J Head Trauma Rehabil
Vol. 38, No. 2, pp. E126–E135

Examining the Evidence From Single-Case Experimental Designs to Treat Challenging Behaviors Following Moderate to Severe Traumatic Brain Injury

Cynthia L. Beaulieu, PhD; Chris Persel, MA; Tracy Shannon, PsyD; John Whyte, MD, PhD; Daniel Hurlburt, PsyD; Natasha Huffine, MS, OTR/L; Jennifer Bogner, PhD

Objective: To evaluate evidence on the effectiveness of behavioral interventions using single-case experimental design (SCED) methodology and to identify behavioral interventions with sufficient evidence for possible inclusion in the development of guidelines for the management of challenging behaviors in adults following moderate to severe traumatic brain injury (TBI). **Methods:** As a subinvestigation of a larger systematic review process designed to identify evidence for guidelines development, the current review focused on studies using SCED methodology applied to persons with challenging behaviors following moderate to severe TBI. Articles were identified from a search of the published literature through January 2021, identifying studies in CINAHL, Cochrane Database of Systematic Reviews, EMBASE, MEDLINE/Ovid, and PsycINFO. Articles meeting inclusion criteria were assessed for design rigor to allow for effect size determination. The identified cases were then critically appraised using the RoBiNT (Risk-of-Bias in N-of-1 Trails) Scale to determine strength of evidence for causal inference. **Results:** Thirty-four studies met inclusion criteria, with a total of 44 cases evaluated for effect of the treatment intervention on defined target behaviors. Seventeen cases had effect sizes rated as large, 22 cases as medium, 3 cases as small, and 3 as no effect. An observed trend was for large and medium effect sizes to be associated with lower RoBiNT Scale internal validity scores. Randomization, blinded provider and assessor, and assessment of treatment adherence were the internal validity items unlikely to meet criteria. **Conclusions:** SCED methodology was found to produce large and medium effect sizes for behavioral interventions targeting challenging behaviors following moderate to severe TBI. However, the strength of the evidence is limited because of weaknesses in study designs. Most of the studies failed to meet established internal validity criteria designed to reduce risk of bias in SCED studies as such rigor is difficult to establish or often not practical in clinical settings. Suggestions and recommendations are outlined for improving the quality of published cases using SCED methodology, which, in turn, will improve credibility of evidence and better inform the development of treatment guidelines for behavior regulation. **Key words:** *behavior*, *brain injury*, *single case*, *single-case experimental design*, *single subject*, *systematic review*, *traumatic brain injury*, *treatment*

IN 2014, the Brain Injury Association of America (BIAA) commissioned a panel of experts to conduct systematic reviews and develop guidelines for chronic

Author Affiliations: *Division of Rehabilitation Psychology, Department of Physical Medicine and Rehabilitation, The Ohio State University College of Medicine, Columbus (Drs Beaulieu, Shannon, Hurlburt, and Bogner); Centre for Neuro Skills, Bakersfield, California (Mr Persel); Moss Rehabilitation Research Institute, Elkins Park, Pennsylvania (Dr Whyte); and Origami Brain Injury Rehabilitation Center, Mason, Michigan (Ms Huffine).*

This study was supported by the Brain Injury Association of America Treatment Guidelines Project awarded to Wayne A. Gordon and Marcel Dijkers of the Icahn School of Medicine at Mount Sinai.

The authors thank the following researchers who assisted in the early stages of this project: Marcel Dijkers, PhD; Wayne Gordon, PhD; David Arciniegas, MD; Charles Bombardier, PhD; Thomas Hall, MA; Harvey Jacobs, PhD; Megan Heinicke, PhD; Carolyn Rocchio, Jean Rush, MSN, Jonathan Silver, MD; and Nicole Feeling, PhD.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website (www.headtraumarehab.com).

The authors declare no conflicts of interest.

Corresponding Author: *Cynthia L. Beaulieu, PhD, Division of Rehabilitation Psychology, Department of Physical Medicine and Rehabilitation, The Ohio State University College of Medicine, 480 Medical Center Dr, Columbus, OH 43210 (cynthia.beaulieu@osumc.edu).*

DOI: 10.1097/HTR.0000000000000795

management of moderate to severe traumatic brain injury (TBI).[1] Five panels were developed (behavior, cognition, function, long-term medical, and vocational-community), and each was charged with developing relevant questions to be addressed through the Population, Intervention, Comparator, Outcome, Time frame (PICOT) process.[1] The behavior issues panel identified the need to review studies comprising single-case experimental design (SCED). Specifically, the panel recognized the need to evaluate the unique and independent contribution of SCED evidence for development of behavioral treatment guidelines.

Prior systematic reviews investigating behavioral interventions to treat various behaviors following brain injury have been completed. Many prior reviews have focused on pediatric populations, young adults, or a mixture of pediatric and adult groups following acquired brain injury.[2–4] Conclusions from these reviews include the following: (1) behavioral interventions are effective for reducing challenging behaviors or increasing adaptive skills in children and adolescents, and effectiveness is seen across multiple settings[3]; (2) behavioral interventions should be considered for both acute and chronic stages of recovery and for both adults and children[4]; (3) treatments employing applied behavior analysis should be considered an evidence-based option for psychosocial or challenging behaviors following brain injury[5]; and (4) a wide range of challenging behaviors are targeted, either for reduction or for acquisition, but study methodology was found to be relatively poor.[2]

Weak rigor on which conclusions of effectiveness have been drawn has led to attempts to improve methodological weaknesses.[6–10] Standards have been developed and criteria established along with scoring systems to determine whether an SCED satisfies established criteria. Core features of SCED methodology targeted for standards development include characteristics of participants and practitioners, observational periods and phases, baseline conditions, target behavior, dependent and independent variables, and experimental control.[6,10,11] Development of SCED methodology standards has evolved over time. The intent of the standards is to allow for the evaluation of causal validity within SCED.[8] Evaluating causal validity has become particularly important since single-case methodology can be considered level 1 evidence for informing treatment decisions by some evidence-based bodies.[11,12] Critical appraisal provides a systematic process to evaluate quality of evidence and determine cause-effect relationships between intervention and targeted behavior.[13]

The need for critical appraisal of SCED methodology has led to the development of tools to evaluate quality of studies implementing SCED.[10] The Risk-of-Bias in N-of-1 Trails (RoBiNT) Scale, a revision of the previously developed SCED Scale,[14] was created to meet the need for critical appraisal of SCED methodology using scored criteria in a psychometrically sound approach to assessing internal and external validity.[11,13] Relative to other critical appraisal tools, the RoBiNT Scale separates assessment of internal validity from external validity since either can contribute for different reasons to strengths or weaknesses of a study design. During revision, the RoBiNT Scale added items to assess for randomization and blinding, powerful design features that can control for threats to internal validity and thus improve credibility of the design and data derived from the design. Scoring on the RoBiNT Scale is nonbinary, allowing for greater sensitivity for discrimination across studies.[10]

Separate from assessing the quality of an SCED study assessing effect size can assist in determining level of evidence to inform the development of practice guidelines. For example, a high-quality SCED study with a large effect size might carry more weight in guideline development than a high-quality SCED study with no effect size or a low-quality SCED study with a large effect size. Traditional and consistently recommended analytic approach is through systematic visual inspection.[6,15–17] Visual inspection is also the preferred method when assessing for clinical significance.[10] Within-phase visual analysis assesses the data for variability or stability, trend or slope, and level or intensity, while between-phase visual analysis assesses for consistency or sameness in similar phases, overlap across phases, and immediacy or latency of change.[6,7,17]

The present study focused on published studies that used SCED methodology to treat challenging behaviors following moderate to severe TBI. Studies meeting inclusion criteria were evaluated on a number of measures, including size of treatment effect on targeted behavior(s) and quality of experimental design as measured by the RoBiNT Scale. The purpose was to determine strength of the SCED evidence in contributing toward the development of guidelines for the management of challenging behaviors in adults following moderate to severe TBI.

METHODS

Defined behavior search targets

During a 2-day face-to-face meeting in 2014, each domain panel drafted PICOT questions based on extensive discussions of identified critical clinical issues requiring a synthesis of evidence. Drafted questions were finalized in subsequent follow-up conference calls with domain panel representatives. The PICOT questions guided the development of a comprehensive list of specific challenging behaviors to target for systematic search and review (see Table 1).

Literature search and selection criteria

Search terms, MeSH descriptor trees, permutations of specific key words, and algorithms were developed. Systematic search of published literature through January 2021 identified studies in CINAHL, Cochrane Database of Systematic Reviews, EMBASE, MEDLINE/Ovid, and PsycINFO using key words and terms related to Behavior and TBI (for details, see Supplemental Digital Content [SDC] Table A, available at: http://links.lww.com/JHTR/A562). A research assistant with expertise in SCED methodology extracted 323 relevant articles and forwarded these articles to the behavior panel. Two members (J.B., C.L.B.) read through abstracts and selected articles meeting the following inclusion criteria: (1) subjects were 16 years or older; (2) diagnostic etiology was TBI; (3) sustained TBI was moderate to severe in severity; and (4) dependent variable was a behavior included on the list of challenging behaviors (see Table 1).

TABLE 1 *Behaviors included in article identification*

Aggression
Agitation (including psychomotor agitation, hyperactivity, restlessness, mania, akathisia)
Agnosia
Amotivation
Anger
Apathy, abulia
Challenging behavior
Conduct disorder
Dangerous behavior
Destructive behavior
Disinhibition
Disruptive
Impulsivity/impulsivity/impulse control disorders
Inappropriate sexual behavior
Indecent exposure
Inhibition
Interpersonal relations
Marital conflicts
Nonadherence, patient compliance, adherence
Offending (legal) behavior
Parenting skills
Perseveration
Personal boundary violations
Self-destructive behavior
Self-injurious behavior
Self-injury
Self-mutilation
Severe behavior problems/disorders
Sex offenses, hypersexual
Stereotyped behavior
Substance abuse, drinking behavior, alcohol abuse/use, drug abuse/use, dependence, addictive behavior
Violence/violent behavior
Voyeurism

Articles meeting inclusion criteria were forwarded to the behavior panel. A total of 35 studies met inclusion criteria (see Figure 1 for selection tree). Each article was assigned to 2 reviewers who were members of the behavior panel.

Each reviewer independently assessed rigor of experimental design for each case. Some studies included multiple cases that were not part of a multiple baseline across persons design or replication design. In these situations, each case that met inclusion criteria was evaluated individually. The criteria to meet minimum rigor included the following: (1) inclusion of a minimum of 3 phases or 2 opportunities for change in behavior; and (2) inclusion of a minimum of 3 data points per phase. Single-subject studies consisting of A-B or biphasic methodology were excluded. These criteria were deemed necessary in order to have adequate data to determine an effect size. Discrepancies between reviewers in assessing rigor of design were brought back to the entire behavior panel for discussion and resolution. Only cases meeting criteria for rigor moved on to visual analysis. For some published articles, multiple cases from one study moved forward for visual analysis of effect size. Of 35 studies meeting inclusion selection criteria, 45 SCED cases were evaluated for effect size and quality of design.

Visual analysis of effect size

Reviewers assigned the task of determining rigor of each SCED were also tasked with visually analyzing cases for effect size. The objective was to determine experimental or causal effect of the intervention on the target behavior. Reviewers were instructed to consider all phases and series within a given case. Each reviewer independently performed the visual analysis, considering features of variability/stability, trend/slope, intensity, consistency, overlap, and latency of change (see Table 2). The visual analysis followed the approach outlined in Tate and Perdices.[10(p159)] Within-phase analysis assessed for variability or stability, trend or slope, and level or intensity. Between-phase visual analysis assessed for consistency or sameness in similar phases, overlap across phases, and immediacy or latency of change.[6,7,17] Discrepancies between assigned reviewers were brought back to members of the behavioral panel for discussion and consensus resolution. Following initial ratings, reviewers achieved agreement on 22 of 45 cases (49%). Cases with discordant ratings were forwarded to the panel for additional review and subsequently achieved agreement on the remaining 23 cases.

Appraisal of design quality with the RoBiNT Scale

The RoBiNT Scale was completed on all 45 cases from the 35 articles. Completion of the RoBiNT Scale per

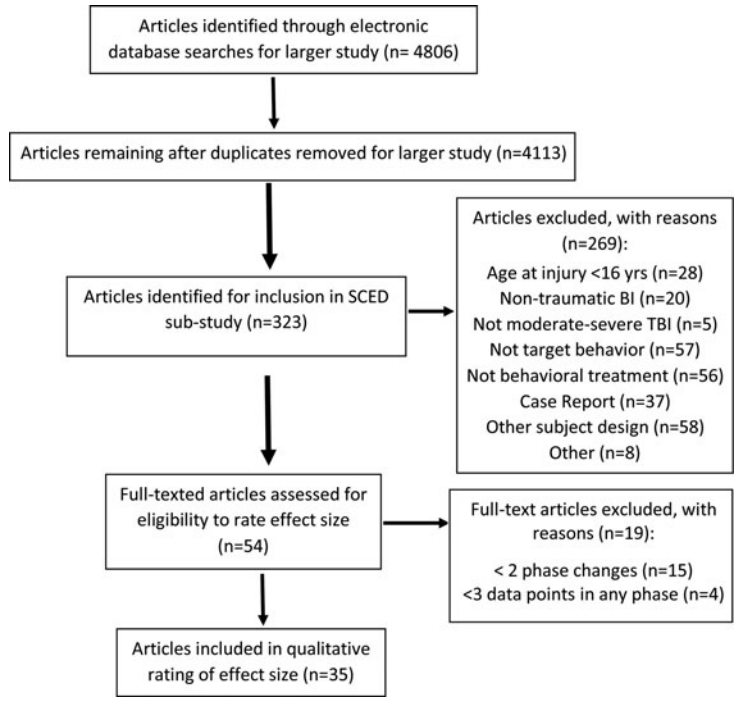

Figure 1. Systematic search and identification. SCED indicates single-case experimental design.

case was justified using the following logic: (1) each case within an article was reviewed, assessed, and selected on the basis of rigor of its design; and (2) aside from cases using multiple baseline design across persons, case designs within an article tended to tailor the design to the subject rather than applying the exact same design to all subjects. Each study was rated independently by 3 reviewers (T.S., D.H., C.L.B.). Discrepancies between reviewers were discussed within the group and then finalized with a fourth reviewer (J.B.).

The RoBiNT Scale consists of 15 items, categorized into 2 subscales that allow independent evaluation of both internal (7 items) and external validity (8 items).[11] Items are scored on a 3-point rating scale, with a possible total score of 30. Each item is rated 0, 1, or 2 for not meeting, partially meeting, or fully meeting the described criteria, respectively.[13] The RoBiNT Scale has demonstrated psychometrically sound results for inter-rater reliability and discriminative validity.[10] To date, there are no established "cutoff" scores for either total score or subscale scores.

Behavioral categorization

Target behaviors identified in each SCED were categorized into one of 5 categories based on a tool designed to assess challenging behaviors.[18] The 5 categories included the following: (1) verbal aggression; (2) physical aggression against objects, self, others; (3) inappropriate behaviors (sexual, social, unwanted); (4) self-regulation deficits (initiation, perseveration, repetition, maintenance); or (5) other. Cases that targeted more than 1 behavior were categorized on the basis of the most challenging behavior.

Intervention categorization

Intervention techniques were also categorized into one of 3 categories–antecedent, consequent, or both (ie, a combination of antecedent and consequent). Antecedent interventions targeted triggers to challenging behaviors, while consequent interventions targeted a response or reaction to challenging behaviors. Examples of antecedent interventions included stress inoculation training, training in self-regulation or self-control, and use of verbal or visual cues. Examples of consequent interventions included differential reinforcement of low rates of behavior, differential reinforcement of appropriate behavior, or other forms of operant conditioning. Examples of studies that employed both antecedent and consequent interventions included environmental management with operant conditioning-based compliance training, or goal-setting, and extinction and token exchange.

RESULTS

The list of the 35 published articles meeting inclusion criteria for this study can be found in Table B of SDC (available at: http://links.lww.com/JHTR/A563). Articles spanned years 1984-2020. Most articles were

TABLE 2 *Criteria for effect size determination*

Determination	Criteria
Large effect	– if change of level, there is at least a small gap between the levels of difference between the phases without any overlap – if change in trend/slope, the direction of the trend changes (eg, if slope is downward in baseline phase, then slope is upward in intervention) – if change in variability, the direction goes from large variability to flat
Medium effect	– if change of level, there is a small amount of overlap between the levels of different phases – if change in trend/slope, a flat slope changes to upward/downward by at least 15° and the change is easily visualized, but not dramatic – if change in variability, the direction goes from large variability to flat
Small effect	– if change in level, there is medium to large amount of overlap between the levels of different phases – if change in trend/slope, the change is barely perceptible – if change in variability, the direction goes from large variability to medium or small
No effect	No changes noted in level, trend/slope, or variability

published prior to 2005, before scrutiny increased and stringent criteria became a focus for risk reduction to improve internal validity in single-case methodology.

Behaviors and interventions

Most frequently treated behavior category was inappropriate behavior, with 42% (19/45) falling in this category, followed by self-regulation, with 27% (12/45) falling in this category. Figure 2 illustrates distribution of behavioral categories across the 3 intervention categories and relative to assessed level of internal validity and effect size.

Interventions were very closely split between antecedent and consequent categories. Antecedent interventions comprised 40% of cases (18/45), while 42% (19/45) implemented consequent interventions. The remaining 8 studies (18%) implemented interventions that combined antecedent and consequent elements.

Effect size

Visual analysis of individual cases revealed effect sizes ranging from no effect to large effect. The largest number of cases were assessed as having a medium effect size (49%; 22/45), with the next largest group of cases assessed as having a large effect size (40%; 18/45). Figure 2 illustrates the distribution of cases for effect size, with 87% (39/45) of cases falling toward the right of the graph and within the medium or large effect columns. These cases were relatively equally distributed across intervention categories and were inclusive of all behavior categories.

RoBiNT Scale

See Table C in SDC (available at: http://links.lww.com/JHTR/A564) for details. The RoBiNT Scale total scores for the 45 cases ranged from 9 to 20, with an overall average of 13.8. Average and range of scores for all 45 cases for the Internal Validity subscale were 4.58 and 2-12, respectively (possible maximum score = 14). Average and range of scores for all 45 cases for the External Validity subscale were 9.42 and 7-14, respectively (possible maximum score = 16). No case achieved maximum scores for either internal or external validity. Despite the individualized approach to rating each case within studies, different cases within an article tended to achieve similar ratings (eg, see cases #4 and #5 in Table B of SDC, available at: http://links.lww.com/JHTR/A563).

Table 3 lists the individual items that comprise the Internal Validity subscale of the RoBiNT Scale and the number of cases that scored 0 (does not meet), 1 (partially meets), or 2 (fully meets) for that specific item. The majority of cases with large and medium effect sizes were rated as 0 (does not meet) criteria for a number of items. Randomization, blinding of both provider and assessor, and treatment adherence were consistently rated low across all cases regardless of effect size. Internal validity items receiving higher ratings included phase count, behavior sampling, and interrater agreement. Figure 2 visually illustrates the relationship between effect size and Internal Validity subscale scores. The majority of cases fall below a subscale score of 7, which is the rating score that could be obtained if a case partially met criteria for all internal validity items.

Data synthesis

Three cases reviewed revealed null effects, with one achieving the highest internal validity score in the set (score 9),[19] and the other 2 among the lowest internal validity scores (scores 2, 3).[20,21] Large and medium effect sizes were also associated with average scores in the low range for internal validity. Despite less than optimal ratings for internal validity across all levels of effect,

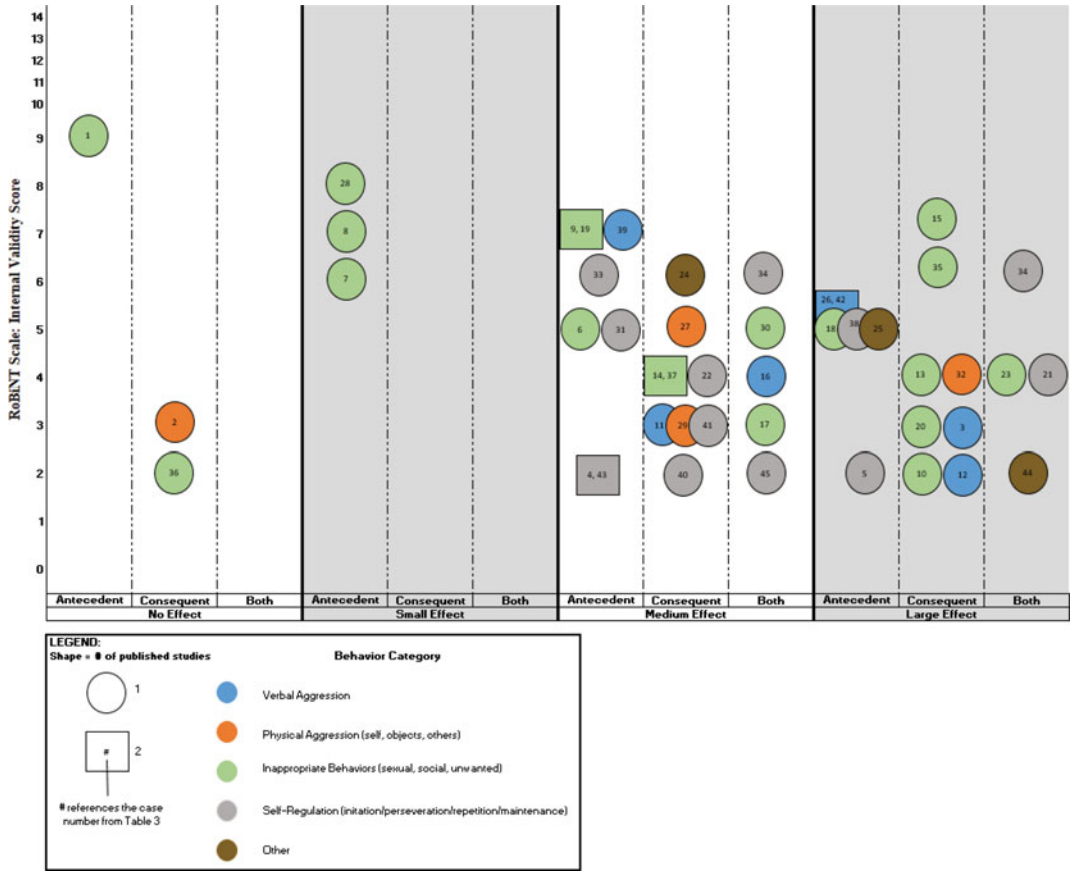

Figure 2. Scatterplot of RoBiNT Scale internal validity subtotal scores for behavioral interventions assessed for effect size.

the data revealed some positive findings. Case studies treating inappropriate behaviors and verbal aggression received the highest ratings for internal validity. Selecting an arbitrary cutoff such as a RoBiNT internal validity score of 7 or greater, which represents an average rating for partially met criteria, a medium effect size was reported for antecedent treatment of verbal aggression and null to medium effect sizes for inappropriate behaviors. A large effect size was reported for a consequent treatment of inappropriate behavior (differential reinforcement used to reduce verbal aggression, sexual comments, and swearing as described by Alderman and Knight[20]; see Table B of SDC, available at: http://links.lww.com/JHTR/A563).

DISCUSSION

Clinicians are encouraged to continue to contribute evidence from single cases. Despite flaws in internal validity identified in this study, the majority of studies were published prior to 2010. Only 4 studies published since 2010 met the inclusion criteria, and only one was published in 2020. The few studies since 2010 suggest a downward trend in single-case publications. This is unfortunate since multiple replications of well-designed single cases are needed to provide evidence for clinical guideline development. Behavioral interventions identified by this study with the greatest promise for treating challenging behaviors can be employed in clinical settings and thus prioritized for additional investigation. Specifically, differential reinforcement used in consequential contexts for inappropriate behaviors demonstrated large effect while contracting, skills training, or cueing used in antecedent contexts for verbal aggression demonstrated medium effect. Each led to improvements in challenging behaviors, indicating these interventions warrant consideration for use in clinical settings and prioritized for SCED investigation. Our study therefore supports a call for more publications employing SCED methodology.

Consistent with Heinicke and Carr,[2] current findings found significant weakness in quality of SCED methodology used to treat challenging behaviors within the published literature. While robust large and medium effect sizes were found for many cases, indicating significant clinical improvement in target behavior following behavioral intervention, the ability to draw causal inference between intervention and target behavior is limited because of low performance on internal validity measures. This is not to say that behavioral

TABLE 3 *Frequency case counts for each of the RoBiNT Scale internal validity items grouped by effect size of treatment intervention*

RoBiNT Scale internal validity item	Large effect cases			Medium effect cases			Small effect cases			No effect cases		
	0 Does not meet	1 Partially meets	2 Fully meets	0 Does not meet	1 Partially meets	2 Fully meets	0 Does not meet	1 Partially meets	2 Fully meets	0 Does not meet	1 Partially meets	2 Fully meets
1. Phase count	1	6	10	3	8	11	0	2	1	0	0	3
2. Randomization	17	0	0	22	0	0	2	1	0	2	0	1
3. Behavior sampling	3	7	7	1	10	11	0	2	1	0	1	2
4. Blinded provider	16	1	0	21	1	0	0	3	0	3	0	0
5. Blinded assessor	12	5	0	11	10	1	0	3	0	2	1	0
6. Interrater agreement	5	10	2	8	10	4	0	2	1	2	0	1
7. Treatment adherence	16	1	0	21	0	1	3	0	0	3	0	0

interventions are ineffective. On the contrary, large and medium effect sizes indicate clinically significant improvements in target behaviors across a number of cases employing a variety of behavioral interventions. What is at issue is the ability to specifically link those improvements to the treatments reported when risk of bias within experimental conditions is not adequately controlled.[2,10,13] Without adequate control of possible sources of bias and subsequent ability to infer causal relationships, the inclusion of evidence derived from SCED methodology may remain limited and possibly eliminated as a source of evidence for practice guideline development. Lack of a strong relationship between RoBiNT internal validity scores and treatment effect size may provide some reassurance about the causal relationship. However, given that even the best studies had relatively low scores, this reassurance is limited.

Based on RoBiNT Scale ratings, randomization, blinding of both provider and assessor, and treatment adherence were all assessed in this study as not meeting any part of the defined criteria in the majority of cases. Interestingly, these were the specific items added to the RoBiNT Scale when it was redesigned from the SCED Scale.[11] These items were identified as the most powerful design elements that could bolster the strength of SCED studies and thus fidelity of data and conclusions.[10] Many cases rated in the current study predate the 2013 publication of the RoBiNT Scale. It is therefore not surprising the criteria for these items were not met. Future SCED studies should strongly consider implementation of these design features, particularly if the goal of publishing treatment using SCED methodology is to provide causal evidence for the effectiveness of the intervention.

A further challenge to deriving causal evidence from SCED methodology is the very reason SCED methodology is used following TBI—to clinically intervene and treat a very specific and usually challenging behavior or a set of challenging behaviors. Environments and persons with TBI are highly diverse, and challenging behaviors following moderate to severe TBI are complex, multifactorial, and individually nuanced. The purpose of the intervention is therefore individualistic, tailored to the need of the individual with the identified target behavior. In clinical settings, interventions are adapted to particular individuals with particular target behaviors. This is in contrast to an approach where one might seek to replicate the efficacy of the same treatment across multiple diverse patients (eg, multiple baseline across individuals or separate replication designs), which would allow greater generalization of the conclusions to future patients. For a clinical provider charged with resolving aberrant behavior quickly, conducting an experimental design that controls for all threats to internal validity is unlikely and possibly ill-advised. Once a challenging

behavior is reduced or eliminated, most providers would be appropriately reluctant to "reverse" the improvement solely to provide efficacy evidence for the intervention. Even if specific designs can be selected to avoid the latter conundrum and even if a well-implemented SCED with robust efforts to meet internal validity criteria could lead one to conclude cause and effect, the conclusion may remain solely tied to the one case. Meeting the robust criteria for external validity would next be required to generalize the cause and effect beyond the single case.

Finally, an additional concern is the risk of publication bias. If one accepts the causal role of behavioral treatments in the improvements reported in these cases, the findings suggest medium to large effects from many treatments applied to many behaviors. Without knowing how many null effects go unpublished, one cannot assume these proportions reflect the actual efficacy of behavioral treatments applied in the field. The publishing of null effects may be necessary for valid evaluation of evidence of efficacy when SCED methodology is employed.

Design elements presenting the greatest challenge are randomization, blinding, and treatment adherence. Practical suggestions for clinicians to consider to improve internal validity and reduce the risk of bias include the following: (1) designating personnel blinded to patient treatment plans and trained to assess performance as requested by clinicians (ie, ready-made assessors) or credentialed behavior analysts; (2) utilizing students/trainees to deliver treatment once designed by primary clinicians; (3) establishing standard protocols and checklists to ensure efficiency and fidelity of data collection; (4) establishing routine, standardized minimum data collection through electronic record systems to provide reliable behavior tracking and immediate baseline data; and (5) implementing treatments and measuring performance across multiple settings by multiple providers, in structured and unstructured settings and at various times of the day. Consider implementing the Model for Assessing Treatment Effect (MATE) developed by Tate and colleagues,[10,22] a framework to guide successive increase in scientific rigor within clinical practice settings. Finally, consider developing libraries of behavior programs targeting specific categories of challenging behaviors so that successive patients are treated with minor variations of standardized treatment approaches, thus allowing more systematic replication and greater ability to infer generalization.

Developing systematic integration of behavior data collection and program planning can aid clinicians in efficiently delivering treatment while potentially contributing to evidence on its effectiveness. Rehabilitation programs face mounting influence from external stakeholders to move patients through a broad continuum of care. Administrative policies and protocols that support assessment, data collection, and program implementation can be key to producing outcomes that meet criteria for methodological rigor. With proper integration of policies and procedures, standardization of protocols, and establishing minimum standardized data sets, behavior treatment can be efficiently captured and data analyzed for the identification of effective interventions.

Limitations of the study

Selected articles and subsequent individual cases resulted from a very specific set of inclusion criteria. The study focused solely on moderate to severe TBI and only on cases that involved those who were 16 years or older at the time of injury. The current results may not be applicable to individuals treated following concussions, those with mild and complex mild TBI, and individuals who sustained TBI prior to the age of 16 years. Cases were excluded if a specific number of phases were not completed and a specific number of behavior sampling data points were not collected. These criteria may have inadvertently eliminated SCED studies with other strong design features. However, these criteria were required to provide the minimum data from which an effect size could be determined. The study used the RoBiNT Scale to assess the quality of SCED features for each case. Despite the RoBiNT Scale having a nonbinary scoring system with well-crafted criteria outlined and examples provided, a formal manual, and training workshops, the scoring criteria require judgment on the part of raters and thus can be subjective. This study managed this limitation through the use of multiple raters and consensus discussion. Despite this, the subjective nature of scoring may not have been entirely eliminated. Finally, the categorization and classification of target behaviors and intervention allowed for ease in visual analysis of trends in the data but details may have been lost during data reduction.

CONCLUSION

Identifying a one-to-one match between evidence-based behavioral interventions and challenging behaviors that can then generalize across all individuals and settings is unlikely an achievable goal. Practitioners who are well versed in behavior modification techniques can certainly contribute to developing more generalizable and practical guidelines by introducing stronger studies with regard to internal and external validity. A more realistic goal may be to request that single-subject publication submissions adhere to operationally defined standards such as SCED versus single-subject methodology versus case study. Alternatively, establishing minimally acceptable internal and external validity scores, such as what the RoBiNT Scale can provide, would improve the quality of evidence for SCED

publications. Setting minimum acceptable standards could then provide clinical providers with quick reference guides on strength of evidence for a range of interventions implemented for a range of target behaviors. In addition, no matter how strong the internal validity, the report of a single case establishes that treatment *can* be effective, but it provides no information on the likelihood that treatment will be similarly effective in other patients. To move from SCED methodology to generalizable conclusions will require practitioners to apply the same treatment to multiple patients, achieving similar effects, to define an algorithm by which a given treatment can be individually tailored, and to demonstrate that individualized treatments, despite their specific differences, deliver similar treatment impact across multiple patients.

REFERENCES

1. Brain Injury Association of America. Guidelines for the rehabilitation and chronic disease management of adults with moderate to severe brain injury. Accessed March 13, 2021. https://www.biausa.org/professionals/research/treatment-guidelines-project
2. Heinicke MR, Carr JE. Applied behavior analysis in acquired brain injury rehabilitation: a meta-analysis of single-case design intervention research. *Behav Interv*. 2014;29(2):77–105. doi:10.1002/bin.1380
3. Gurdin SL, Huber SA, Cochran CR. A critical analysis of data-based studies examining behavioral interventions with children and adolescents with brain injuries. *Behav Interv*. 2005;20(1):3–16. doi:10.1002/bin.172
4. Ylvisaker M, Turkstra L, Coehlo C, et al. Behavioural interventions for children and adults with behavior disorders after TBI: a systematic review of the evidence. *Brain Inj*. 2007;21(8):769–805. doi:10.1080/02699050701482470
5. Cattelani R, Zettin M, Zoccolotti P. Rehabilitation treatments for adults with behavioral and psychosocial disorders following acquired brain injury: a systematic review. *Neuropsychol Rev*. 2010;20(1):50–85. doi:10.1007/s11065-009-9125-y
6. Horner RH, Carr EG, Halle J, et al. The use of single-subject research to identify evidence-based practice in special education. *Except Child*. 2005;71(2):165–179. doi:10.1177/001440290507100203
7. Kratochwill TR, Hitchcock J, Horner RH, et al. *Single-Case Designs Technical Documentation*. What Works Clearinghouse; 2010. Accessed May 13, 2020. http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf
8. Kratochwill TR, Hitchcock J, Horner RH, et al. Single-case intervention research design standards. *Remedial Spec Educ*. 2013;34(1):26–38.
9. Mozzoni MP. Applied behavior analysis evaluation strategies and neurorehabilitation. *Brain Inj Prof*. 2008;5:28–30.
10. Tate RL, Perdices M. *Single-Case Experimental Designs for Clinical Research and Neurorehabilitation Settings*. Routledge; 2019.
11. Tate RL, Perdices M, Rosenkoetter U, et al. 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. *Neuropsychol Rehabil*. 2013;23(5):619–638.
12. Howick J, Chalmers I, Glasziou P, et al. *The 2011 Oxford CEBM Evidence Table (Introductory Document)*. Oxford Centre for Evidence-Based Medicine; 2011. Accessed March 13, 2021. http://www.cebm.net/index.aspx?o=5653
13. Tate RL, Rosenkoetter U, Wakim D, et al. *The Risk of Bias in N-of-1 Trials (RoBiNT) Scale: An Expanded Manual for the Critical Appraisal of Single-Case Reports*. The PsycBITE Group; 2015.
14. Tate RL, McDonald S, Perdices M, et al. Rating the methodological quality of single-subject designs and n-of-1 trials: introducing the single-case experimental design (SCED) scale. *Neuropsychol Rehabil*. 2008;18(4):385–401. doi:10.1080/09602010802009201
15. Parsonson BS, Baer DM. The analysis and presentation of graphic data. In: Kratochwill TR, ed. *Single-Subject Research: Strategies for Evaluating Change*. Academic Press; 1978:101–162.
16. Todman JB, Dugard P. *Single-Case and Small-N Experimental Designs: A Practical Guide to Randomization Tests*. Lawrence Erlbaum Associates; 2001.
17. Ledford JR, Lane JD, Severini KE. Systematic use of visual analysis for assessing outcomes in single case design studies. *Brain Impair*. 2017;19:4–17. doi:10.1017/BrImp.2017.16
18. Kelly G, Kremer D, Simpson G, Kremer P, Martin C. The Overt Behavior Scale (OBS): a tool for measuring challenging behaviors following ABI in community settings. *Brain Inj*. 2006;20(3):307–319.
19. Aeschleman S, Imes C. Stress inoculation training for impulsive behaviors in adults with traumatic brain injury. *J Ration Emot Cogn Behav Ther*. 1999;17(1):51–65. doi:10.1023/A:1023073130972
20. Alderman N, Knight C. The effectiveness of DRL in the management and treatment of severe behavior disorders following brain injury. *Brain Inj*. 1997;11(2):79–101. doi:10.1080/026990597123683
21. Turner JM, Green G, Braunling-McMorrow D. Differential reinforcement of low rates of responding (DRL) to reduce dysfunctional social behaviors of a head injured man. *Behav Interv*. 1990;5(1):15–27. doi:10.1002/bin.2360050103
22. Tate RL, Arid V, Taylor C. Bringing single-case methodology into the clinic to enhance evidence-based practices. *Brain Impair*. 2012;13(3):347–359. doi:10.1017/BrImp.2012.32
23. Arco L. Neurobehavioural treatment for obsessive-compulsive disorder in an adult with traumatic brain injury. *Neuropsychol Rehabil*. 2008;18(1):109–124. doi:10.1080/09602010701656706
24. Arco L, Cohen L, Geddes K. Verbal self-regulation of impulsive behavior of persons with frontal lobe brain injury. *Behav Ther*. 2004;35(3):605–619. doi:10.1016/S0005-7894(04)80034-0
25. Brotherton F, Thomas L, Wisotzek MS, Milan M. Social skills training in the rehabilitation of patients with traumatic closed head injury. *Arch Phys Med Rehabil*. 1988;69(10):827–832.
26. Dixon M, Guercio J, Falcomata T, et al. Exploring the utility of functional analysis methodology to assess and treat problematic verbal behavior in persons with acquired brain injury. *Behav Interv*. 2004;19(2):91–102. doi:10.1002/bin.155
27. Dixon MR, Horner MJ, Guercio J. Self-control and the preference for delayed reinforcement: an example in brain injury. *J Appl Behav Anal*. 2003;36(3):371–374. doi:10.1901/jaba.2003.36-371
28. Gajar A, Schloss PJ, Schloss CN, Thompson CK. Effects of feedback and self-monitoring on head trauma youths' conversation skills. *J Appl Behav Anal*. 1984;17(3):353–358. doi:10.1901/jaba.1984.17-353
29. Gouvier WD, Richards JS, Blanton PD, Janert K, Rosen LA, Drabman RS. Behavior modification in physical therapy. *Arch Phys Med Rehabil*. 1985;66(2):113–116.

30. Hegel MT. Application of a token economy with a noncompliant closed head-injured male. *Brain Inj*. 1988;2(4):333–338. doi:10.3109/02699058809150904
31. Johnson K, Davis PK. A supported relationship intervention to increase the social integration of persons with traumatic brain injuries. *Behav Modif*. 1998;22:502–528.
32. Kirsch NL, Shenton M, Spirl E, Simpson R, LoPresti E, Schreckenghost D. An assistive-technology intervention for verbose speech after traumatic brain injury: a single case study. *J Head Trauma Rehabil.* 2004;19(5):366–377.
33. Knight C, Rutterford NA, Alderman N, Swan LJ. Is accurate self-monitoring necessary for people with acquired neurological problems to benefit from the use of differential reinforcement methods? *Brain Inj.* 2002;16(1):75–87. doi:10.1080/02699050110082188
34. Lane IM, Wesolowski MD, Burke WH. Teaching socially appropriate behavior to eliminate hoarding in a brain injured adult. *J Behav Ther Exp Psychiatry.* 1989;20(1):79–82.
35. Lane-Brown A, Tate R. Evaluation of an intervention for apathy after traumatic brain injury: a multiple-baseline, single-case experimental design. *J Head Trauma Rehabil.* 2010;25(6):459–469.
36. Lennox DB, Brune P. Incidental teaching for training communication in individuals with traumatic brain injury. *Brain Inj.* 1993;7(5):449–454.
37. Maki AL, Rudrud EH, Schulze KA, Rapp JT. Increasing therapeutic exercise participation by individual with acquired brain injury using self-recording and reinforcement. *Behav Interv.* 2008;23(2):75–86. doi:10.1002/bin.255
38. O'Reilly MF, Green G, Braunling-McMorrow D. Self-administered written prompts to teach home accident prevention skills to adults with brain injuries. *J Appl Behav Anal.* 1990;23(4):431–446. doi:10.1901/jaba.1990.23-431
39. Pace GM, Ivancic MT, Jffereson G. Stimulus fading as treatment for obscenity in a brain-injured adult. *J Appl Behav Anal.* 1994;27(2):301–305. doi:10.1901/jaba.1994.27-301
40. Persel CS, Persel CH, Ashley MJ, Krych DK. The use of noncontingent reinforcement and contingent restraints to reduce physical aggression and self-injurious behavior in a traumatically brain injured adult. *Brain Inj.* 1997;11(10):751–760. doi:10.1080/026990597123124
41. Schloss PJ, Thompson CK, Gajar AH, Schloss CN. Influence of self-monitoring on heterosexual conversational behaviors of head trauma youth. *Appl Res Ment Retard.* 1985;6:269–282. doi:10.1016/0270-3092(85)90001-3
42. Slifer KJ, Cataldo MD, Babbit RL, Kane AC, Harrison KA, Cataldo MF. Behavior analysis and intervention during hospitalization for brain trauma rehabilitation. *Arch Phys Med Rehabil.* 1993;74(8):810–817. doi:10.1016/0003-9993(93)90006-v
43. Slifer KJ, Tucker CL, Gerson AC, et al. Antecedent management and compliance training improve adolescents' participation in early brain injury rehabilitation. *Brain Inj.* 1997;11(12):877–890.
44. Sohlberg MM, Sprunk H, Metzelaar K. Efficacy of an external cueing system in an individual with severe frontal lobe damage. *Cogn Rehabil.* 1988;6:36–42.
45. Stewart I, Alderman N. Active versus passive management of post-acquired brain injury challenging behaviour: a case study analysis of multiple operant procedures in the treatment of challenging behaviour maintained by negative reinforcement. *Brain Inj.* 2010;24(13/14):1616–1627. doi:10.3109/02699052.2010.523050
46. Tasky KK, Rudrud EH, Schulze KA, Rapp JT. Using choice to increase on-task behavior in individuals with traumatic brain injury. *J Appl Behav Anal.* 2008;41(2):261–265. doi:10.1901/jaba.2008.41-261
47. Tate RL, Wakim D, Sigmundsdottir L, Longley W. Evaluating an intervention to increase meaningful activity after severe traumatic brain injury: a single-case experimental design with direct intersubject and systematic replications. *Neuropsychol Rehabil.* 2020;30(4):641–672. doi:10.1080/09602011.2018.1488746
48. Travis R, Sturmey P. Functional analysis and treatment of the delusional statements of a man with multiple disabilities: a four-year follow-up. *J Appl Behav Anal.* 2010;43(4):745–749. doi:10.1901/jaba.2010.43-745
49. Wesolowski MD, Zencius A, Burke WH. Effects of feedback and behavior contracting on head trauma persons' inappropriate sexual behavior. *Behav Interv.* 1993;2:89–96.
50. Wesolowski MD, Zencius AH, Rodriguez IM. Mini-breaks: the use of escape on a fixed-time schedule to reduce unauthorized breaks from vocational training sites for individuals with brain injury. *Behav Interv.* 1999;14(3):163–170.
51. Zencius AH, Wesolowski MD. Reducing verbal aggression in adults with brain injury. *Behav Interv.* 1991;6(3):155–164. doi:10.1002/bin.2360060302
52. Zencius A, Wesolowski MD, Burke WH. Comparing motivational systems with two non-compliant head-injured adolescents. *Brain Inj.* 1989;3(1):67–71. doi:10.3109/02699058909008076
53. Zencius AH, Wesolowski MD, Burke WH. The use of a visual cue to reduce profanity in a brain injured adult. *Behav Interv.* 1990;3:143–147.
54. Zencius AH, Wesolowski MD, Burke WH, McQuade P. Antecedent control in the treatment of brain-injured clients. *Brain Inj.* 1989;3(2):199–205. doi:10.3109/02699058909004553