Journal of Communication Disorders 67 (2017) 1–13

Contents lists available at ScienceDirect

Journal of Communication Disorders

journal homepage: www.elsevier.com/locate/jcomdis

Demonstrating treatment efficacy using the single subject randomization design: A tutorial and demonstration

Susan Rvachew*, Tanya Matthews

School of Communication Sciences and Disorders, McGill University, Centre de recherche interdisciplinaire en réadaptation du Montréal métropolitain, Canada

ARTICLE INFO

Keywords:
Single case research
Single subject randomization design
Evidence-based practice
Treatment efficacy
Speech sound disorders
Childhood apraxia of speech
Tutorial

ABSTRACT

Single case research refers to the broader category of research designs in which each case serves as his or her own control. A single subject randomization design is a specific form in which sessions are randomly allocated to treatment conditions within subjects. Two tutorials on the application of these designs are provided. In the single subject randomized phase design, baseline probes are administered repeatedly during a no-treatment or usual care phase; subsequently probes are administered repeatedly during the treatment phase of the experiment; the starting point for the treatment phase is determined by random selection. In the single subject randomized alternation design, any session can be randomly allocated to any treatment condition. In either case, the test statistic can be the mean of probe performance during the treatment sessions after subtracting the baseline mean. The significance of the obtained test statistic is determined by resampling test. Specifically, the obtained test statistic is interpreted relative to a distribution of test statistics generated by all possible random allocations. This distribution yields a P value which represents the probability of obtaining a test statistic as large as that obtained by the selected allocation. In addition to the tutorials, two experiments using these designs with a single 8-year-old participant with Childhood Apraxia of Speech are presented to demonstrate the utility of these designs and the application of the associated statistical analysis procedures.

1. Introduction

Interventions for young children with communication disorders are expected to conform to the principles of evidence-based practice following the definition provided by Sackett, Rosenberg, Muir, Haynes, and Richardson (1996). This definition focused attention on the "conscientious, explicit, and judicious use of current best evidence" when deciding when and how to treat individuals requiring speech and language therapy. The guidelines for identifying *best evidence* (i.e., American Speech-Language & Hearing Association, 2004) emphasize the importance of randomization as a key characteristic of a well-designed intervention study; randomization of participants to treatment conditions controls for many possible threats to internal validity especially history, maturation, and selection effects (Altman & Bland, 1999). In contrast, if a researcher provides an experimental intervention to a group of participants and then compares their outcomes to a comparison group selected without randomization, for example a convenience sample of participants receiving "usual care" in the community, it is difficult to conclude that experimental group outcomes were caused by the experimental treatment. Any differences in outcomes between the group that received the

Abbreviations: SCR, single case research; SSRD, single subject randomization design; CAS, childhood apraxia of speech

* Corresponding author at: School of Communication Sciences and Disorders, McGill University, 1266 Pine Avenue West, Montréal, Québec, Canada.
E-mail address: susan.rvachew@mcgill.ca (S. Rvachew).

http://dx.doi.org/10.1016/j.jcomdis.2017.04.003
Received 2 August 2016; Received in revised form 23 March 2017; Accepted 10 April 2017
Available online 22 April 2017

experimental treatment and the group that received usual care might be attributed to other differences between those groups: greater time resources in the families that chose to participate in the experiment (selection effect), naturally slower rate of progress by the participants who received usual care (maturation effect), or higher uptake of a co-occurring treatment program by the experimental group (history effect). Given these threats to internal validity, researchers, practitioners or policy makers will find it difficult to determine whether the new intervention should be adopted. Some decision makers might discount the threats to the validity of the research when there is a large advantage to the experimental group, but in fact large effects can occur due to biases of the sort described above and also due to chance.

The advantage of a randomized control design is not that such differences between subjects and groups are eliminated via randomization but that they are not *systematically* related to the independent or dependent variable; furthermore, the probability that the outcome is due to chance differences in the distribution of such variables across the experimental and control groups can be estimated (Greenland & Morgenstern, 2001). Attempts may be made to control for confounding variables in nonexperimental designs via careful selection and matching of participants at intake or post-study analytical strategies. However, it is well established that these factors are often impossible to identify, their effects are frequently large and unpredictable, and known control strategies may be insufficient when applied (Valentine & Thompson, 2013). Therefore, randomization is the best method for controlling confounds, and ensuring the internal validity of a study designed to establish the causal relationship between a treatment and its purported benefits.

Notwithstanding these known benefits of a randomized control trial, single case research (SCR) methods are often used to investigate the efficacy of interventions for communicative disorders (McReynolds & Kearns, 1983; Murray, McCabe, & Ballard, 2014; Sommers, Logsdon, & Wright, 1992; Thompson, 2006). For example, SCR methods have been used to evaluate the efficacy of interventions to improve articulation accuracy (Skelton, 2004), productive phonological knowledge (Rudolph & Wendt, 2014), and vocalization frequency (Hulme, Bain, Hardin, McKinnon, & Waldron, 1989) in children with primary and secondary speech sound disorders. The primary advantage of SCR is that of feasibility given the much reduced resource requirements relative to a randomized controlled trial, particularly when targeting participant populations with rare disorders (Byiers, Reichle, & Symons, 2012). A single subject design is also ideal when the behaviors to be targeted by the intervention are idiosyncratic and must be selected individually for each client, as is often the case for speech error patterns. A single subject design is also recommended when the target population is extremely heterogeneous and children with speech sound disorders meet this criteria at phenotype, endophenotype and genotype levels of description.

SCR methods are also favoured for their external validity, given that these studies can be conducted in the clinical setting to answer questions of direct clinical relevance (Kearns, 1986; Kratochwill & Levin, 2010). Proponents of SCR also point out that randomized control trials describe the response of an “average client” to a manualized treatment protocol, a situation far removed from normal clinical decision making. Kazdin (1983) suggested that it can be difficult to apply this kind of information to an individual who requires a nonstandardized therapeutic solution to their unique multifaceted problem (see also Bernstein Ratner, 2006). A particular challenge is the difficulty of matching the unique characteristics of the client presenting at a given time for treatment to the characteristics of the subgroup of individuals that actually benefited from the intervention tested in the trial (McReynolds & Thompson, 1986). Subgroup effects are commonly reported in the literature. For example, the core vocabulary approach is reported to be an effective intervention to improve speech intelligibility, but only for those children with “inconsistent phonological disorder” (Crosbie, Holm, & Dodd, 2005). Yoder, Camarata, and Gardner (2005) assessed the impact of broad target recasts on speech intelligibility; although a main effect of treatment versus control was not observed, the authors suggested the intervention was effective for children who began treatment with the most unintelligible speech. SCR may be well suited for exploring the relationship between specific client characteristics and response to certain treatment procedures or combinations of procedures.

Given these compelling arguments in favor of SCR, tutorials demonstrating the use and reporting of these methods as well as innovative strategies for interpreting the results of single case studies multiply (e.g., Byiers et al., 2012; Campbell, 2004; Horner et al., 2005; Kratochwill et al., 2013; Lanovaz & Rapp, 2016; Moeller, Dattilo, & Rusch, 2015). Although there are many variations on the details of these designs, a single case study consists simply of repeated measurements of the participant’s performance on the dependent variable across baseline and treatment phases with replication. The internal validity of the single case study rests primarily on the stability, or at least predictability, of the participant’s performance during the baseline phases of the experiment; a successful outcome is indexed by a marked change in performance co-incident with each introduction of treatment. Relatedly, there must not be performance changes in untreated behaviors or alternatively, on the treated behavior during no-treatment phases of the study. The behavior change during the treatment phase may be observed in terms of level, trend or variability in performance on the dependent variable in relation to baseline performance. Replication of the effect may occur within one subject, as in a withdrawal (ABAB) design, or across multiple subjects, as in a multiple baseline across subjects design. Alternatively, a multiple baseline across behaviors design permits replications across multiple targeted behaviors within a single subject. Horner, Swaminathan, Sugai, and Smolkowski (2012) suggest that a functional relationship between the treatment and the outcome is established when the predicted change in performance is observed at least three times in response to the treatment.

Ideally the results of a single case study will be straightforward and the treatment effect can be discerned through visual inspection of the data points, especially when using interpretative criteria such as the conservative dual criterion method to reduce type I and type II error (Fisher, Kelley & Lomas, 2003). However, many strategies for statistical analysis of single case experiments have been proposed (Campbell, 2004; Parker, Vannest, & Davis, 2011). These statistics have been developed largely to facilitate *meta*-analyses; that is, the statistics allow for the calculation of effect sizes that can be combined across experiments to improve estimation of the actual treatment effect. The ability to conduct *meta*-analyses of SCR is important because *n*-of-1 trials are now accepted as a high level of evidence by the Oxford Centre for Evidence Based Medicine (2011); however, SCR is acceptable as evidence of a