RESEARCH · RESEARCH DESIGN

Single-Case Experimental Designs for Generating Rigorous Idiographic Knowledge

How single-case experimental designs establish cause and effect for one person, team or unit through repeated measurement and planned changes, the main designs (reversal, multiple baseline, alternating treatments, changing criterion), standards for rigour, analysis, and uses in health, education and business.

ModQu research article · 1,727 words · 17 references (APA 7th edition)

Most research learns about people by studying groups and averaging across them. **Single-case experimental designs** take the opposite approach. They study one individual, or one small unit such as a classroom, team or store, intensively over time, and they establish whether an intervention works for that case by measuring an outcome repeatedly while the intervention is introduced, withdrawn or staggered in a planned way. The aim is **idiographic** knowledge, about a particular case, rather than the **nomothetic** knowledge, about populations, sought by group studies, a distinction Allport (1937) brought into psychology. Single-case designs are widely used in clinical psychology, special education, rehabilitation and behaviour analysis, and in medicine as "n-of-1" trials. This article explains their logic, the main designs, the standards for judging them and their uses.

Why study one case?

Group studies answer the question "what works on average?" But an average effect may describe few individuals. A treatment that helps half of patients a great deal and harms the other half may show no average effect at all. Fisher et al. (2018) analysed intensive repeated measurements of individuals and showed that relationships estimated across a group often differed substantially from those found within individuals. Molenaar (2004) argued, on mathematical grounds, that findings from groups can be generalised to individuals only under strict conditions that human behaviour rarely meets.

When the question is "does this intervention work for this person, class or team?", a single-case design answers it directly. There is also a long history of scientific discovery from intensive study of individuals. Ebbinghaus (1913/1885) established basic laws of memory and forgetting by testing himself repeatedly, and Skinner (1938) built the experimental analysis of behaviour on the detailed study of individual organisms. Normand (2016) argues that studying fewer participants more intensively often yields more reliable knowledge than studying many participants briefly.

The logic of single-case experiments

A single-case experiment replaces the comparison between groups with comparisons within one case over time. Three features make this work (Kazdin, 2011).

Repeated measurement. The outcome is measured many times, often daily or weekly, so that its natural level, trend and variability can be seen.

A baseline phase. Before the intervention, the outcome is measured long enough to establish a stable pattern. The baseline serves as the prediction of what would happen without intervention.

Replication of the effect. A single change from baseline to intervention could coincide with something else that happened at the same time. Confidence comes from showing the effect repeatedly: the outcome changes each time the intervention is introduced, and only then. Horner et al. (2005) proposed that a design should demonstrate the effect at least three times, at three different points in time, before it is considered experimental evidence.

Main designs

Reversal (ABAB) designs

In a **reversal** or **withdrawal** design, a baseline phase (A) is followed by an intervention phase (B), then a return to baseline (A) and a reintroduction of the intervention (B). If the outcome improves each time the intervention is introduced and returns towards baseline each time it is withdrawn, the intervention is very likely responsible (Barlow et al., 2009).

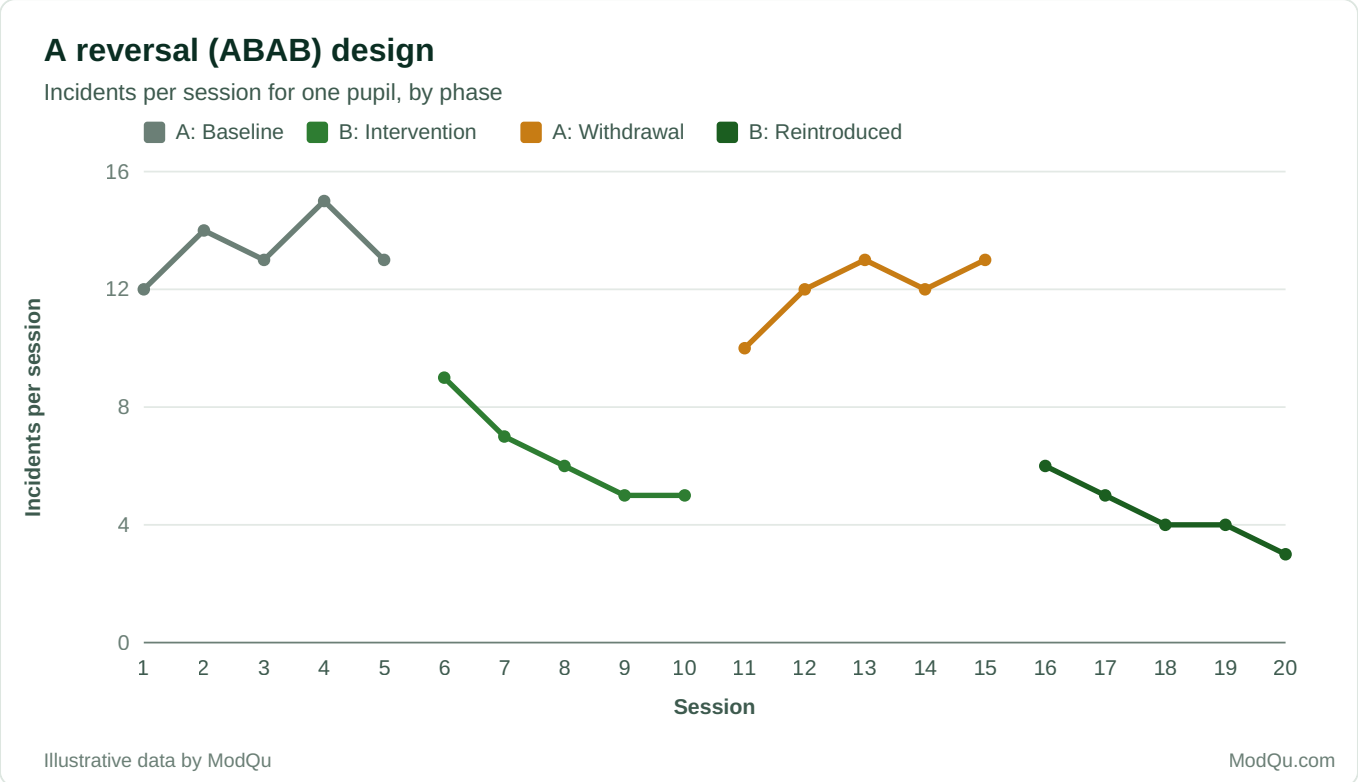


Figure 1. An illustrative ABAB (reversal) design. The outcome changes each time the intervention is introduced or withdrawn, which makes it very likely that the intervention, not some other event, produced the change. Illustrative data by ModQu.

The weakness is that some interventions cannot or should not be withdrawn. A skill once learned cannot be unlearned, and withdrawing an effective treatment may be unethical.

Multiple baseline designs

In a **multiple baseline** design, the intervention is introduced at different times across several people, settings or behaviours. For example, a new coaching method might be introduced for one sales team in week 4, a second in week 7 and a third in week 10. If each team's performance changes only when the coaching begins, and not before, the staggered timing rules out explanations based on events that affected all teams at once. No withdrawal is needed, which makes this design suitable for interventions with lasting effects.

Alternating treatments designs

An **alternating treatments** design switches rapidly between two or more interventions, such as different instructions or schedules, in a random or counterbalanced order, and compares the outcomes under each. It is efficient for comparing options within one case.

Changing criterion designs

A **changing criterion** design sets a series of increasingly demanding targets, for example gradually reducing the number of cigarettes allowed per day. If behaviour closely tracks each new criterion, the intervention is shown to control it.

Four single-case designs

Each shows the effect repeatedly within one case

Design	How the effect is shown	Best when
Reversal (ABAB)	Outcome changes each time the intervention starts or stops	Effects reverse and withdrawal is ethical
Multiple baseline	Outcome changes only when the staggered start reaches each person or site	Effects are lasting or withdrawal is not possible
Alternating treatments	Outcomes differ consistently between rapidly switched conditions	Comparing options quickly
Changing criterion	Behaviour tracks each new target as it is set	Gradual change towards a goal

Based on Kazdin (2011); Barlow et al. (2009) ModQu.com

Figure 2. Four main single-case experimental designs. Based on Kazdin (2011) and Barlow et al. (2009); table by ModQu.

Standards for rigour

Single-case research has developed explicit quality standards. Kratochwill et al. (2013) set out design standards for the What Works Clearinghouse, a United States review body for education research. Among other things, they require that the independent variable be systematically manipulated by the researcher, that outcomes be measured by more than one assessor with acceptable agreement, that each phase contain enough data points, typically at least three to five, and that the design allow at least three demonstrations of the effect.

Kratochwill and Levin (2010) argued that credibility is strengthened further by **randomisation**: for example, randomly choosing when the intervention starts, or the order of conditions in an alternating treatments design. Randomisation guards against the researcher unconsciously timing changes to coincide with favourable moments and allows formal statistical tests. Smith (2012) reviewed published single-case studies and found that many did not meet current standards, particularly in the number of data points and the reporting of measurement reliability.

Analysing the data

Single-case data have traditionally been analysed by **visual analysis**: inspecting a graph of the outcome over time and judging changes in level, trend and variability between phases, the immediacy of any change and the overlap between phases (Kazdin, 2011). Visual analysis is transparent, but different analysts can reach different conclusions from the same graph.

Statistical methods complement it. **Non-overlap indices** measure the proportion of intervention data points that exceed the baseline; Parker et al. (2011) reviewed nine such techniques. Shadish et al. (2014) developed a standardised mean difference for single-case designs that is comparable to effect sizes in group studies, allowing single-case results to be combined in meta-analyses.

N-of-1 trials in medicine

In medicine, single-case experiments are known as **n-of-1 trials**. A patient receives a treatment and a comparison, such as a placebo or an alternative drug, in several randomised, often blinded, periods, and outcomes are compared. Guyatt et al. (1986) described how such trials could determine the best therapy for individual patients when clinicians were uncertain. Lillie et al. (2011) argue that n-of-1 trials are well suited to personalised medicine, especially for chronic conditions with stable symptoms and treatments that act quickly and wear off quickly.

Uses in applied and business settings

The logic of single-case designs applies well beyond clinics and classrooms. Baer et al. (1968), in a founding statement of applied behaviour analysis, emphasised that applied research should address socially important behaviour, measure it reliably and show that the intervention, rather than something else, produced the change. Organisations face similar questions: does a new scheduling system improve one warehouse's accuracy, or does a new layout increase sales in one flagship store? With repeated measurement, a stable baseline and a staggered or reversed introduction, such questions can be answered with experimental rigour for the unit in question, without needing many comparable units.

A multiple baseline design is particularly practical. Introducing a change in a few sites at staggered times, and tracking a key measure weekly in each, provides much stronger evidence than a simple before-and-after comparison, at little extra cost.

A worked example

A distribution company wants to know whether a new pick-list format reduces picking errors in its warehouses (illustrative). It has only three warehouses, too few for a group comparison. Using a multiple baseline design, it records the daily error rate in all three for several weeks. The new format is introduced at the first warehouse in week 3, at the second in week 6 and at the third in week 9, with the order of warehouses chosen at random.

If errors fall at each warehouse shortly after the new format arrives there, and not before, the pattern is hard to explain any other way. A seasonal change, a new manager or a software update would be expected to affect all three warehouses at the same time, not in the staggered sequence of the intervention. If instead errors fall everywhere in week 5, the company learns that something other than the new format was responsible.

The design also shows things a simple before-and-after comparison would hide: how quickly the effect appears, whether it fades, and whether it is larger in one warehouse than another. Each of these is useful for deciding whether, and how, to roll out the change elsewhere.

Reporting single-case studies

Clear reporting allows readers to judge a single-case study and to combine it with others. A report should describe the participant or unit and setting in enough detail for others to judge relevance, define the outcome and how it was measured, state how many measurements were taken in each phase and how phase changes were decided, report measurement agreement, show the full graphed data for every phase and participant, and describe any randomisation used. Smith (2012) found that omissions in these areas were common in published studies and made results difficult to evaluate.

Limitations

Single-case designs show whether an intervention works for the cases studied. Generalising to other cases depends on replication across individuals and settings, and on understanding why the intervention worked. They require many measurements, which can be burdensome. They work best when outcomes can be measured frequently and respond fairly quickly to the intervention. Slow, irreversible or rare outcomes are poorly suited to them.

Conclusion

Single-case experimental designs provide rigorous, causal evidence about individual cases. By measuring outcomes repeatedly, establishing a baseline and introducing, withdrawing or staggering an intervention so that its effect is shown several times, they answer the question that matters most to practitioners: does this work here, for this person or unit? Combined with replication across cases, they also contribute to general knowledge, and they offer organisations a practical way to test changes when large comparison groups are not available.

References

Allport, G. W. (1937). *Personality: A psychological interpretation*. Henry Holt.

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*(1), 91–97. https://doi.org/10.1901/jaba.1968.1-91

Barlow, D. H., Nock, M. K., & Hersen, M. (2009). *Single case experimental designs: Strategies for studying behavior change* (3rd ed.). Pearson.

Ebbinghaus, H. (1913). *Memory: A contribution to experimental psychology* (H. A. Ruger & C. E. Bussenius, Trans.). Teachers College, Columbia University. (Original work published 1885)

Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences, 115*(27), E6106–E6115. https://doi.org/10.1073/pnas.1711978115

Guyatt, G., Sackett, D., Taylor, D. W., Chong, J., Roberts, R., & Pugsley, S. (1986). Determining optimal therapy: Randomized trials in individual patients. *New England Journal of Medicine, 314*(14), 889–892. https://doi.org/10.1056/NEJM198604033141406

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children, 71*(2), 165–179. https://doi.org/10.1177/001440290507100203

Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings* (2nd ed.). Oxford University Press.

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education, 34*(1), 26–38. https://doi.org/10.1177/0741932512452794

Kratochwill, T. R., & Levin, J. R. (2010). Enhancing the scientific credibility of single-case intervention research: Randomization to the rescue. *Psychological Methods, 15*(2), 124–144. https://doi.org/10.1037/a0017736

Lillie, E. O., Patay, B., Diamant, J., Issell, B., Topol, E. J., & Schork, N. J. (2011). The n-of-1 clinical trial: The ultimate strategy for individualizing medicine? *Personalized Medicine*, *8*(2), 161–173. https://doi.org/10.2217/pme.11.7

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement: Interdisciplinary Research and Perspectives*, *2*(4), 201–218. https://doi.org/10.1207/s15366359mea0204_1

Normand, M. P. (2016). Less is more: Psychologists can learn more by studying fewer people. *Frontiers in Psychology, 7*, Article 934. https://doi.org/10.3389/fpsyg.2016.00934

Parker, R. I., Vannest, K. J., & Davis, J. L. (2011). Effect size in single-case research: A review of nine nonoverlap techniques. *Behavior Modification*, *35*(4), 303–322. https://doi.org/10.1177/0145445511399147

Shadish, W. R., Hedges, L. V., & Pustejovsky, J. E. (2014). Analysis and meta-analysis of single-case designs with a standardized mean difference statistic: A primer and applications. *Journal of School Psychology, 52*(2), 123–147. https://doi.org/10.1016/j.jsp.2013.11.005

Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. Appleton-Century.

Smith, J. D. (2012). Single-case experimental designs: A systematic review of published research and current standards. *Psychological Methods*, *17*(4), 510–550. https://doi.org/10.1037/a0029312