaps

PND

PEM

PAND

NAP



Non-overlaps

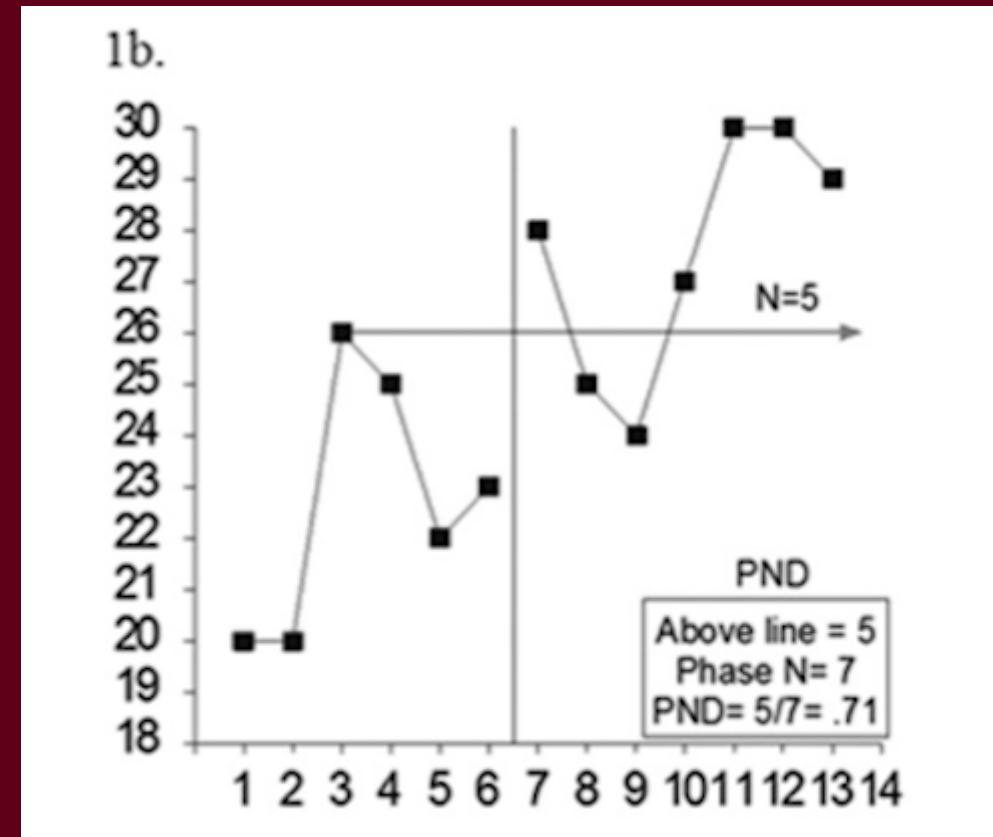
PND

PEM

PAND

NAP

Percentage of non- overlapping data



Original article = ([Parker et al., 2011](#))



Non-overlaps

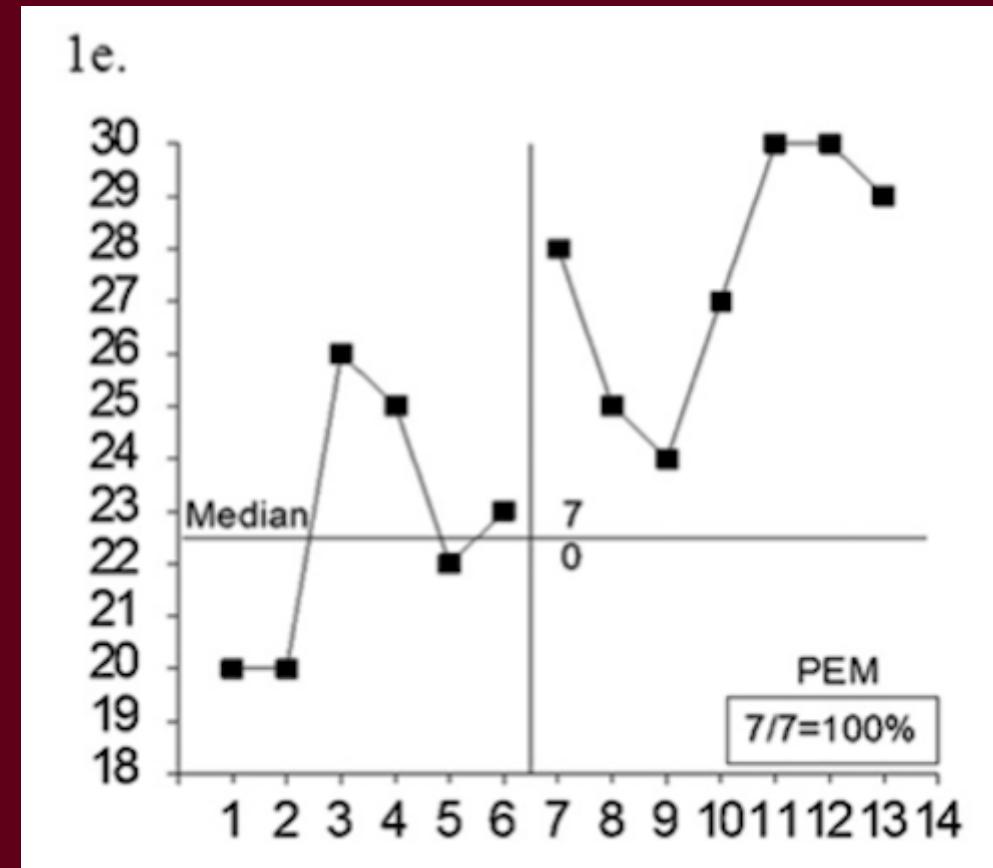
PND

PEM

PAND

NAP

Percentage
of
data
exceeding
the
medium



Original article = [\(Parker et al., 2011\)](#)



Non-overlaps

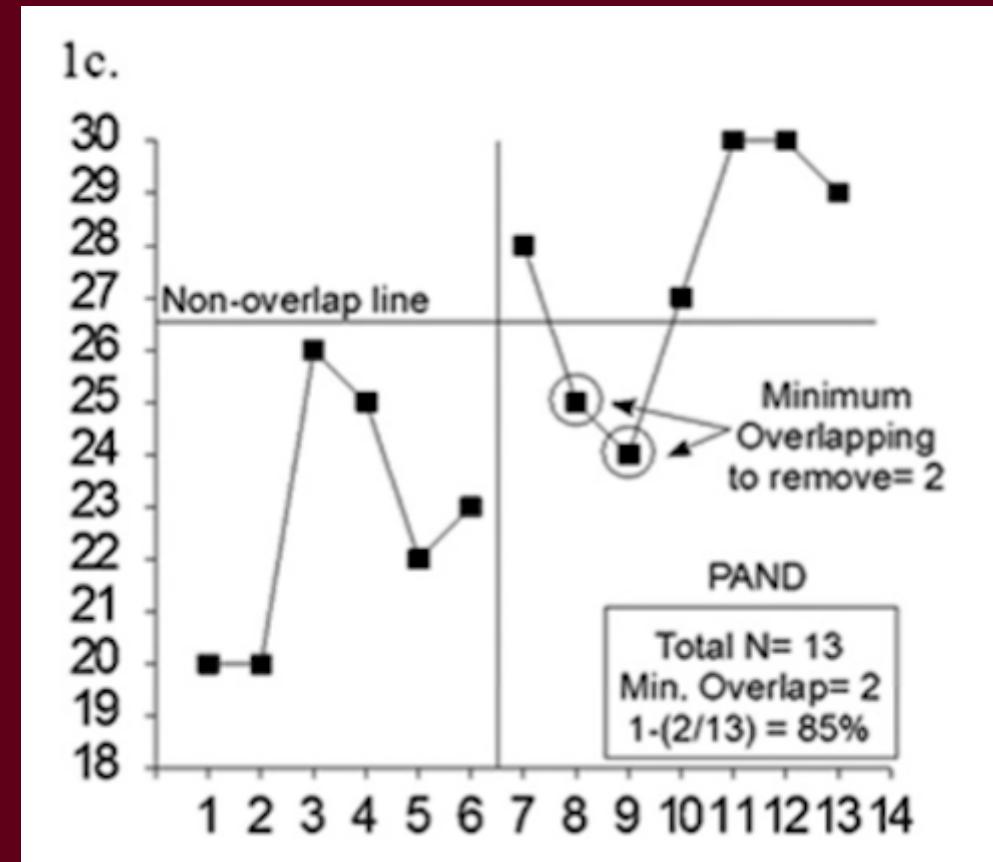
PND

PEM

PAND

NAP

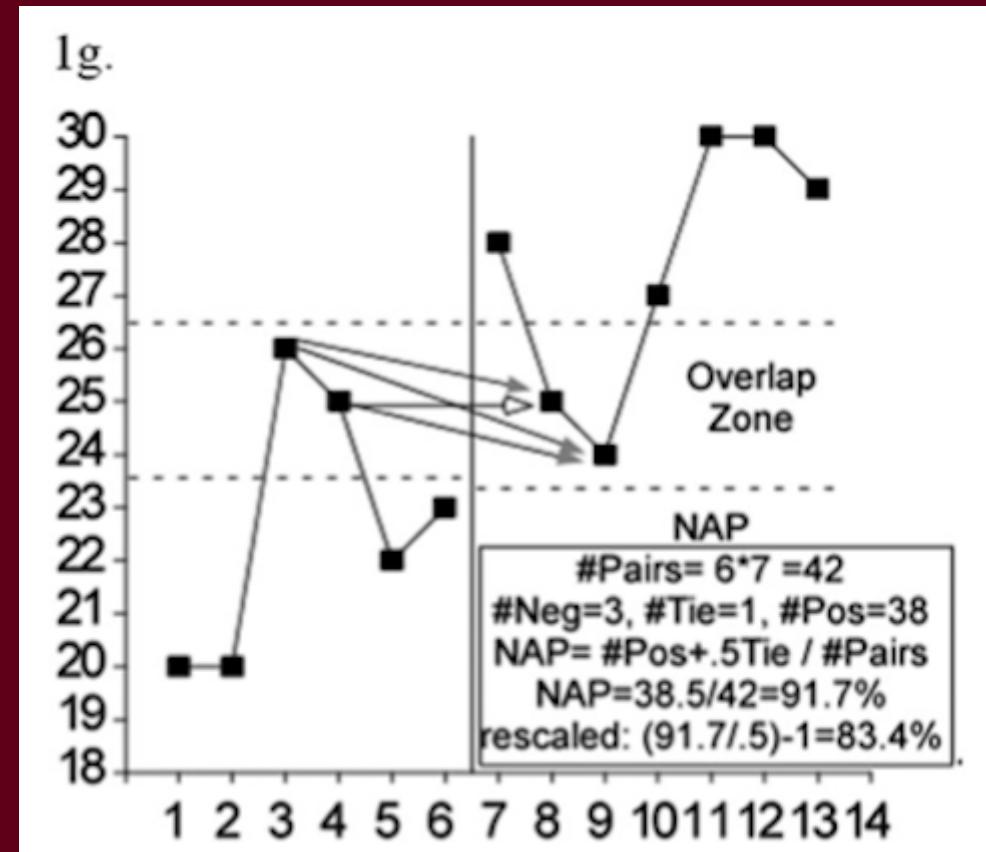
Percentage of all non- overlapping data



Original article = (Parker et al., 2011)

[Non-overlaps](#)[PND](#)[PEM](#)[PAND](#)[NAP](#)

Non-overlap of all pairs



Original article = [\(Parker et al., 2011\)](#)



Example: Percentage exceeding the median (PEM)

- Calculate median of baseline
 - =3
 - Count number of intervention phase datapoints that are above the median
 - =37 out of 42
 - Convert to a %
 - $37/42 * 100 = 88.10\% \text{ PEM}=88.10\%$



The
University
Of
Sheffield.

Psychotherapy
Evaluation
And
Research
Lab @
Sheffield



Example: Percentage of all non-overlapping data (PAND)

- Count how many extreme datapoints would need to be removed to there is no overlap
 - =6 (red circles)
 - Convert to %
 - $6/63 * 100 = 9.52\%$
 - Take the inverse
 - $100 - 9.52 = 90.48\% \text{ PAND} = 90.48\%$



Interpreting non-overlap effect sizes

Effects are provided as a proportion (or *100 to make a %)

The larger the effect size, the more effective the treatment

No agreed way to interpret effect sizes but Scruggs & Mastropieri (1998) suggested the following interpretation:

- 0.90+(or 90%) indicative of a very effective treatment
- 0.70-0.89 (or 70-89%) represent moderate effectiveness
- 0.50-0.69 (or 50-69%) are debatably effective
- <0.50 (50%) are regarded as not effective



Tau/Tau-U for baseline trend

Tau/Tau-U family of effect sizes calculates overlap with adjustments for baseline trend and is robust to serial dependency (autocorrelation).

If baseline data is stable (for time series data) then Tau (τ AvsB) is a sufficient index of outcome. If there is significant trend then use Tau-U (τ AvsB-Atrend).

so first assess whether there is a significant trend in the baseline phase (Tau trend A). If there is a significant baseline trend ($p < .05$) then adjust using Tau-U.

Dont do this by hand!

- Online Tau calculator: <http://ktarlow.com/stats/tau> OR
- Shiny apps: <https://manolov.shinyapps.io/Overlap/>

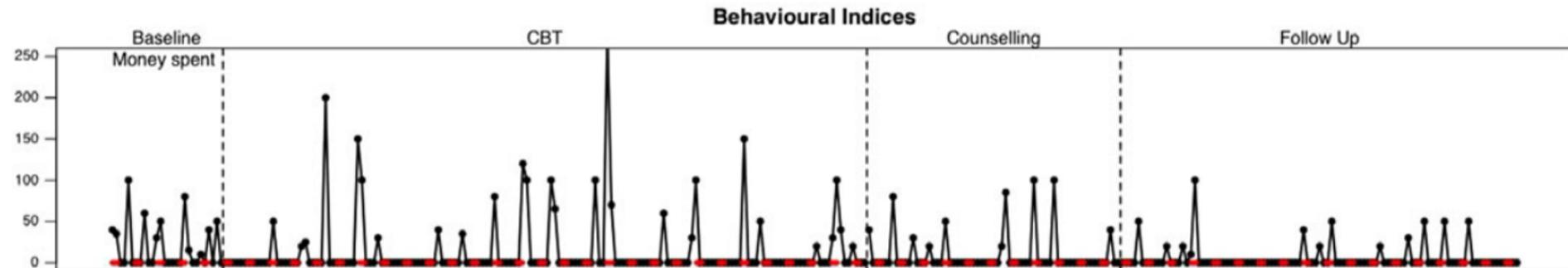


Comparing multiple phases

If you have more than two phases (i.e. any design other than A/B) then do pairwise comparisons for each bi-phasic combination of interest.

In example of AB-F/U design:

- Baseline (A) vs intervention (B)
- Baseline (A) vs follow-up (FU)
- See example published papers of different designs in your pack





Assessing nomothetic outcomes

Does pre/post phase scores on nomothetic measures meet criteria for reliable and clinical significant change (Jacobson & Traux, 1991)?

1. Has the client shown a **reliable change** (RC)?

- Reduction in scores beyond that which could be due to measurement error.
- The RC threshold will be published for most widely used measures.
- If there isn't one then this can be calculated! Use the reliability (Cronbach's Alpha) of the measure and non-clinical and clinical norms to calculate the RC.

2. Has the client shown **clinically significant change**? (CSC)

- Moved from above to below a clinical cut-off threshold.
- Use defined cut-off's (if available) or calculate using clinical and non-clinical norms.
- Process or generalisation measures may not make conceptual sense to assess clinically significant change (e.g. activation or quality of life).

An Excel workbook & manual is provided in the Resource Pack that will allow you to do this!



Evaluating nomothetic outcomes using reliable & clinically significant change

- Nomothetic measure of general distress: CORE-OM
 - Connell et al (2007) paper reports psychometric evaluation including;
 - Reliability of the measures (Cronbach's Alpha = 0.91)
 - Clinical and non-clinical norms (mean & SD)
 - Established reliable change and clinical cut-off thresholds
 - Pre-post change ≥ 6 shows Reliable Change
 - Pre-post score moving from above to below 10 shows Clinically Significant Change



What do you need to include in your assignment?

1. Visual analysis:

- Visual plots of every idiographic measure.
- Include phase trend lines and baseline median line.

2. Summary descriptives:

- Report means & SDs for each idiographic measure in each phase

3. Statistical analysis:

- Assess for baseline trend (appropriate Tau correction if trend is present)
- Report at least 3 non-overlap statistics
- Assess autocorrelation in data (instructions provided in resources).

4. Nomothetic outcomes:

- Calculate whether RC and CSC have occurred for the primary nomothetic measure.
- For the other nomothetic measures calculate RC and if relevant, calculate CSC.



Deciding what software to use

Data input and management: MS Excel

Lots of the analysis can be done by hand (if you want!)

More sophisticated techniques require software/web calculators (they will also do the simpler analyses/visualisation as well!)

See SingleCaseSoftware table (next slides) for summary of some of the available programs/web calculators for SCED analysis.

- What you can do in each program
- Benefits and limitations of each program
- Complexity to learn



Quick intro to SCED software

- Microsoft Excel
- Tau online calculator: <http://ktarlow.com/stats/tau/>
- Jepusto Shiny app: <https://jepusto.shinyapps.io/SCD-effect-sizes/>
- SCDA Shiny app: <https://tamalkd.shinyapps.io/scda/>
- Overlap Shiny app: <https://manolov.shinyapps.io/Overlap/>
<http://ktarlow.com/stats/tau/> <https://jepusto.shinyapps.io/SCD-effect-sizes/>
<https://tamalkd.shinyapps.io/scda/> <https://manolov.shinyapps.io/Overlap/>

	IDIOGRAPHIC VARIABLES							NOMOTHETIC VARIABLES	ASSIGNMENT REQUIREMENT
	Microsoft Excel	ShinyApps	R	Web calculator	SPSS	Microsoft Excel			
Data Management									
Data file requirements	any	Copy + Paste	.txt, .csv or .xls	any	Copy & paste	SPSS data	Single-case-V8 excel spreadsheet calculator	n/a	
Data input requirements	any	Phase A and B scores	Phase & score columns	"score" & "phase"	Phase A and B scores	Scores & phase	manual input	n/a	
Imputation features	No	No	No	Yes	No	Yes	n/a	No	
Assess Outliers	No	No	No	Yes	No	Yes	n/a	No	
Design Plotting Compatability									
A/B	Manual	Yes	Yes	Yes	No	Yes	n/a	Any quasi-experimental or SCED design	
A/B-F/U	Manual	No	Yes	No	No	Yes	n/a	Any quasi-experimental or SCED design	
ABC / ABA	Manual	No	Yes	No	No	Yes	n/a	Any quasi-experimental or SCED design	
ABAB	Manual	No	Yes	No	No	Yes	n/a	Any quasi-experimental or SCED design	
Visual Analysis									
Plot Function	Yes	Yes	Yes	Yes	No	Yes	n/a	Yes	
Customisability	Highly	No	Some (can add manually)	Some (can add manually)	Highly	No	Some	n/a	
Central Tendency lines	Manual	No	Yes	Yes	Via non SCED packages	No	Some	n/a	Yes - median
Trend lines	Manual	No	Yes	Yes	Yes	No	Some	n/a	Yes
SD Bands + Stability Lines	Manual	No	Yes	Yes	Via non SCED packages	No	No	n/a	No
Statistical Analysis									
Descriptive Statistics	Manual	No	No	Yes	Yes	No	Yes	n/a	Yes
Auto-correlation	No	No	No	No	Yes	No	Yes	n/a	Yes
Assess baseline trend	No	No	No	Yes	Yes	Yes	No	n/a	Yes
Normal distribution	No	No	No	No	Via non SCED packages	No	Yes	n/a	No
PND	Manual	Yes	No	Yes	Yes	No	Manual	n/a	Minimum of 3 non-overlap tests
PEM	Manual	Yes	No	Yes	Yes	No	Manual	n/a	Minimum of 3 non-overlap tests
PAND	Manual	Yes			Yes	No	Manual	n/a	Minimum of 3 non-overlap tests
NAP	No	Yes	No	Yes	Yes	No	Potentially	n/a	Minimum of 3 non-overlap tests
IRD	No	Yes	No	No	Yes	No	Potentially	n/a	Minimum of 3 non-overlap tests
Tau/TauU	No	Yes	No	Yes	Yes	Yes	Potentially	n/a	Yes
PEM-T	No	No	No	Yes	Yes	No	No	n/a	No
Log Odds + Response Ratios	No	Yes	No	Yes	Yes	No	Potentially	n/a	No
SMD	Manual	Yes	No	Yes	Yes	No	Potentially	n/a	No
Piecewise regression	No	No	No	No	Yes	No	Manual	n/a	No
Reliable change analysis	n/a	n/a	n/a	n/a	n/a	n/a	Yes	Yes (for nomothetic outcomes)	
Complexity to learn	Fair	Low	Low	Low	High	Low	Fair	Low	n/a
Best for	Flexibility and customising of plots.	Provides all the major non-overlap statistics and parametric ES for pairwise phase comparisons (statistical analysis can compare each pair of phases in any design).	Plotting all types of design with lots of trend line options.	Most detailed analysis of baseline trend using Tau & Tau-u ES. Also can do all stages of analysis in one place (apart from autocorrelation).	Able to perform all analyses and more for all types of design. High level of customisability to be able to produce high quality plots.	Very easy and quick assessment of baseline trend and clearly directed whether to use Tau or apply a baseline correction and use Tau-U.	Able to assess autocorrelation and provide lag plots easily.	Can calculate RCI and CSC for a single-case.	n/a
Limitations	All manual so more time consuming. Not able to assess baseline trend or Tau. More prone to errors if statistics done by hand.	Can only plot A/B phase designs and no option of adding trend lines.	Does not perform any statistical analysis. Trend line plots do not join up data points.	Can only handle A/B phase designs. Plots can be limited i.e. Trend line plots do not join up data points. Uses .txt file and upload errors are common.	Very difficult to learn and use if not familiar with R.	Only performs Tau/Tau-U analysis so cannot be used alone. Tau analysis output is not as detailed as Manolov-Overlap shiny app.	Not obvious how/difficult to apply other types of SCED analysis and visualisation.	Need psychometric data of measure: reliability (alpha) for RCI and clinical/healthy norms for CSC	n/a



SCED Resource Pack

- 1. Teaching slides (2 day workshop)
- 2. SCED data analysis stages and guides
 - a. AB design using Overlap Shiny App
 - b. AB-FU design using SCDA and SCD-effect-sizes Shiny Apps
 - c. AB-FU design using Excel (including Tau online calculator)
- 3. Example Excel datasets for 4 SCED designs and summary of results
 - a. AB (Adult inpatient)
 - b. AB-F/U (Adult mental health) version for SCDA/SCD Shiny guide
 - c. AB-F/U (Adult mental health) version for Excel guide
- 4. Glossary of SCED statistics
- 5. SCED Software summary table (*SingleCaseSoftware Table*)
- 6. Single-case RCSI Excel workbook and Manual
- 7. Guide to calculating autocorrelation in SPSS including
 - a. Template SPSS data file (*Autocorrelation_data_template.sps*)
 - b. SPSS syntax file (*Autocorrelation_syntax.sps*)
- 8. SCED Reporting Guidelines (SCRIBE)
- 9. Published examples of SCED designs
- 10. SCED guidance papers
 - a. SCED Practical Guide
 - b. SCED Data Analysis Considerations
 - c. Effect size calculation
 - d. Overlap statistics interpretation (Scruggs & Mastropieri (1998))
 - e. Excel guide for producing SCED plots



Practical task

We will divide you into groups of 4

Two example datasets – each group will be assigned a dataset either Adult mental health setting (AB-F/U) OR Adult inpatient setting (AB)

Each has a guide to walk you through conducting and reporting the analysis in different types of software

There are multiple idiographic measures – you don't have to do all of them in the practical – the examples are provided so you have a guide for the full analysis process you will need to do for your assignment

- You may want to divide the measures up between your group



Feedback your findings

Present informal overview of findings (~5-10 mins per group).

PowerPoint template provided to copy and paste your results into.

Things to report on;

- Visual plot of each idiographic measure
- Assess whether there is a baseline trend using Tau/Tau-U
- Calculate x3 non-overlap statistics (Tau/Tau-U counts as one!)
- Descriptive statistics for each phase (means and SDs)
- RCSI analysis of nomothetic outcomes

Don't worry too much about having a detailed interpretation!

- This is about practicing, troubleshooting any becoming familiar with the resources.
- Each dataset has a summary of the results so you can sense check if you have produced the same findings – just don't cheat!



Feedback

- How did you find using the different SCED software?
- Any problems?
- Any questions/queries?
- Anything you are concerned/unsure about?

\



References

- Cohen et al (2013) Single-Case Research Design in Pediatric Psychology: Considerations Regarding Data Analysis. *Journal of Pediatric Psychology*, 39, 124-127.
 - Krasny-Pacini & Evans (2018) Single-case experimental designs to assess intervention effectiveness in rehabilitation: A practical guide. *Annals of Physical and Rehabilitation Medicine*, 61, 164–179