Single Case Experimental Designs:

Introduction to a Special Issue of Neuropsychological Rehabilitation

Jonathan J. Evans[1], David L. Gast[2], Michael Perdices[3,4] and Rumen Manolov[5,6]

[1]Institute of Health and Wellbeing, University of Glasgow, Scotland, UK

[2]Communication Sciences and Special Education, University of Georgia, USA

[3]Department of Neurology, Royal North Shore Hospital, Australia

[4]Discipline of Psychiatry, Sydney Medical School, University of Sydney, Australia

[5]Department of Behavioural Sciences Methods, University of Barcelona, Spain

[6]ESADE Business School, Ramon Llull University, Spain

Running head: Introduction to SCED Special Issue

Correspondence:

Professor Jonathan Evans
Professor of Applied Neuropsychology
University of Glasgow
Institute of Health and Wellbeing
The Academic Centre
Gartnavel Royal Hospital
1055 Great Western Road
Glasgow
G12 0XH

Tel: +44 (0) 141 211 0694
Fax: +44 (0) 141 211 0356
Email: jonathan.evans@glasgow.ac.uk

ABSTRACT

This paper introduces the Special Issue of Neuropsychological Rehabilitation on Single Case Experimental Design (SCED) methodology. SCED studies have a long history of use in evaluating behavioural and psychological interventions, but in recent years there has been a resurgence of interest in SCED methodology, driven in part by the development of standards for conducting and reporting SCED studies. Although there is consensus on some aspects of SCED methodology, the question of how SCED data should be analysed remains unresolved. This Special Issues includes two papers discussing aspects of conducting SCED studies, five papers illustrating use of SCED methodology in clinical practice, and nine papers that present different methods of SCED data analysis. A final Discussion paper summarises points of agreement, highlights areas where further clarity is needed, and ends with a set of resources that will assist researchers conduct and analyse SCED studies.

INTRODUCTION

Single case experimental design studies (SCEDs) have a long history of use in evaluating interventions in education and psychology (e.g. Skinner, 1938). They provide a detailed examination of the impact of programmed interventions on behaviour over time (Barlow, Nock, & Hersen, 2009), and enable decision making regarding the nature and timing of interventions to be guided by ongoing behavioural responses (Edgington, 1983). Many people have advocated for use of SCEDs in a variety of fields (e.g., Galassi & Gersh, 1993; Gedo, 2000; Rabin, 1981; Wilson, 2011) and they have been carried out and published in a wide range of journals in the fields of special education, behavioural therapy, and neuropsychological rehabilitation (see Shadish & Sullivan, 2011, and Smith, 2012). In some fields, including education and intellectual disabilities, SCED studies have continued to influence practice, but in many fields including rehabilitation it could perhaps be said that for a period of time their perceived value as a source of evidence for the effectiveness of clinical interventions was diminished. The emergence of the evidence-based-practice movement placed a strong focus on group studies, and in particular randomised controlled trials (RCTs), systematic reviews and meta-analyses of RCTs as the primary source of evidence. In most systems for classifying levels of evidence, if they were considered at all, single case experimental design studies were classified along with all other forms of single case report as low level evidence, and therefore have had little influence on the development of clinical guidelines and practice standards. However, it appears that recently there has been a resurgence of interest in SCEDs (Smith, 2012; Tate et al., 2013), in part reflected in recent research funding initiatives (Institute of Education Sciences, 2013).

Before discussing the drivers of this resurgence of SCEDs, a brief note on terminology is required. We are using the term SCEDs to describe studies in which one (or more) participant (or unit, which could for example be a ward or a classroom) is studied in an experiment in which the participant(s) acts as their own control. In SCEDs, repeated measurement of outcome variables is made within different phases, typically involving baseline phase(s) and intervention phase(s). The aim is to determine whether the intervention affects the outcome variables of interest. There are several variations of experimental design that come under the SCED heading (e.g., reversal designs such as ABA or ABAB design and their variations, multiple baseline and multiple probe designs, alternating treatments design, adapted alternating treatments design, parallel treatments design, changing criterion design) and many

different names have been used to describe this type of study including N-of-1 Trials (see Shadish and Sullivan, 2011 for a list of other names used). A challenge for describing or naming this class of investigation is that studies often include more than one participant! Indeed methodologically rigorous studies will involve replication across three or more participants. However, the reason for the using the terms 'single-case', single subject, or 'N-of-1' is that in SCED studies participants serve as their own control, with repeated measurement of the dependent variable(s) across different phases for each participant. Tate et al. (2013) provide a useful taxonomy of common designs using a single participant, only some of which meet the criteria for being described as SCEDs. For example case descriptions, pre-post designs (where a single measurement is taken before an intervention and a single measurement is taken during or after an intervention) are not SCEDs. Perhaps more controversially they also classify AB designs (where multiple measurements are taken during a baseline and then during an intervention phase) as not being SCEDs, as others have done previously (e.g. Barlow, Nock, & Hersen, 2009; Byiers, Reichle, & Symons FJ, 2012). The reason for this is that simple AB designs lack sufficient experimental control to be able to determine whether the intervention is responsible for any observable change in the dependent variable.

What then are the drivers for the renewed interest in SCED metholology in rehabiliation? Tate et al. (2013) highlight that the Oxford Centre for Evidence-Based Medicine (www.cebm.net) now rank the randomised *n*-of-1 trial as Level 1 evidence for treatment decision purposes in individual patients, alongside systematic reviews of RCTs (Howick et al., 2011). Another driver has been the development of quality assessment tools and reporting guidelines, aimed at improving the methodological quality, and consistency in reporting, of SCEDs. Like all areas of science, there has been a wide range in methodological rigor of published SCED research reports, but consensus on criteria for judging the quality of a SCED study has not been attained. For group studies, and RCTs in particular, there are well established quality criteria and reporting standards (such as the CONSORT statement, Schulz, Altman, Moher, and the CONSORT Group, 2010), but as Smith (2012) notes, for SCEDs there is less consensus on what makes a methodologically rigorous SCED, although SCED guidelines have been proposed (Gast, 2010; Horner et al., 2005). Nevertheless in recent years several different sets of methodological standards have been published. Smith (2012) identifies six sets of standards, noting that the Single Case Experimental Design Scale of Tate et al. (2008), was 'perhaps the only psychometrically validated tool for assessing the

rigor of SCED methodology' (p. 512). The Tate et al., (2008) scale has also recently been expanded and updated to incorporate currently accepted evidence standards (e.g., Kratochwill et al., 2013) and is now known as the Risk of Bias in N-of-1 Trials (RoBiNT) scale (Tate et al., 2013). Partly based on this scale, reporting standards have been developed and used as the basis for a set of reporting guidelines referred to as the Single Case Reporting in BEhavioural interventions (SCRIBE; Tate et al., this issue). The combination of a recognition of the level of evidence provided by SCEDs and an emerging consensus on what constitutes methodologically strong SCEDs perhaps explains why Smith (2012) views SCEDs as 'poised for resurgence' (p. 510).

A third driver of the SCED revival in rehabilitation research is the development of methods of analysis suitable for SCED data. How to analyse SCED data has, and continues to be, a much-debated topic. In group design studies, such as RCTs, the methods are well established and there is reasonable consensus on the rules regarding what methods should be applied to what study designs and under what circumstances. For SCEDs this is not the case, and as Smith (2012) notes in his review of published research and methodological standards, 'analytic method emerged as an area of discord' (p. 510). The debate has included the question of *whether* SCED data should be analysed using any statistics at all - one view of SCED data is that if you can't see the effect of the intervention from a graph of the data then the intervention has not had sufficient impact to be of any practical use, i.e., social validity, or clinical significance. Others have argued that in addition to the use of formal visual analysis methods (see Lane and Gast, this issue, for an account of visual analysis methods), it is necessary to include some form of statistical analysis. There are several reasons for this, including the argument that human observers (particularly those who are not blind to conditions and have a vested interest in the success of an intervention) are subject to bias, particularly because people are poor at seeing randomness in data (Williams & Griffiths, 2008). Thus the use of statistical analysis provides a means of systematically examining data in a quantitative way, produces standardized measures of the size of an effect and quantifies the probability that any apparent effect is due to the intervention and not just chance. The last decade has seen many new methods of analysis emerge, beginning with modifications to the frequently used 'Percent of Non-overlapping Data' in 2006 (Ma, 2006) and 2007 (Parker, Hagan-Burke, & Vannest, 2007), and continuing with the procedures discussed in the current Special Issue, all of which emerged in the period 2008-2013, with several still undergoing modification.

Perhaps the most convincing reason for the inclusion of statistical analysis in SCED studies, particularly using measures of effect size, is to allow for the possibility of combining studies to produce meta-analyses of SCEDs. Returning to the issue of levels of evidence of treatment intervention efficacy, whilst SCEDs are level 1 evidence in relation to the effectiveness of an intervention with a particular participant in a particular situation, for a clinician or educator, the value of this evidence in selecting an intervention approach for a new participant depends on the extent to which the characteristics of the participant and situation in the original SCED study are similar to those of the new participant. As SCED evaluations of particular treatment approaches are replicated over more and more participants (i.e., systematic replication), so confidence in the range of participants and settings in which the intervention is likely to be effective grows. Thus, as with group design studies, there is potential value in meta-analysis in combining information from multiple SCEDs, broadening out the range of participants and situations for which the intervention is effective (i.e., increasing external validity of an intervention).

Meta-analysis makes it possible to compare and combine studies in terms of intervention effectiveness (Jenson, Clark, Kircher, & Kristjansson, 2007), but is subject to the same challenges as the analysis of individual studies in terms of the selection of effect size measure and the integrating procedure to be used (Beretvas & Chung, 2008). Integrating outcomes of individual studies necessarily provides an average effect in which individual differences are lost. That is, the general information about effectiveness masks information on whom the intervention was more or less effective than the average. A possible solution may be including moderator variables that might account for the variability in outcomes as in multilevel models (Van den Noortgate & Onghena, 2008).

PURPOSE OF THE SPECIAL ISSUE

The idea for this Special Issue arose at a meeting being held to discuss Single Case Reporting guidelines In BEhavioural interventions (the SCRIBE project, Tate et al., this issue) in Sydney in December 2011. Participants at that meeting (all experts in SCED methodology) agreed that whilst recent developments in setting standards for the conduct and reporting of SCED studies are stimulating a resurgence of interest in SCED methodology, there remains a

number of issues yet to be resolved, the most challenging and controversial of which is the question of how SCED data should be analyzed. In putting together this Special Issue, our aim was to encourage submission of papers that addressed aspects of the conduct and analysis of SCED studies, as well as examples of application of SCED methodology in clinical studies. The studies illustrating the use of SCED methodology are focused on neuropsychological rehabilitation interventions, reflecting the focus on the journal that is hosting these papers, though the issues that arise from the studies relating to implementing SCED methodology are widely applicable. Furthermore, the papers that discuss conduct or analysis are clearly not field-specific and are applicable to a great variety of behavioural areas.

The specific aim of papers focused on analytical techniques is to present the myriad of existing possibilities, with each author highlighting the rationale, the application, and the strengths and limitations of each method. Readers are invited to assess for themselves the degree to which each technique is applicable in their area of practice and research, although some tentative criteria for selecting a method of analysis are presented in the editorial discussion paper that closes this Special Issue. Given that the debate about how best to analyse SCED data is ongoing, we aim to stimulate further research and discussion on how to improve how SCED studies are conducted and findings analysed. We do not claim to identify a single optimal solution as yet.

CONTENT OF THE SPECIAL ISSUE

The Special Issue includes 16 contributions, which can be grouped into three categories: 1) papers relating to the conduct of SCED studies; 2) original applied investigations using SCED methodology; 3) papers that describe and discuss techniques for analysing SCED data. The two initial articles by Tate and colleagues and by Ledford and Gast offer general guidelines about how a SCED study should be carried out and reported, highlighting different aspects of studies that should be taken into account in order for the study to provide solid scientific evidence on the effectiveness of the intervention being tested.

The papers presenting examples of the application of SCED methodology are not artificial or idealised examples of SCED studies. They are real clinical studies that reflect many of the challenges of applying SCED methodology in clinical practice. Winkens et al.'s study,

focuses on aggressive behaviour in acquired brain injury, using a simple AB design. O'Neill and Findlay's study also looks at aggressive challenging behaviour (and emotional dysregulation, considered to be a contributing factor) and also uses an AB design, replicating it across two participants. Both papers include randomisation. Tunnard and Wilson present an investigation of treatment interventions for unilateral neglect using a more complex design structure in which each different intervention (of which there are five) is preceded by a baseline phase, making the design ABACADAEAF. The last two contributions, from Svanberg and Evans and from Jamieson, Cullen, McGee-Lennon, Brewster, and Evans deal with applications of the SCED methodology in relation to memory impairments. Whilst Svanberg and Evans present a single patient, Jamieson and colleagues present a meta-analysis of studies of technology for memory rehabilitation. The paper meta-analyses both group and SCED studies. Related to the issue of combining data, amongst the papers on analytical techniques, the last two contributions (by Baek et al. and by Solmi and Onghena) also deal with integrating the results across SCED studies.

The order of the papers describing methods analysing single-case data is related to their content and the relationships between the procedures. The first of these papers, by Lane and Gast, focusses on visual analysis, which is arguably the most straightforward analytical approach, and the usual starting point for analysis of SCED data. The subsequent contribution by Brossart and colleagues presents one of the simplest types of quantification, which is based on one of the data features assessed in visual analysis, namely, the amount of overlap between the data pertaining to adjacent phases. The two articles that follow both present techniques that entail computing a statistic or an effect size index for the original data set and comparing it with the value of the same measure for a set of hypothetical data: Borckardt and Nash describe Simulation Modelling Analysis in which the hypothetical data are generated according to the features of the data at hand and the measure used is the point-biserial correlation, whereas Heyvaert and Onghena discuss randomisation tests in which the hypothetical data are created by placing markers of phase change at all possible points, with the specific statistic used being chosen by the researcher (e.g., a *t* statistic, a nonoverlap index, or a standardized mean difference). The next contribution, from Shadish and colleagues, is focused on a new way of estimating the standardized mean difference (i.e., a *d* statistic) for single-case data, with the intention of making it comparable to the *d* statistic that researchers are familiar with for group-designs. A standardized difference between conditions similar to the *d* statistic can also be obtained following a regression analysis of the SCED

data, an approach described by Swaminathan and colleagues. Swaminathan et al.'s procedure is related to the contributions from Rindskopf and from Baek and colleagues, in that they are all based on a similar, albeit not identical, method of modelling SCED data features such as level, trend, and autocorrelation. Rindskopf uses the multilevel models described in Baek et al. to illustrate how Bayesian statistics can be useful in the SCED context. As noted above, the final two analytical papers deal with data synthesis, either via hierarchical linear models (in the paper by Ferron and colleagues) or combining the probabilities associated with statistical tests such as randomisation tests, as presented by Solmi and Onghena.

Following the analysis papers, we provide a paper reflecting on the issues arising from the collection of papers in this Special Issue and the wider SCED literature. There are now many points of consensus with regard to what makes a methodologically rigorous evaluation SCED research. But there are still issues to be addressed, which we will highlight and discuss. Finally, our aim with this Special Issue is to encourage clinicians and researchers to become confident in their use of SCED methodology, and so we propose, in our closing discussion paper, criteria to support the selection of a method of analysis, and end with a set of resources that are available to assist with analysing SCED study data.

REFERENCES

Baek, E., Moeyaert, M., Petit-Bois, M., Beretvas, S. N., Van de Noortgate, W., & Ferron, J. (2014). The use of multilevel analysis for integrating single-case experimental design results within a study and across studies. *Neuropsychological Rehabilitation, 24,* XXX-XXX.

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

Beretvas, S. N., & Chung, H. (2008). A review of meta-analyses of single-subject experimental designs: Methodological issues and practice. *Evidence-Based Communication Assessment and Intervention, 2*, 129-141.

Borckardt, J., & Nash, M. (2014). Simulation modelling analysis for small sets of single-subject data collected over time. *Neuropsychological Rehabilitation, 24,* XXX-XXX.

Brossart, D. F., Vannest, K., Davis, J., & Patience, M. (2014). Incorporating nonoverlap indices with visual analysis for quantifying intervention effectiveness in single-case experimental designs. *Neuropsychological Rehabilitation, 24,* XXX-XXX.

Byiers B.J., Reichle J. & Symons F.J. (2012). Single-Subject Experimental Design for Evidence-Based Practice. American Journal of Speech-Language Pathology, 21, 397–414.

Edgington, E.S. (1983). Response-guided experimentation. *Contemporary Psychology*, *28*, 64-65.

Galassi, J. P., & Gersh, T. L. (1993). Myths, misconceptions, and missed opportunity: Single-case designs and counseling psychology. *Journal of Counseling Psychology*, *40*, 525-531.

Gast, D. L. (Ed.). (2010). *Single subject research methodology in behavioral sciences*. London, UK: Routledge.

Gedo, P. M. (2000). Single case studies in psychotherapy research. *Psychoanalytic Psychology*, *16*, 274-280.

Heyvaert, M., & Onghena, P. (2014). Analysis of single-case data: Randomisation tests for measures of effect size. *Neuropsychological Rehabilitation, 24,* XXX-XXX.

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*, 165-179.

Howick, J., Chalmers, I., Glasziou, P., Greenhaigh, T., Heneghan, C., Liberati, A., et al. (2011). The 2011 Oxford CEBM Evidence Table (Introductory Document). *Oxford Centre for Evidence-Based Medicine*. http://www.cebm.net/index.aspx?o=5653

Institute of Education Sciences. (2013). *Request for applications: Statistical and research methodology in education*. Retrieved from http://ies.ed.gov/funding/pdf/2014_84305D.pdf

Jamieson, M., Cullen, B., McGee-Lennon, M., Brewster, S., & Evans, J. J. (2014). The efficacy of cognitive prosthetic technology for people with memory impairments: A systematic review and meta-analysis. *Neuropsychological Rehabilitation, 24,* XXX-XXX.

Jenson, W. R., Clark, E., Kircher, J. C., & Kristjansson, S. D. (2007). Statistical reform: Evidence-based practice, meta-analyses, and single subject designs. *Psychology in the Schools, 44,* 483-493.

Lane, J., & Gast, D. L. (2014). Visual analysis in single case experimental design studies: Brief review and guidelines. *Neuropsychological Rehabilitation, 24,* XXX-XXX.

Ledford, J., & Gast, D. L. (2014). Measuring procedural fidelity in behavioural research. *Neuropsychological Rehabilitation, 24,* XXX-XXX.

Ma, H. H. (2006). An alternative method for quantitative synthesis of single-subject research: Percentage of data points exceeding the median. *Behavior Modification, 30,* 598-617.

O'Neill, B., & Findlay, G. (2014). Single case experimental designs in neurobehavioural rehabilitation: Preliminary findings on biofeedback in the treatment of challenging behaviour. *Neuropsychological Rehabilitation, 24,* XXX-XXX.

Parker, R. I., Hagan-Burke, S., & Vannest, K. J. (2007). Percentage of all non-overlapping data: An alternative to PND. *Journal of Special Education, 40,* 194-204.

Rabin, C. (1981). The single-case design in family therapy evaluation research. *Family Process, 20,* 351-366.

Rindskopf, D. (2014). Bayesian analysis of data from single case designs. *Neuropsychological Rehabilitation, 24,* XXX-XXX.

Schulz K. F., Altman D. G., Moher D., and the CONSORT Group. (2010). CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMC Medicine, 8*:18.

Shadish, W. R., Hedges, L. V., Pustejovsky, J. E., Boyajian, J. G., Sullivan, K. J., Andrade, A, & Barrientos, J. L. (2014). A d-statistic for single-case designs that is equivalent to the usual between-groups d-statistic. *Neuropsychological Rehabilitation, 24,* XXX-XXX.

Shadish, W. R., & Sullivan, K. J. (2011). Characteristics of single-case designs used to assess intervention effects in 2008. *Behavior Research Methods, 43,* 971-980.

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

Smith, J. D. (2012). Single-case experimental designs: A systematic review of published research and current standards. *Psychological Methods, 17,* 510-550.

Solmi, F., & Onghena, P. (2014). Combining *p*-values in replicated single-case experiments with multivariate outcome. *Neuropsychological Rehabilitation, 24,* XXX-XXX.

Svanberg, J., & Evans, J. J. (2014). Impact of SenseCam on memory, identity and mood in Korsakoff's syndrome: A single case experimental design study. *Neuropsychological Rehabilitation, 24,* XXX-XXX.

Swaminathan, H., Rogers, H. J., Horner, R., Sugai, G., & Smolkowski, K. (2014). Regression models for the analysis of single case designs. *Neuropsychological Rehabilitation, 24,* XXX-XXX.

Tate, R. L., McDonald, S., Perdices, M., Togher, L., Schultz, R., & Savage, S. (2008). Rating the methodological quality of single subject designs and n-of-1 trials: Introducing the Single-Case Experimental Design (SCED) Scale. *Neuropsychological Rehabilitation, 18,* 385-401.

Tate, R. L., Perdices, M., McDonald, S., Togher, L., & Rosenkoetter, U. (2014). The conduct and report of single-case research: Strategies to improve the quality of the neurorehabilitation literature. *Neuropsychological Rehabilitation, 24,* XXX-XXX.

Tate, R. L., Perdices, M., Rosenkoetter, U., Wakima, D., Godbee, K., Togher, L., & McDonald, S. (2013). Revision of a method quality rating scale for single-case experimental designs and n-of-1 trials: The 15-item Risk of Bias in N-of-1 Trials (RoBiNT) Scale. *Neuropsychological Rehabilitation, 23,* 619-638.

Tunnard, C., & Wilson, B. (2014). Comparison of neuropsychological rehabilitation techniques for unilateral neglect: An ABACADAEAF single-case experimental design. *Neuropsychological Rehabilitation, 24,* XXX-XXX.

Van den Noortgate, W., & Onghena, P. (2008). A multilevel meta-analysis of single-subject experimental design studies. *Evidence-Based Communication Assessment and Intervention, 2,* 142-151.

Williams, J. J., & Griffiths, T. L. (2008). *Why are people bad at detecting randomness? Because it is hard*. Proceedings of the 30th Annual Conference of the Cognitive Science Society. Retrieved from http://cocosci.berkeley.edu/tom/papers/hard.pdf

Wilson, B. A. (2011) Cutting edge developments in neuropsychological rehabilitation and possible future directions. *Brain Impairment, 12,* 33-42.

Winkens, I., Ponds, R., Pouwels-van den Nieuwenhof, C., Eilander, H., & van Heugten, C. (2014). Using single-case experimental design methodology to evaluate the effects of the

ABC method for nursing staff on verbal aggressive behaviour after acquired brain injury. *Neuropsychological Rehabilitation, 24,* XXX-XXX.