Australasian Journal of Special and Inclusive Education (2020), **44**, 1–14
doi:10.1017/jsi.2020.4

CAMBRIDGE UNIVERSITY PRESS

ORIGINAL ARTICLE

Using a Systematic Review to Explore Single-Subject Design Use Within Special Education Over Time†

Jessica Zoe Zanuttini

The University of Sydney, Australia
Corresponding author. Email: jessica.zanuttini@sydney.edu.au

(Received 20 October 2019; revised 26 February 2020; accepted 27 February 2020)

Abstract

While teachers and researchers of students with disability are commonly interested in individual students and their progress towards individualised goals, traditional approaches to educational research that aim to illustrate generalisation between cause and effect are still commonly used. Traditional approaches, such as group-comparison designs, are used to demonstrate improved performance for the average student; however, they also obscure individual student data. Within special and inclusive education, single-subject designs may provide more appropriate conclusions to particular types of questions than traditional research methods as they allow for the examination of a functional relationship between the dependent and independent variables and rely on the participant serving as their own control. This article provides a summary of the use of single-subject designs within the field of special education over time and the findings of a systematic review conducted according to the PRISMA statement are presented. Through this systematic review, published articles from 7 prominent special education journals that involved intervention for school-aged students with disability were examined. In addition, a sample of exemplar studies that have employed single-subject designs are discussed.

Keywords: single-subject design; systematic review; special and inclusive education; experimental design; PRISMA

Group designs and single-subject designs are the two major types of experimental research designs used to examine whether a change in an independent variable causes a change in the dependent variable within the field of special education (Cook & Cook, 2016). Group experimental designs rely on the scientific method that is largely quantitative, compares behaviours between a group receiving treatment and another that is not, and predicts future data (Cook, Cook, Landrum, & Tankersley, 2008; Forbes, Ross, & Chesser, 2011). Although group designs with large sample sizes and randomly selected participants may produce statistically significant findings to indicate effective practices for a typical student of a particular age or stage, these findings cannot be simply applied to individual students in specific contexts (Forbes et al., 2011). Such research provides little value to special educators who are commonly interested in selecting pedagogical approaches to meet individual students' needs and measuring their progress towards individualised goals.

Single-subject designs, as an alternative to group designs, are regularly associated with the field of special education and tend to provide more appropriate conclusions than traditional research approaches within education (Cakiroglu, 2012; Mastropieri et al., 2009). This is because special education is considered a problem-solving discipline that focuses on the unique characteristics of individual students from low-incidence populations (Horner et al., 2005). The primary aim of research

†This manuscript was accepted under the Editorship of Michael Arthur-Kelly.

© The Author(s) 2020

https://doi.org/10.1017/jsi.2020.4 Published online by Cambridge University Press

within this field is to regularly measure the effects of active intervention on individual students' behaviours (i.e., increase of desirable behaviour or decrease of undesirable behaviour) within authentic settings (Horner et al., 2005). Single-subject designs not only narrow the focus onto an individual and regularly measure student performance, but also highlight the role of functional interventions and practicality of solutions, with an emphasis on socially valid goals.

History of Single-Subject Designs

By the early 20th century, the intense investigation of individuals was an accepted research method within the fields of psychology and physiology (Kazdin, 2011). Pavlov and Thorndike famously used few participants in some of their most notable investigations (Kazdin, 2011). However, by the 1930s, small-scale research studies were less frequently published in journals, and the publication of larger studies involving statistical analyses became more common (Kazdin, 2011). Group-comparison designs gained popularity in the 1940s but were soon criticised for their limitations, including the variability of participants involved in research, the costs (e.g., time, financial), and ethical issues of withholding treatment (Hacker, 1980).

The single-subject methodology was first formally operationalised in the 1960s and has since been considered a valid experimental approach across the disciplines of education, psychology, social work, and nursing (Alnahdi, 2015; Cohen, Manion, & Morrison, 2000; Forbes et al., 2011; Kratochwill & Levin, 1992; Sidman, 1960). Before being employed by special education professionals, single-subject designs were typically used by medical practitioners who were the primary service providers for children and young people with disabilities (Odom et al., 2005).

Single-subject research designs became the primary research methodology for applied behaviour analysis research in the fields of education, occupational therapy, and psychology (Cooper, Heron, & Heward, 2007). Within applied behaviour analysis research, the primary aim has been to demonstrate a repeated effect of an intervention on socially important behaviours (Cooper et al., 2007). Over the past 50 years, however, the use of single-subject designs has evolved. Although single-subject research continues to explore the effects of the independent variable on the dependent variable, it is now more commonly used to help identify evidence-based practices (Horner et al., 2005).

Along with the reasons for using single-subject designs, the ways single-subject designs are conducted have been refined through the introduction of single-subject designs standards, such as a preferred set of standards published by the What Works Clearinghouse (Kratochwill et al., 2010; Smith, 2012; Wolery, Dunlap, & Ledford, 2011). Such standards clearly define how quality indicators can be maintained regardless of the uniqueness and variability of single-subject experimental designs. These standards also provide researchers with standardised criteria for the design, implementation, and evaluation of research involving single-subject methodology (Smith, 2012).

Given that the single-subject design standards can be applied to the design of single-subject research within any field or discipline, an increased range of educational research now employs single-subject designs (Mastropieri et al., 2009). Such diversity suggests a need for systematic reviews that produce synthesised reports of their use within specific fields, disciplines, or specific areas of inquiry. Even within the field of special and inclusive education, there is variation in the use of single-subject designs.

Particularly within the last 35 years, the publication of studies involving single-subject designs within special education journals has become more common (Hammond & Gast, 2010). Moeyaert, Ugille, Ferron, Beretvas, and Van den Noortgate (2013) found the most frequently used type of research design, across these published studies, was single-subject multiple-baseline design. A number of reviews have systematically investigated the use of single-subject designs within the broad field of special and inclusive education. For example, Mastropieri et al. (2009) conducted an extensive study by analysing articles published between 1988 and 2006 in 11 prominent journals within the field of special education. As part of this review, articles classified as intervention studies conducted within preschool to Grade 12 underwent a research design coding process. All studies were coded as one group only pre–post research, experimental or quasi-experimental group designs, or single-subject research.

https://doi.org/10.1017/jsi.2020.4 Published online by Cambridge University Press

Analysis of coding across all journals revealed that more single-subject research designs were published than any other research design (i.e., 50.5%; Mastropieri et al., 2009).

Hammond and Gast (2010) conducted a descriptive analysis of single-subject research designs in articles published across similar years. Their analysis reviewed the types of single-subject research designs used to answer applied research questions published between 1983 and 2007 in eight special education peer-reviewed journals. They selected 196 journal issues for review, containing 1,936 articles, in which 456 articles employed at least one single-subject design. Results revealed that multiple-baseline designs were published more frequently than withdrawal/reversal designs. Of the multiple-baseline designs, replication of effects was most commonly conducted across participants/participant groups (Hammond & Gast, 2010).

More recently, Alnahdi (2015) reviewed articles published from January 2000 to February 2012 that involved the use of any single-subject design within a special education context. Alnahdi's review evaluated the use of single-subject designs within special education according to their popularity, reported strengths, and common concerns. Conclusions suggested that single-subject designs have been one of the most commonly used research designs within special education over the past 30 years, with trends suggesting a continuation of their use in the future (Alnahdi, 2015). The main reported strengths of the designs were their cost effectiveness and flexibility, while external validity and generalisability of effect were among the common concerns (Alnahdi, 2015).

Other reviews of single-subject designs within the field of special and inclusive education have concentrated on more specific areas of inquiry, such as positive behaviour support and interventions for students with autism spectrum disorder. For example, Conroy, Dunlap, Clarke, and Alter (2005) systematically reviewed positive behaviour interventions for young children presenting challenging behaviour. In their study, they reviewed articles across 23 peer-reviewed journals published between 1984 and 2003. Review of 63 intervention studies revealed that 62 employed single-subject designs, while the remaining one used a group design (Conroy et al., 2005). Further, Wong et al. (2015) explored the use of interventions for children with autism spectrum disorder. Their review of 456 studies revealed 89% involved single-subject designs, while only 11% used group designs.

Single-Subject Designs

Single-subject designs reflect the tenets of educational psychology that emphasise each student's individualised learning needs and allow for simple methods of data collection and analysis (Forbes et al., 2011). Single-subject designs are primarily used to examine the existence of a causal, or functional, relationship between the dependent and independent variables; they rely on the participant serving as their own control across various settings, rather than using another participant or participant group (Cakiroglu, 2012; Maggin, Cook, & Cook, 2018). The demonstration of a functional relationship in single-subject design research is particularly important when assessing whether an educational intervention should be considered evidence based (Horner & Kratochwill, 2012).

While there are a number of single-subject research design types, some characteristics are common to all designs and reflect the single-case design standards. For example, single-subject research involves the use of direct observation to collect repeated measures of behaviour over time, allowing comparisons to be made between at least two conditions (e.g., baseline and intervention; Horner, Swaminathan, Sugai, & Smolkowski, 2012). A baseline condition requires behaviour to be observed repeatedly before the introduction of the intervention to determine a stable, predictable, and reliable pattern of the participant's current performance (Tankersley, Harjusola-Webb, & Landrum, 2008). Once the baseline data are stable, an intervention can be introduced, and behaviour is again measured repeatedly. Observational data from across multiple conditions are typically represented visually (e.g., line graph) and analysed for within-participant data patterns of trend, level, and variability (Tankersley et al., 2008). Behaviour analysis through within-study replication (i.e., repeated measurement across participants/participant groups, behaviours/tasks, conditions/settings) improves the data's reliability and allows researchers to explore the intervention's effect on student behaviour and determine whether

https://doi.org/10.1017/jsi.2020.4 Published online by Cambridge University Press

a functional relationship between the independent and dependent variables exists (Horner et al., 2005; Maggin et al., 2018). To illustrate a causal relationship between variables, single-subject research must also repeatedly and systematically introduce interventions across conditions. As an individual participant acts as their own control within single-subject designs, the systematic introduction of an intervention across phases demonstrates the intervention's effectiveness for that participant (Maggin et al., 2018).

Single-Subject Designs in Special and Inclusive Education

The features of single-subject research designs align with the key principles of inclusive and special education, and specifically individualised pedagogy in response to student goals and ongoing progress monitoring towards goals (Tankersley et al., 2008). McMillan (2004) identified five defining characteristics of single-subject designs that rendered them an appropriate method for research within this field. These characteristics include the use of repeated measures over a minimum of two conditions, inclusion of a baseline and treatment condition, reliable means of measurement, explicit description of conditions, and adherence to the single-variable rule. While Horner and colleagues (2005) have been particularly critical of the ways some researchers conduct single-subject designs, they also compiled a list of the benefits of single-subject designs within the field of special education: single-subject research methods are focused on the individual and do not rely on the use of populations to fulfil specific conditions, they are cost-effective, and allow for practical examination of behavioural or educational intervention data within applied contexts (Horner et al., 2005). The continued use of single-subject designs within the field of special and inclusive education is considered appropriate as the design requires relatively few participants, permits research on low-incidence populations, and allows research in authentic settings in which other designs are difficult to use (Alnahdi, 2015; Pustejovsky & Ferron, 2017).

Types of Single-Subject Designs

The choice to use one type of single-subject design over another should be determined by an investigation's research question(s) and the nature of the dependent and independent variables (Cakiroglu, 2012; Cook & Cook, 2016). Typically, single-subject research designs address questions interested in determining the performance of an individual participant in response to a condition or independent variable(s) and the existence of a functional relationship between dependent and independent variables (Horner et al., 2005). Once practitioners have selected the type of design, they must be aware of the data collection implications, strengths, limitations, and design standards for the chosen design (Cakiroglu, 2012; Horner et al., 2005). The following describes three of the most commonly used single-subject designs used within the field of special education.

Withdrawal designs (ABAB)

The most basic single-subject research design is a withdrawal design, commonly called an ABAB or reversal design, and involves two alternating baseline phases (A) and two alternating intervention phases (B). In the first phase (A1), baseline data are collected to establish a participant's typical performance prior to the intervention's implementation. The intervention is then applied during the following phase (B1) before being removed in the second baseline phase (A2). The intervention is applied again during the second intervention phase (B2). Withdrawal designs are used to determine a functional relationship, which can be demonstrated if a subject's behaviour during the second baseline phase returns to a similar level as the first baseline phase, and if behaviour during the second intervention phase returns to a similar or higher level during the second intervention phase (Hammond & Gast, 2010). Withdrawal designs are typically appropriate when investigating student behaviour and less acceptable when investigating student learning (Forbes et al., 2011). This is because

https://doi.org/10.1017/jsi.2020.4 Published online by Cambridge University Press

student learning outcomes should not be expected to return to baseline levels once learning has taken place during intervention phases. However, the researcher must be aware of the logistical and ethical issues that may arise from returning to the baseline condition, particularly if the behaviour of interest is destructive or self-injurious (Hammond & Gast, 2010).

Multiple-baseline designs

When returning to baseline is deemed problematic (i.e., targeting behaviours that are dangerous or cannot be unlearnt), a multiple-baseline design can be used. Multiple-baseline designs involve the use of a series of baseline-intervention (AB) studies across participants/participant groups, settings/conditions, or behaviours/tasks, thus achieving within-experiment replication (Maggin et al., 2018). Multiple-baseline designs use a lagged introduction of the intervention; therefore, a stable baseline is required for the initial participant, setting, or behaviour before the intervention is introduced (Maggin et al., 2018). A change in the target behaviour, according to the criteria set by the researcher, is required before the intervention is consecutively initiated across the participants, settings, or behaviours. An experimental effect must be demonstrated at three different points in time to suggest a functional relation between the dependent and independent variables (Kazdin, 2011).

Alternating treatment designs

An alternating treatment design is suitable for researchers interested in comparing the effectiveness of more than a single intervention (Gast & Ledford, 2014). Like multiple-baseline designs, alternating treatment designs do not require the withdrawal of the intervention. Alternating treatment designs, however, also do not require a stable baseline prior to the intervention's implementation. As there is no baseline requirement, and because the phases are short in duration, these designs can show carry-over effects of an intervention across phases (Gast & Ledford, 2014). Therefore, intervention data need to demonstrate robust and immediate effects on the dependent variable, typically evaluated by visual data analysis (Kazdin, 2011).

Analysis of Single-Subject Design Data

Data collected using single-subject designs are typically organised, graphed, and analysed using visual data analysis (Kazdin, 2011; Kennedy, 2005). Graphic displays of data are simple to create, and patterns within the data can instantly be read. Visual analysis is an ongoing process where sessions are scored, visually plotted on a graph, and then analysed session by session (Kennedy, 2005). The visual comparison of data points occurs both within phase and across phases.

Within-phase, or within-condition, patterns analyse phase length, level, and trend (Horner et al., 2005). Condition length simply refers to the number of data points within a phase. Level refers to the clustering of scores around a value on the vertical axis (Cooper et al., 2007). The level of a dataset is determined by level stability and level change. Level stability is the range of the data point values and described as stable or variable. Level change refers to the amount of change within the level of the same condition and is defined as either therapeutic or contratherapeutic (Gast & Spriggs, 2014).

Trend refers to the slope or angle of a dataset over time and is determined by trend direction and trend stability (Kratochwill et al., 2010). Trend direction is described as accelerating (i.e., increasing in interval scale value over time), decelerating (i.e., decreasing in interval scale value over time), or zero-celerating (i.e., no increase or decrease over time). Trend is estimated in one of two ways: by either the freehand method or the more reliable split-middle method (Gast & Spriggs, 2014). Trend stability uses the trend line and the range of data point values to determine the existence of variable data patterns.

Visual data analysis of between-phase patterns identifies level change, trend direction change, and percentage of nonoverlapping data (Horner et al., 2005). Level change refers to the immediacy of level change from one phase to the next and is used to determine the immediate impact of an intervention on behaviour as improving or deteriorating. Trend direction change is based on the results of trend

https://doi.org/10.1017/jsi.2020.4 Published online by Cambridge University Press

direction observations from within-phase analysis. If a baseline phase illustrates an accelerating trend and the following intervention phase demonstrates a decelerating trend, then the trend direction change between phases is labelled as accelerating to decelerating. Finally, the percentage of nonoverlapping data is a common method of determining the effectiveness of an intervention (Banda & Therrien, 2008). Typically, the higher the percentage of nonoverlapping data, the greater the impact of the independent variable on the dependent variable.

Although visual data analysis uses some simple metrics to interpret intervention effectiveness (e.g., percentage of nonoverlapping data), statistical indexes have not historically been used for single-subject design data (Derenne & Baron, 1999). Recently, however, it has become more common for researchers to use methods in addition to visual data analysis to report intervention effectiveness to support conclusions from single-subject designs (Olive & Smith, 2005). Both regression and nonregression effect size calculations (e.g., standard mean difference, percentage reduction measure, per cent reduction, percentage exceeding the median) have therefore been developed and reviewed (e.g., Olive & Franco, 2008; Olive & Smith, 2005; Şen & Şen, 2019).

Through such reviews it has been confirmed that nonregression methods are successful in identifying intervention effectiveness within single-subject experimental data (Olive & Smith, 2005). However, regression-based calculation methods have been shown to be affected by the serial dependency of single-subject design data and should be avoided (Olive & Smith, 2005; Şen & Şen, 2019). Therefore, it has been recommended that a nonregression effect size calculation be used in combination with visual data analysis to provide statistical evidence of intervention effectiveness (Olive & Franco, 2008).

Challenges of Single-Subject Designs

While embracing the foundations of single-subject designs, McConnell (2002) highlighted that the individuality of each single-subject design creates a challenge for the analysis of results against similar studies and even across cases within the same study. Difficulties in synthesising data often arise because of the varying scales used within single-subject designs. Also due to the unique nature of each single-subject design, the external validity of these designs is often considered weak and difficult to control (Fraenkel & Wallen, 2006). To strengthen the external validity of single-subject designs, data must demonstrate replication of effects across individuals, and the characteristics of each study must be operationally defined (Cakiroglu, 2012; Fraenkel & Wallen, 2006). In particular, the articulation of dependent variables in 'objective, observable and measurable terms' (Cakiroglu, 2012, p. 22) is essential in minimising variance between data collectors and between subject data. Although single-subject designs are well suited to answering research questions that determine whether and to what extent an intervention is effective, they cannot adequately determine the characteristics of a representative group of learners (Cook & Cook, 2016).

According to literature reviewing the use of single-subject designs within the field of special and inclusive education, findings have commonly indicated an overall increase in the use of single-subject designs over the past 30 years, with a preference for multiple-baseline designs across participants. It remains unclear, however, whether these trends have been maintained in recent years. The purpose of this review is to identify the current trends in the use of single-subject research designs in special education literature by answering the following questions:

1. Which are the preferred research methods used for intervention studies involving school-aged students with disability within current special education research?
2. Which types of single-subject designs are most frequently used within intervention studies involving school-aged students with disability within current special education research?
3. How have trends in the use of single-subject designs within the field of special education changed over the last five years?

https://doi.org/10.1017/jsi.2020.4 Published online by Cambridge University Press

Method

To answer the research questions, a systematic review of the literature was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009). The PRISMA statement was developed to improve the way that authors report systematic reviews and meta-analyses (Moher et al., 2009). The statement is accompanied by a 27-item checklist and the PRISMA elaboration document that provides a rationale for each of the checklist items. The PRISMA statement and associated documents can be found at http://www.prisma-statement.org. The following section describes how this systematic review was conducted to meet the requirements of the PRISMA statement through the checkpoints associated with systematic reviews.

Procedure

To satisfy the first PRISMA checkpoint, this review was given a title that identified it as a systematic review. To address the second checkpoint, a structured summary has been provided to establish a background for the review, the objectives of the review, and the eligibility criteria for inclusion in the review. To meet the third checkpoint, a rationale for the review was provided with reference made to what is already known. To address the fourth checkpoint, the rationale was used to write specific research questions that specify particular participants, interventions, and study designs. As a systematic review has not yet been conducted to answer the same specific questions in the past, the registration number of an existing review protocol has not been provided and therefore satisfies checkpoint five.

Checkpoint six requires eligibility criteria to be defined, and checkpoint nine deals with the study selection process. As this review is concerned with providing a snapshot of the current use of single-subject designs within the field, journal articles published within the seven selected journals between January 2014 and August 2019 were screened for eligibility to be included in this review. In addition, only articles reporting on intervention studies involving K–12 students with disability within school settings were included. This satisfies checkpoint seven, which requires information sources and inclusion dates to be described.

Checkpoint eight requires an account of the search strategy used. Like other systematic reviews (e.g., Hammond & Gast, 2010; Mastropieri et al., 2009), an electronic database search was replaced with a manual screening of articles published within the set time frame. To do this, BrowZine (i.e., a tool used to access, browse, search, and read entire e-journal archives from a variety of publishers from one online platform) was used to screen the complete archives of each journal over the defined period and identify those that were eligible for inclusion in this review. This process ensured that every article in each of the journals was screened rather than relying on electronic database search results.

As required by checkpoint 11, the following describes the variables for which data were sought. Once articles were identified as eligible, a set of coding rules were used to examine each article according to (a) year of publication, (b) first author's surname, and (c) research design (i.e., single-subject design, group design, single-group pre–post design, or other). Any study coded as a single-subject design was additionally coded according to the type of single-subject design employed (i.e., multiple-baseline across participants/participant groups, multiple-baseline across behaviours/tasks, multiple-baseline across conditions/settings, multiple-probe across participants/groups, multiple-probe across behaviours/tasks, multiple-probe across conditions/settings, withdrawal/reversal, alternating treatment, or changing-criterion design).

To address checkpoint 10, the data collection process must be provided. For this review, data were extracted from the journal articles. If an article included more than one research design, the types of designs used were coded separately. In addition, if a combination design was used, the designs were each identified and recorded.

To further satisfy checkpoint 10, the process for collecting and confirming data from multiple investigators follows. The author and a research assistant used BrowZine to access all journal articles

https://doi.org/10.1017/jsi.2020.4 Published online by Cambridge University Press

published in the seven selected journals between January 2014 and August 2019. Both coders initially worked together to screen and code a selection of articles from each journal according to the coding rules previously defined. Once coders reliably and consistently applied the coding rules, they continued to code the remaining articles independently. Reliability checks were conducted at regular intervals during the screening and coding process by dividing the total number of agreements by the total number of agreements and disagreements. Analysis revealed an interrater agreement rate of 99% for article eligibility for inclusion and between 82% and 100% for article coding. At the completion of coding, the coders checked data for consistency, and differences between the data of the two coders were identified and resolved. Once all differences between the coders were resolved, the number of articles published and screened within the selected sample of journals was recorded, in order to satisfy checkpoint 17. These figures can be found in the Results section that follows.

To meet the demands of checkpoints 24, 25, and 26, the Discussion section then provides a summary of the key findings within the context of what is already known, describes the relevance of findings for stakeholders (e.g., researchers, teachers, journal editors), and states the implications for future research and practice while discussing the limitations of this review at both the outcome level and review level.

Sample

Seven prominent special education journals were selected to provide a representative sample of high-quality article publication within the field of special education. These journals were selected based on their high visibility rates (i.e., as indicated by the 2018 figures for full-text usage, journal impact factor, and rank within the education, special category as reported by the journal home pages, Ulrich's Periodicals Directory, InCites Journal Citation Reports, and Cabells periodical directory), publication of articles that involve intervention studies, inclusion of research targeting cross-categorical or high-incidence disabilities (e.g., autism, intellectual disabilities, learning disabilities, emotional/behavioural disorders), and based on their use within previous reviews (e.g., Lessen, Dudzinski, Karsh, & Van Acker, 1989; Mastropieri et al., 2009). The following journals were therefore selected for inclusion in the review: *Behavioral Disorders* (*BD*), *Exceptional Children* (*EC*), *Focus on Autism and Other Developmental Disabilities* (*FOCUS*), *Journal of Emotional and Behavioral Disorders* (*JEBD*), *Journal of Learning Disabilities* (*JLD*), *Journal of Special Education* (*JSE*), and *Learning Disabilities Research & Practice* (*LDRP*).

EC and *JSE* represent the cross-categorical journals in the field. *EC* was selected for inclusion in this review as it is ranked 1/41, has a journal impact factor of 2.854, and a full-text use figure of 116,091. Further, *EC* publishes intervention studies involving qualitative, quantitative, and single-subject designs. *JSE* was selected for this review according to a similarly high full-text usage of 114,846 and publication of primary intervention research.

LDRP and *JLD* were both selected as journals to represent a broad range of learning disabilities. *LDRP* was selected for inclusion in this review as it is ranked 7/41, has a journal impact factor of 2.077, and publishes intervention articles with a practical emphasis. *JLD* was also selected for this review because of its remarkably high full-text usage figure of 252,686, along with its rank of 4/41 and journal impact factor of 2.578. In addition, *JLD* publishes intervention and primary research articles.

In this review, *JEBD* and *BD* represent journals dedicated to emotional and behavioural disorders. *JEBD* was selected for inclusion in the review as it is ranked 12/41, has a journal impact factor of 1.775, and publishes intervention and treatment studies. *BD* was included in the review because of its focus on publishing research-based articles on evidence-based practices.

Finally, *FOCUS* was selected as the journal to represent journals dedicated to developmental disorders, specialising in autism. *FOCUS* was selected for inclusion based on its emphasis on the publication of articles discussing practical interventions and treatments while having a large full-text usage figure of 132,313.

https://doi.org/10.1017/jsi.2020.4 Published online by Cambridge University Press

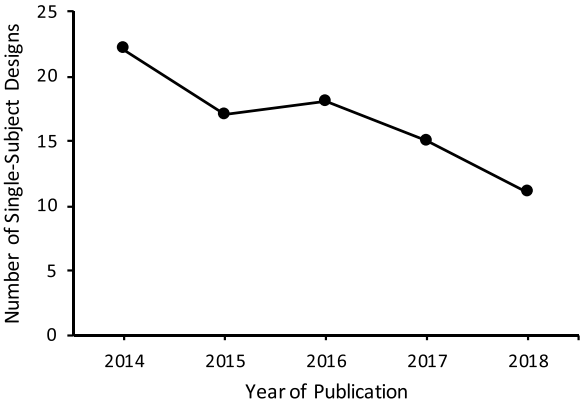


Figure 1. Number of single-subject designs over 5 years.

Results

A total of 1,009 articles were screened for their eligibility to be included in this review. Of the 1,009 published articles, 203 reported an intervention study involving K–12 students within school settings and therefore met the inclusion criteria. As each research design was coded separately, some articles included multiple designs and resulted in the coding of 209 designs.

Preferred Research Designs

Of the 209 eligible studies, 94 (45.0%) were coded as single-subject designs. This total suggests that single-subject designs are one of the preferred research designs used for intervention studies involving school-aged students within current special education research. However, there was a substantial discrepancy between the percentages of single-subject designs published within each journal. For example, the majority of eligible studies were coded as single-subject designs in *BD* (77.8%), *FOCUS* (93.6%), and *JSE* (86.7%), whereas single-subject designs made up a much smaller proportion in *EC* (20.0%), *JEBD* (22.2%), *JLD* (8.0%), and *LDRP* (37.5%).

By a marginal difference, group designs were coded as the second most preferred research design across the sample. Of the 209 studies, 87 (41.6%) research designs were coded as group designs. The three journals contributing most significantly to this figure were *EC*, *JLD*, and *LDRP*, with group designs representing 63.0%, 81.0%, and 41.7% of all coded research designs, respectively. Both of the journals representing research in learning disabilities (*JLD* and *LDRP*) have published a greater proportion of group designs than any other research design.

Most Frequently Used Types of Single-Subject Designs

Of the 45.0% of studies coded as single-subject designs, the majority involved a multiple-baseline (35.1%) or multiple-probe design (39.4%). The bulk of multiple-baseline and multiple-probe designs were employed across participants/participant groups (77.1%). Multiple-baseline or multiple-probe designs across behaviours/tasks represented 14.3%, while multiple-baseline and multiple-probe designs across conditions/settings represented only 8.6%. Alternating treatment design was the second most frequently used single-subject design (18.1%), followed by a reversal/withdrawal design (5.3%). Changing-criterion designs accounted for only 2.1% of the designs coded as single subject across the sample.

Single-Subject Design Trends in the Last Five Years

Figure 1 shows the total number of research designs coded as single subject within the sample of journals for each year between 2014 and 2018. The data for 2019 have not been included in this figure as they were not collected for the second half of the year.

https://doi.org/10.1017/jsi.2020.4 Published online by Cambridge University Press

Microsoft Excel was used to calculate the trend of the data represented in Figure 1. First, a linear trend line superimposed onto the graph illustrates a decelerating trend, indicating that the number of single-subject designs used within the sample of special education journals has decreased over the last 5 years. Second, a logarithmic trend line was used to extrapolate the data and provide an estimate of the trend over the next 5 years (2019–2023). The logarithmic trend line showed the continuation of a decelerating trend over time, suggesting that the use of single-subject designs in intervention research for school-aged students with disability may continue to decline in the future.

Discussion

The findings emerging from the current review provide a modest insight into the past, present, and future trends of published research within the field of special and inclusive education. Overall, the researcher found that 20.1% of the screened articles were intervention studies involving school-aged students conducted within school settings. When compared to past studies, this figure suggests that there has been a growth in the publication of intervention studies involving students within special education journals over time. For example, Mastropieri et al. (2009) found that 15.9% of the studies within their sample were intervention studies involving students from preschool to Year 12. This sample was inclusive of articles published within 11 special education journals between 1988 and 2006. Even though Mastropieri et al. (2009) included research involving preschool-aged participants in addition to K–12 students, their figures remain less than those obtained in the current study. A possible explanation for this suggested increase is the formal requirement for data-informed reporting by schools. Particularly since the introduction of the Nationally Consistent Collection of Data in Australia, schools have been expected to provide evidence of the effectiveness of programs and interventions. Journal articles on intervention research demonstrate the link between research, practice, and evaluation of performance, and are a valuable source of guidance for educators and service providers as they evaluate interventions and modify practices in a timely manner. Therefore, it is important for journal editors to maintain the rigour of single-subject designs, including those published in practitioner-focused journals.

As predicted by Alnahdi (2015) and indicated through the current review, it seems that researchers within the field of special education continue to use single-subject designs and group designs as their preferred experimental methods. This finding could be attributed to manuscript reviewing and editing processes. For example, journal editors and reviewers may accept and recommend articles for publication based on their own preferred experimental methods. This could result in a distorted representation of current research and makes it difficult to accurately determine which experimental methods are most commonly being used within the field. Rather, findings more accurately illustrate which experimental methods are most frequently published within the field. Researchers conducting future reviews in the area may therefore benefit from investigating publication rates rather than usage rates or preferences.

Within the field of special and inclusive education, single-subject designs are used as an alternative to group designs because of their cost effectiveness, difficulty securing large numbers of participants with low-incidence disabilities, and avoiding the ethical concerns of withholding treatment from a control group (Cakiroglu, 2012; Shadish, Cook, & Campbell, 2002). As single-subject designs are often used as a necessary alternative, it remains difficult to distinguish between preference and need (Hammond & Gast, 2010; McDonnell & O'Neill, 2003). This further demonstrates the need for future research to consider investigating rates of use rather than preferences.

Reflecting the findings of Hammond and Gast (2010) and Moeyaert et al. (2013), the current review provides evidence that multiple-baseline designs, particularly those across participants/participant groups, continue to be published more frequently than withdrawal/reversal designs. This is possibly because withdrawal/reversal designs require the target behaviour to be reversible while multiple-baseline designs do not. As many social, behavioural, and academic behaviours cannot be reversed once acquired, withdrawal/reversal designs are typically less appropriate than multiple-baseline

https://doi.org/10.1017/jsi.2020.4 Published online by Cambridge University Press

designs. In addition, multiple-probe designs may also be popular because data do not need to be collected as frequently. Although experimental control can be demonstrated through the use of multiple-baseline and multiple-probe across participant designs, they do not enable intrasubject replication. In order to demonstrate a functional relationship between the dependent and independent variables, practitioners should illustrate replication of effect across at least three conditions or behaviours with a minimum of three participants. Thus, those interested in conducting or publishing more rigorous research should consider multiple-baseline and multiple-probe designs across conditions/settings or behaviours/tasks.

There remains a pattern of difference in the types of research designs published by different journals. Historically, some journals have published more intervention studies involving single-subject designs (e.g., *BD*, *FOCUS*, *JSE*) than others (e.g., *EC*, *JEBD*, *JLD*, *LDRP*), represented in both past reviews and the current review (Mastropieri et al., 2009). This trend is difficult to explain, as most journals within the current sample explicitly state that they welcome the submission of manuscripts of single-subject design research. In addition, two journals in the sample, *JLD* and *LDRP*, do not list single-subject designs in their guidelines for authors but continue to publish single-subject research. Perhaps some journals attempt to maintain a balance among the various research designs whereas others have a preference for single-subject designs. Another possible reason for this could be that journals hold higher standards for single-subject research that make manuscript acceptance more difficult, particularly since a number of analyses have identified the limited number of published high-quality single-subject designs (Gersten et al., 2005; Seethaler & Fuchs, 2005). It is therefore important for journal editors to have clear goals in relation to the publication rates of single-subject designs and the standards expected. Further, it would be advantageous for publishers to clearly state these goals and standards in their guidelines for authors.

The comprehensive review conducted by Mastropieri et al. (2009) found that 50.5% of intervention research used single-subject designs, while the current review found that a comparable 45.0% of intervention research used single-subject designs. The logarithmic trend line used to extrapolate the data from the current review also suggests a continuation of decreased single-subject design use over the next 5 years. These figures illustrate a possible decline in the use, or publication, of single-subject designs over the last 10 years that may continue into the future. This decelerating trend may be associated with teachers' increased recognition of the importance of inclusive practices that demonstrate benefits for all students rather than focusing only on the instructional impacts for students with specific disabilities or characteristics (Maciver et al., 2018). Therefore, group designs that involve a more varied student sample may be favoured over single-subject designs that typically follow the effects of intervention on individual students. While group designs offer greater simplicity, single-subject designs should be considered, as they enable educators to systematically evaluate the effectiveness of interventions for students with specific characteristics and lead to well-informed professional decision-making within diverse mainstream classrooms (Cook & Cook, 2016; Maggin et al., 2018; Tankersley et al., 2008).

Limitations

This review is limited to the literature published within seven special education journals from January 2014 to the end of August 2019. The inclusion of additional or alternative journals in this review, or a longer time period, may have provided different or more robust findings. In addition, the interpretation of findings from the current study are only applicable to the field of special and inclusive education, and it would therefore be appropriate to explore these findings within the broader context of education. For future studies, researchers could investigate the topic on a larger scale through the selection of additional journals, extending the time frame, or expanding the field of investigation beyond special and inclusive education.

Since single-subject designs have long been evaluated against a set of standards (Kratochwill et al., 2013; Smith, 2012), research investigating the quality of published single-subject designs has been

https://doi.org/10.1017/jsi.2020.4 Published online by Cambridge University Press

conducted in the past (e.g., Hsieh et al., 2005; Seethaler & Fuchs, 2005). Considering the age of such research, it would have been advantageous for the researcher to have examined the quality of published research in addition to publication frequency. Researchers should consider this need in the future.

While it is acknowledged that there was risk of bias across the studies included in this review, it was difficult to confirm the existence of bias and therefore extremely difficult to manage; for example, publication bias, whereby the selected journals may have published studies that use specific research designs and therefore could have affected the cumulative findings. It is important to be aware of such bias when interpreting the overall results of the current analysis.

Regardless of the limitations, this review provides a valuable snapshot of the recent trends in research design use within the field of special and inclusive education, although replication and extension of this review is warranted.

Single-Subject Design Exemplars

The following describes two recent special education studies that have employed different single-subject designs to investigate the impacts of interventions on school-aged students with disability. These studies have been selected as exemplars for their social validity, ease of replication, and adherence to single-case design standards.

Kouo (2019) explored the use of a package intervention to teach social initiations to five kindergarten students on the autism spectrum. The package intervention included point-of-view modelling, video priming, verbal praise, reinforcement, and prompting. A multiple-baseline-across-participants design was used to investigate the impact of the intervention on the participants' use of greeting with trained typically developing peers who acted as their communication partners. The multiple-baseline design allowed for experimental control to be demonstrated, and visual data analysis confirmed that of the five participants, three established a stable and rapid increase in the dependent variable (i.e., social initiations) as the independent variable (i.e., video-modelling package intervention) was implemented. Although the other two participants also reached the established criterion, there was a greater level of variability within their datasets. Kouo (2019) aligned their study with the single-case design standards defined by Kratochwill et al. (2010) and the single-case quality indicators suggested by Horner et al. (2012), while demonstrating socially important findings within authentic learning settings.

Satsangi, Hammer, and Hogan (2019) also followed the single-case design standards suggested by Kratochwill et al. (2013) to conduct a single-subject alternating treatment design across participants to compare the use of video modelling and explicit instruction to teach geometry. The study involved three female high-schoolers with learning disabilities. After five baseline sessions, participants received five sessions of video modelling alternated with five sessions of explicit instruction to determine the effect on students' performance of solving word problems in geometry. Visual data analysis was used in combination with the effect size measure, Tau-*U*. Results of this analysis established the existence of a functional relationship between both video modelling and explicit instruction and student performance. In addition, the explicit instruction intervention was shown to be slightly more effective for two of the three students.

Conclusion

This article provides teachers and researchers of students with disability with a rationale for using single-subject designs as an alternative to group designs to evaluate interventions and monitor student performance. The past, present, and future of single-subject design research methods within the field of special education have been discussed through the examination of data emerging from systematic reviews of the literature. Although the use of single-subject designs has gained popularity within special education publications over the last 35 years, there is some evidence of a decline in the current and future use of the research method within published articles. As single-subject designs remain relevant

https://doi.org/10.1017/jsi.2020.4 Published online by Cambridge University Press

within the field of special education according to their cost effectiveness, ease of use, and ability to be individualised, teachers and researchers of students with disability should continue to consider using the method to demonstrate the effectiveness of interventions within authentic contexts.

References

Alnahdi, G. H. (2015). Single-subject designs in special education: Advantages and limitations. *Journal of Research in Special Education Needs*, *15*, 257–265. https://doi.org/10.1111/1471-3802.12039

Banda, D. R., & Therrien, W. J. (2008). A teacher's guide to meta-analysis. *Teaching Exceptional Children*, *41*(2), 66–71. https://doi.org/10.1177/004005990804100208

Cakiroglu, O. (2012). Single subject research: Applications to special education. *British Journal of Special Education*, *39*, 21–29. https://doi.org/10.1111/j.1467-8578.2012.00530.x

Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education* (5th ed.). London, England: Routledge Falmer. https://doi.org/10.4324/9780203224342

Conroy, M. A., Dunlap, G., Clarke, S., & Alter, P. J. (2005). A descriptive analysis of positive behavioral intervention research with young children with challenging behavior. *Topics in Early Childhood Special Education*, *25*, 157–166. https://doi.org/10.1177/02711214050250030301

Cook, B. G., & Cook, L. (2016). Research designs and special education research: Different designs address different questions. *Learning Disabilities Research & Practice*, *31*, 190–198. https://doi.org/10.1111/ldrp.12110

Cook, L., Cook, B. G., Landrum, T. J., & Tankersley, M. (2008). Examining the role of group experimental research in establishing evidence-based practices. *Intervention in School and Clinic*, *44*, 76–82. https://doi.org/10.1177/1053451208324504

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd ed.). Upper Saddle River, NJ: Pearson/Merrill-Prentice Hall.

Derenne, A., & Baron, A. (1999). Human sensitivity to reinforcement: A comment on Kollins, Newland, and Critchfield's (1997) quantitative literature review. *The Behavior Analyst*, *22*, 35–41. https://doi.org/10.1007/BF03391976

Forbes, S. A., Ross, M. E., & Chesser, S. S. (2011). Single-subject designs and action research in the K–12 setting. *Educational Research and Evaluation*, *17*, 161–173. https://doi.org/10.1080/13803611.2011.599555

Fraenkel, J. R., & Wallen, N. E. (2006). *How to design and evaluate research in education* (6th ed.). New York, NY: McGraw-Hill.

Gast, D. L., & Ledford, J. R. (Eds.). (2014). *Single case research methodology: Applications in special education and behavioral sciences* (2nd ed.). New York, NY: Routledge. https://doi.org/10.4324/9781315150666

Gast, D. L., & Spriggs, A. D. (2014). Visual analysis of graphic data. In D. L. Gast & J. R. Ledford (Eds.), *Single case research methodology: Applications in special education and behavioral sciences* (pp. 176–210). New York, NY: Routledge. https://doi.org/10.4324/9780203521892-9

Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children*, *71*, 149–164. https://doi.org/10.1177/001440290507100202

Hacker, B. (1980). Single subject research strategies in occupational therapy, part 1. *American Journal of Occupational Therapy*, *34*, 103–108. https://doi.org/10.5014/ajot.34.2.103

Hammond, D., & Gast, D. L. (2010). Descriptive analysis of single subject research designs: 1983–2007. *Education and Training in Autism and Developmental Disabilities*, *45*, 187–202.

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. https://doi.org/10.1177/001440290507100203

Horner, R. H., & Kratochwill, T. R. (2012). Synthesizing single-case research to identify evidence-based practices: Some brief reflections. *Journal of Behavioral Education*, *21*, 266–272. https://doi.org/10.1007/s10864-012-9152-2

Horner, R. H., Swaminathan, H., Sugai, G., & Smolkowski, K. (2012). Considerations for the systematic analysis and use of single-case research. *Education and Treatment of Children*, *35*, 269–290. https://doi.org/10.1353/etc.2012.0011

Hsieh, P., Acee, T., Chung, W.-H., Hsieh, Y.-P., Kim, H., Thomas, G. D., … Robinson, D. H. (2005). Is educational intervention research on the decline? *Journal of Educational Psychology*, *97*, 523–529. https://doi.org/10.1037/0022-0663.97.4.523

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

Kennedy, C. H. (2005). *Single-case designs for educational research*. Boston, IL: Pearson Education.

Kouo, J. L. (2019). The effectiveness of a packaged intervention including point-of-view video modeling in teaching social initiation skills to children with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities*, *34*, 141–152. https://doi.org/10.1177/1088357618815887

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). *Single-case designs technical documentation*. Retrieved from What Works Clearinghouse website: https://files.eric.ed.gov/fulltext/ED510743.pdf

https://doi.org/10.1017/jsi.2020.4 Published online by Cambridge 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*, 26–38. https://doi.org/10.1177/0741932512452794

Kratochwill, T. R., & Levin, J. R. (1992). *Single-case research design and analysis: New directions for psychology and education*. Hillsdale, NJ: Erlbaum. https://doi.org/10.4324/9781315725994

Lessen, E., Dudzinski, M., Karsh, K., & Van Acker, R. (1989). A survey of ten years of academic intervention research with learning disabled students: Implications for research and practice. *Learning Disabilities Focus*, *4*, 106–122.

Maciver, D., Hunter, C., Adamson, A., Grayson, Z., Forsyth, K., & McLeod, I. (2018). Supporting successful inclusive practices for learners with disabilities in high schools: A multisite, mixed method collective case study. *Disability and Rehabilitation*, *40*, 1708–1717. https://doi.org/10.1080/09638288.2017.1306586

Maggin, D. M., Cook, B. G., & Cook, L. (2018). Using single-case research designs to examine the effects of interventions in special education. *Learning Disabilities Research & Practice*, *33*, 182–191. https://doi.org/10.1111/ldrp.12184

Mastropieri, M. A., Berkeley, S., McDuffie, K. A., Graff, H., Marshak, L., Conners, N. A., … Cuenca-Sanchez, Y. (2009). What is published in the field of special education? An analysis of 11 prominent journals. *Exceptional Children*, *76*, 95–109. https://doi.org/10.1177/001440290907600105

McConnell, S. R. (2002). Interventions to facilitate social interaction for young children with autism: Review of available research and recommendations for educational intervention and future research. *Journal of Autism and Developmental Disorders*, *32*, 351–372. https://doi.org/10.1023/A:1020537805154

McDonnell, J., & O'Neill, R. (2003). A perspective on single/within subject research methods and "scientifically based research". *Research and Practice for Persons with Severe Disabilities*, *28*, 138–142. https://doi.org/10.2511/rpsd.28.3.138

McMillan, J. H. (2004). *Educational research: Fundamentals for the consumer* (4th ed.). Boston: Pearson/Allyn & Bacon.

Moeyaert, M., Ugille, M., Ferron, J. M., Beretvas, S. N., & Van den Noortgate, W. (2013). The three-level synthesis of standardized single-subject experimental data: A Monte Carlo simulation study. *Multivariate Behavioral Research*, *48*, 719–748. https://doi.org/10.1080/00273171.2013.816621

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Public Library of Science Medicine*, *6*(7), 1–6. https://doi.org/10.1371/journal.pmed.1000097

Odom, S. L., Brantlinger, E., Gersten, R., Horner, R. H., Thompson, B., & Harris, K. R. (2005). Research in special education: Scientific methods and evidence-based practices. *Exceptional Children*, *71*, 137–148. https://doi.org/10.1177/001440290507100201

Olive, M. L., & Franco, J. H. (2008). (Effect) size matters: And so does the calculation. *Behavior Analyst Today*, *9*, 5–10.

Olive, M. L., & Smith, B. W. (2005). Effect size calculations and single subject designs. *Educational Psychology*, *25*, 313–324. https://doi.org/10.1080/0144341042000301238

Pustejovsky, J. E., & Ferron, J. M. (2017). Research synthesis and meta-analysis of single-case designs. In J. M. Kauffman, D. P. Hallahan, & P. C. Pullen (Eds.), *Handbook of special education* (2nd ed., pp. 168–186). New York, NY: Routledge. https://doi.org/10.4324/9781315517698-15

Satsangi, R., Hammer, R., & Hogan, C. D. (2019). Video modeling and explicit instruction: A comparison of strategies for teaching mathematics to students with learning disabilities. *Learning Disabilities Research & Practice*, *34*, 35–46. https://doi.org/10.1111/ldrp.12189

Seethaler, P. M., & Fuchs, L. S. (2005). A drop in the bucket: Randomized controlled trials testing reading and math interventions. *Learning Disabilities Research & Practice*, *20*, 98–102. https://doi.org/10.1111/j.1540-5826.2005.00125.x

Şen, N., & Şen, S. (2019). Calculation of effect size in single-subject experimental studies: Examination of non-regression-based-methods. *Journal of Measurement and Evaluation in Education and Psychology*, *10,* 30–48. https://doi.org/10.21031/epod.419625

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, IL: Houghton Mifflin.

Sidman, M. (1960). *Tactics of scientific research: Evaluating experimental data in psychology*. New York, NY: Basic Books.

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

Tankersley, M., Harjusola-Webb, S., & Landrum, T. L. (2008). Using single-subject research to establish the evidence base of special education. *Intervention in School and Clinic*, *44*, 83–90. https://doi.org/10.1177/1053451208321600

Wolery, M., Dunlap, G., & Ledford, J. R. (2011). Single-case experimental methods: Suggestions for reporting. *Journal of Early Intervention*, *33*, 103–109. https://doi.org/10.1177/1053815111418235

Wong, C., Odom, S. L., Hume, K. A., Cox, A. W., Fettig, A., Kucharczyk, S., … Schultz, T. R. (2015). Evidence-based practices for children, youth, and young adults with autism spectrum disorder: A comprehensive review. *Journal of Autism and Developmental Disorders*, *45*, 1951–1966. https://doi.org/10.1007/s10803-014-2351-z

Cite this article: Zanuttini, J. Z. (2020). Using a systematic review to explore single-subject design use within special education over time. *Australasian Journal of Special and Inclusive Education* **44**, 1–14. https://doi.org/10.1017/jsi.2020.4

https://doi.org/10.1017/jsi.2020.4 Published online by Cambridge University Press