Program Evaluation of a Specialized Treatment Home for Adults with Severe Challenging
Behaviour

Thurka Thillainathan, BSc (Hons)

Department of Applied Disability Studies, Brock University

Submitted in partial fulfillment of the requirements for the degree of

Master of Arts in Applied Disability Studies

Faculty of Social Sciences, Brock University

St. Catharines, Ontario

Abstract

Individuals with intellectual and developmental disabilities who engage in severe challenging behaviour comprise 5-10% of the population and experience significant limitations in meaningfully participating in everyday activities due to associated risks (e.g., substantial injury to self and others, extreme property destruction, outward physical aggression targeting others). Unfortunately, research featuring adult participants who engage in severe challenging behaviour is relatively scare compared to child participants. Further, challenging behaviour literature tends to emphasize efficacy (e.g., Does the intervention work?) more often than effectiveness (e.g., Does the intervention work in *real world* settings?). The current project thus was a systematic program evaluation conducted to evaluate the effectiveness of a comprehensive behavioural treatment package at reducing severe challenging behaviour and generating adaptive skills in adults with intellectual and developmental disabilities. A hybrid nonexperimental consecutive case series design was employed featuring all participants ($n = 8$) who experienced the treatment package, regardless of their *success*. The results depicted primarily therapeutic outcomes with a substantial decrease in challenging behaviour from baseline to intervention for majority of participants ($n = 5$) and an increase in adaptive behaviour (i.e., number of mastered skills targets) for participants ($n = 7$) across the intervention condition. Treatment fidelity suggests frontline staff were largely implementing the interventions as intended ($M = 84\%$, range 82-90%). Social validity surveys administered to participants, caregivers, and case managers provide support for the acceptability of treatment goals, procedures, and effects. Project limitations, clinical considerations, and future directions are discussed.

Keywords: intellectual and developmental disabilities, severe challenging behaviour, program evaluation, effectiveness, consecutive case series

Acknowledgements

It is my absolute honour to be able to share this work that was made possible by the unwavering support of my incredible mentors, colleagues, family, and friends.

To my advisor, Dr. Alison Cox, who inspired me to go *above and beyond* in everything that I do (even though it first started off as a requirement for Statistics assignments) and whose innovative thinking I could only hope to match someday, thank you for all the time and effort you put in to make sure that not only my thesis, but that also my graduate journey was a successful one.

To the partnering agency, the clinical team, the frontline staff, the participants, and my two research assistants, James Padua, and Maria Sushcheva, thank you for welcoming me to the team and for making the data collection process as seamless as possible.

To my colleagues in Dr. Cox's lab and the friends I made along the way, thank you for being the support system I could turn to over the past two years. I am so lucky to be a part of such an outstanding group of individuals.

To my family and friends, and especially my mom and dad, thank you for believing in me. These barriers were only broken because I had your shoulders to stand on.

Finally, I would like to dedicate this thesis in loving memory to my beloved dog, Chico, who passed away unexpectedly earlier this year.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ADHD	Attention Deficit Hyperactivity Disorder
ASD	Autism Spectrum Disorder
BCBA	Board-Certified Behaviour Analyst
BPD	Borderline Personality Disorder
BSP	Behaviour Support Plan
CCCS	Consecutive Controlled Case Series
DAS	Daily Activity Schedule
DAS-Q	Daily Activity Schedule – Quality Rating Scale
DR	Differential Reinforcement
DRA	Differential Reinforcement of Alternative Behaviour
DTT	Discrete Trial Training
FAST	Functional Assessment Screening Tool
FCT	Functional Communication Training
IOA	Interobserver Agreement
NCR	Noncontingent Reinforcement
NOS	Not Otherwise Specified
QABF	Questions About Behavioural Function
RIRD	Response Interruption and Redirection
SAP	Skill Acquisition Program
SCED	Single-Case Experimental Design
SIB	Self-injurious Behaviour
TIC	Treatment Integrity Checklist

Program Evaluation of a Specialized Treatment Home for Adults with Severe Challenging Behaviour

Challenging behaviour can be defined as "behaviour that is aggressive or injurious to self or to others or that causes property damage or both and that limits the ability of the person with a developmental disability to participate in daily life activities and in the community or to learn new skills or that is any combination of them" (Ministry of Community and Social Services [MCSS], 2008, s.15(2)). Approximately 50% of individuals with intellectual and developmental disabilities engage in challenging behaviour and approximately 5%–10% of these individuals' challenging behaviour may be described as *severe* (Bowring et al., 2017; Emerson et al., 2001; Lowe et al., 2007). Severe challenging behaviour can take several forms, including aggression, property destruction, and self-injury, although the term *severe* has been applied inconsistently in the literature. That is, it has been described using subjective terms (i.e., dangerous) as well as by referencing topography, injury, frequency, and/or risk to name a few (Bonner & Borrero, 2019; Fahmie & Iwata, 2011; Lowe et al., 2007; Oropeza et al., 2018; Poppes et al., 2010). Support for individuals engaging in severe challenging behaviours may cost the healthcare system (or related ministries) more than $3000 per day and upwards of $600,000 per year, per person (Butterill et al., 2009). Importantly, this reference is dated and, therefore, it is possible that associated support costs today (2023) likely exceed those outlined in the reference from 2009.

Applied Behaviour Analytic Interventions for Severe Challenging Behaviour

Typically, treatment approaches for challenging behaviour in persons with intellectual and developmental disabilities include pharmacological interventions (i.e., psychotropic medications; Valdovinos, 2019) and non-pharmacological interventions (i.e., applied behaviour analysis; Kurtz et al., 2021). Applied behaviour analysis is among one of the non-

pharmacological intervention options, and has been described as an objective, scientific approach used to determine environmental variables that may reliably predict socially significant behaviour in order to implement evidence-based strategies that result in therapeutic behavioural change (Cooper et al., 2020). Current research evaluating interventions informed by this discipline place emphasis on efficacy. That is, there exists substantial research examining treatment effects under highly controlled conditions in order to demonstrate a causal relation between the treatment and behaviour change (Ghaemmaghami et al., 2021; Singal et al., 2014). Unfortunately, effectiveness research has tended to lag behind (Ghaemmaghami et al., 2021; Gover et al., 2019). Compared to efficacy research, relatively few studies examine the effect of the treatment under real-world circumstances to demonstrate the generality, acceptability, and clinical utility of the treatment (Chartier & Feldman, 2015; Phillips et al., 2017; Singal et al., 2014). There are several behaviour analytic intervention options that have been employed to reduce challenging behaviour, though few have an established effectiveness evidence-base (Ghaemmaghami et al., 2021).

Function-Based Interventions

Function-based interventions, defined as interventions that directly address consequences that are maintaining a challenging behaviour (e.g., access to preferred items/activities, escape from instructions), are predominantly implemented within the field of applied behaviour analysis to decrease challenging behaviour (Geiger et al., 2010; Wilder et al., 2019). There are many different function-based *protocols* that can be enacted. However, it may be reasonable to state that, broadly speaking, there are three overarching categories of function-based treatments: (1) modification of motivating operations (e.g., altering the environment to decrease the value of access to functional reinforcer; Carr et al., 2009; Şenkal et al., 2023); (2) extinction (e.g.,

withholding reinforcement for functional reinforcer in response to challenging behaviour; Iwata et al., 1994); and (3) differential reinforcement (e.g., withholding reinforcement for functional reinforcer in response to challenging behaviour plus teaching alternative appropriate behaviour to access functional reinforcer; DR; Wilder et al., 2019). Best-practice in behaviour analysis indicate that the results of a comprehensive functional behaviour assessment will inform which category(ies) to select (Melansen & Fahmie, 2023).

Noncontingent Reinforcement. Noncontingent reinforcement (NCR) is a common behaviour analytic intervention. Typically, implementers deliver reinforcement on a predetermined fixed- or variable-time schedule independent of the individual's behaviour (Cooper et al., 2020). NCR has been shown to be efficacious in treating challenging behaviour maintained by social and automatic reinforcement, however, marked, sustained success is often necessitated by additional components, such as extinction (Lindberg et al., 2003; Saini et al., 2017; Vollmer et al., 1995). In evaluating the effectiveness of NCR, Phillips et al. (2017) examined 27 consecutive applications of NCR with participants ranging in age from 5 to 33 years and reported outcomes regardless of treatment success. Their analysis found that NCR was effective in reducing challenging behaviour maintained by social reinforcement in 14 of 15 applications when either functional or alternative reinforcers were used. However, additional treatment components were required to supplement NCR when challenging behaviour was maintained by automatic reinforcement in order to observe clinically significant outcomes. Though it appeared that the effectiveness of NCR was explored in this study, the authors described several important limitations. First, a restricted sample was included, which may not be representative of the broader clinical population. Second, brief assessment sessions were conducted by highly trained staff in a tightly controlled environment. As such, the study

personnel and settings were not reflective of typical treatment settings thereby limiting the study's ecological validity (Phillips et al., 2017).

Extinction. Extinction has been defined as withholding reinforcement for a behaviour that was previously reinforced to decrease the occurrence of that behaviour (Cooper et al., 2020). Behaviour analytic best-practice advises against applying extinction in isolation for the treatment of challenging behaviour (Fisher et al., 2004; Newcomb & Hagopian, 2018). These recommendations are based on hypotheses suggesting an increased likelihood of secondary effects due to the extinction procedure being implemented in isolation, including emotional reactivity, extinction bursts (i.e., momentary increases in the occurrence, magnitude, and/or duration of behaviour put on extinction), response variation (i.e., occurrence of diverse and novel forms of behaviour during the extinction process), and spontaneous recovery (i.e., reappearance of behaviour put on extinction after it was decreased or eliminated; Cooper et al., 2020; Lerman & Iwata, 1995; Newcomb & Hagopian, 2018). As such, contemporary behaviour analysis always showcases extinction as one element in a treatment package that may offset competing contingencies (Newcomb & Hagopian, 2018).

Escape-Extinction. Escape-extinction refers to the application of extinction for behaviours maintained by negative reinforcement. That is, the perceived aversive stimulus (e.g., instructional demand) is *not* removed contingent on the behaviour (e.g., through physical prompting, repeating task directions, blocking escape attempts) thereby precluding escape from the individuals' perceived aversive situation (e.g., completing a task; Cooper et al., 2020; Iwata et al., 1990). Over the years, researchers have demonstrated escape-extinction as an efficacious strategy to decrease severe challenging behaviour, such as self-injurious behaviour (SIB), physical aggression, and property destruction (e.g., Anderson & Long, 2002; Iwata et al., 1990;

Lalli et al., 1994; Tereshko & Sottolano, 2017). Although the efficacy of escape-extinction is well-established in the literature, research examining the implementation of this procedure have primarily focused on children with developmental disabilities (Chazin et al., 2022; Dart et al., 2018). It is possible that a learner's size and the intensity with which they engage in challenging behaviour may be one reason that research has largely been conducted with children. Although, there have been methods describing how one might safely enact extinction in the context of supporting adult clients who engage in severe challenging behaviour. For example, *safe extinction* has been described as a form of escape-extinction wherein contingent physical interventions and/or mechanical restraints are utilized to safely enact extinction to treat escape-maintained behaviours in large adolescents or adults (Salameh & Linder, 2016). Currently, there is a relative dearth of research evaluating escape-extinction (in combination with reinforcement-based procedures) implementation with adults or large adolescents who engage in severe challenging behaviour by natural change agents in treatment contexts (i.e., *effectiveness;* Geiger et al., 2010; Mcconnachie & Carr, 1997).

Differential Reinforcement. Extinction applied alongside reinforcement-based procedures can facilitate a simultaneous increase in adaptive behaviour (e.g., skills and socially appropriate behaviours) and decrease in challenging behaviour (Cengher et al., 2020). Such interventions are referred to as DR, wherein reinforcement is provided for one response class, but withheld for another. In the context of reducing challenging behaviour, DR involves: (1) contingently providing reinforcement for a behaviour other than the targeted challenging behaviour; and (2) withholding reinforcement for said challenging behaviour (Cooper et al., 2020). There are several variations of DR procedures, including differential reinforcement of alternative behaviour (DRA), differential reinforcement of other behaviour, and differential

reinforcement of low rates of behaviour, that have been used in the treatment of challenging behaviour (Cooper et al., 2020). Recent research suggests DRA has been more extensively researched with respect to severe challenging behaviour in individuals with intellectual and developmental disabilities (Newcomb & Hagopian, 2018).

Functional Communication Training. Functional Communication Training (FCT) may be conceptualized as a variation of DRA. That is, an alternative, functional communication response is reinforced by offering access to the same reinforcer(s) that were originally obtained by engaging in challenging behaviour (Carr & Durand, 1985). FCT with extinction has been demonstrated as efficacious for challenging behaviours of diverse topographies and functions as well as for participants of different ages, intellectual abilities, and diagnoses under tightly controlled conditions and rich reinforcement schedules (Chowdury & Benson, 2011; Petscher et al., 2009). Ghaemmaghami et al. (2021) conducted a quantitative review of the literature to determine whether there exists sufficient empirical evidence to support the effectiveness of FCT; that is, to evaluate the extent to which the implementation of FCT in natural contexts result in general and socially valid changes. They found that while there is impressive evidence supporting the efficacy of FCT involving researcher implementation and rich reinforcement schedules, there has yet to be any applications of FCT involving natural change agents in relevant contexts, in which maintenance, generality, and social validity have been demonstrated (Ghaemmaghami et al., 2021).

Punishment-Based Interventions

In the event that function-based interventions in isolation have failed to produce meaningful challenging behaviour change (e.g., Foxx & Garito, 2007), additional consequence-based procedures that transcend function, such as punishment, may need to be implemented to

produce clinically meaningful outcomes and achieve socially significant goals (Newcomb & Hagopian, 2018). Such cases typically include circumstances in which: (1) the challenging behaviour presents a serious risk of harm (e.g., physical injury; Busch et al., 2018); (2) extinction- and reinforcement-based intervention have not resulted in challenging behaviour reduction to clinically meaningful levels (Behavior Analyst Certification Board, 2020; Manente & LaRue, 2014); and (3) the reinforcer maintaining challenging behaviour cannot be identified with certainty (i.e., through functional analysis) or withheld due to behavioural and/or environmental conditions (Cooper et al., 2020; Lerman & Vorndran, 2002).

Response Blocking. Response blocking refers to immediate physical intervention following the occurrence of challenging behaviour as a means to prevent the completion of the response (Cooper et al., 2020). Response interruption and redirection (RIRD) is a variation of response blocking in which stereotypic behaviour is interrupted at its onset and the individual is redirected to complete high-probability behaviours instead (Cooper et al., 2020). Historically, response blocking and RIRD have been found to be efficacious for behaviour that is automatically maintained (e.g., Ahearn et al., 2007; Fisher et al., 1996; Reid et al., 1993). Importantly, automatically maintained behaviour can take the form of SIB (e.g., face slapping, head hitting, skin picking). Further, DeRosa et al. (2019) compared response blocking and RIRD and found support for the efficacy of both interventions in reducing motor stereotypy. Although these procedures' efficacy is relatively established in the literature, there appears to be a lack of effectiveness research. That is, research evaluating response blocking and RIRD have primarily focused on children (defined as less than 18 years of age) with autism spectrum disorder (ASD) who engage in either motor or vocal stereotypy (Lydon et al., 2013). In addition, most studies have evaluated response blocking and RIRD with near perfect implementation by a highly

trained therapist which may not be feasible in the natural environment with caregivers (e.g., parents, teachers). Further, little is known regarding the social validity and maintenance of the treatment effects under these conditions (Kliebert et al., 2011).

Overcorrection. Overcorrection is a behaviour reduction strategy in which an individual is required to engage in related effortful behaviour contingent on the occurrence of a target behaviour (Cooper et al., 2020; Foxx & Azrin, 1972). Overcorrection consists of one or both of the following components: (1) restitution; and (2) positive practice (Cooper et al., 2020; McAdam & Knapp, 2021). Restitutional overcorrection refers to requiring an individual to restore the environment to a condition better than its original state contingent on the occurrence of a target behaviour (e.g., clean up a mess, repair broken items; Cooper et al., 2020). Positive practice overcorrection is defined as requiring an individual to repeatedly engage in either the correct form of the behaviour or an incompatible behaviour for a predetermined frequency of responses or duration of time contingent on the occurrence of a target behaviour (e.g., colouring with crayons instead of throwing them; Cooper et al., 2020; Peters & Thompson, 2013). Prior research has demonstrated the efficacy of overcorrection in decreasing a variety of topographically diverse behaviours, including vocal and motor stereotypy, SIB, and pica (e.g., Agosta et al., 1980; Anderson & Le, 2011; Matson et al., 1978; Peters & Thompson, 2013). Further, an early review by Miltenberger and Fuqua (1981) found support for the effectiveness of overcorrection. That is, consistent decelerative effects were reported across a wide range of clinical populations (e.g., children with ASD, neurotypical college students, adults with intellectual and developmental disabilities; Miltenberger & Fuqua, 1981). Lending further support for the effectiveness of overcorrection, Halpern and Andrasik (1986) found maintenance of low rates of head-banging behaviour 1 year following the implementation of an overcorrection

intervention and Linscheid et al. (1981) reported greater acceptability of overcorrection procedures by caregivers. Importantly, these references are dated. As such, current best practice recommendations promote the supplementation of punishment procedures, such as overcorrection, with complementary interventions (i.e., antecedent and reinforcement strategies; Cooper et al., 2020). More recently, in line with best practice, overcorrection implemented as one component in a comprehensive treatment package has been found to be efficacious in decreasing inappropriate sexual behaviour in individuals with intellectual and developmental disabilities and stereotypy in children with ASD (Falligant & Pence, 2020; Steinhauser et al., 2021). Further, Williams et al. (2009) reported on treatment outcomes in a series of consecutive cases of 41 individuals with intellectual disability engaging in pica. All individuals experienced the same unique program consisting of: (1) staff training and monitoring; and (2) implementation of a hierarchical, function-based behaviour plan with both differential reinforcement and restrictive procedures (i.e., overcorrection, brief contingent restraint). The results found that the program reduced pica to substantial levels (i.e., 75%–100% reduction), eliminated the use of protective restraints, decreased the application of restrictive behaviour management practices, and facilitated the transfer of a minority of individuals to less restrictive settings (Williams et al., 2009). Continued evaluations of overcorrection in conjunction with reinforcement-based procedures is required for a contemporary understanding of effectiveness (i.e., feasibility, generality, acceptability).

Response Cost. Response cost refers to the contingent loss of positive reinforcers (e.g., generalized conditioned reinforcers, tangibles, time with preferred activities) resulting in the decrease of future occurrences of the target behaviour (Bagwell et al., 2022; Cooper et al., 2020; Rapport & Begolli, 2007). Response cost is more commonly found as a component in token and

point systems, but has also been used as a stand-alone procedure (Pritchard, 2013). Early

research, such as Woods (1982), implemented response cost in isolation to decrease the

occurrence of severe SIB in an adolescent with ASD. Tickets were provided at the start of each

day and were removed contingent on the target behaviour. The results found support for the

efficacy of response cost. That is, engagement in the target behaviours (i.e., headbanging, biting,

and aggression) were decreased to near zero levels. Similar findings were demonstrated more

recently by Bartlett et al. (2011) in which the removal of a highly preferred item for 10 s

decreased spitting behaviours to near zero levels for a child diagnosed with ASD. Further,

response cost in conjunction with reinforcement-based procedures have been found to be

efficacious in decreasing behaviours, such as inappropriate vocalizations and behaviours

maintained by automatic reinforcement (Falcomata et al., 2004). Beyond efficacy, there exists

some support for the effectiveness of response cost. For example, caregivers were able to

successfully implement response cost in the natural environment (i.e., home). They observed

decreases in target behaviour that was maintained up to four months post-treatment, while low

rates persisted with continued increases in response cost intervals (e.g., from 1 min to 12 min;

Bartlett et al., 2011; Conyers et al., 2004; Woods, 1982). As with overcorrection, it is prudent to

continue to conduct additional research evaluating the effectiveness of response cost in

combination with reinforcement-based interventions.

Efficacy versus Effectiveness of Interventions for Severe Challenging Behaviour

Despite the established prevalence and resultant need for *effective* interventions, adults

with intellectual and developmental disabilities who engage in severe challenging behaviour are

largely underrepresented in the behaviour analytic literature (Cox et al., 2021). Research patterns

also suggest an emphasis on antecedent and reinforcement-based strategies which, for

individuals displaying truly severe challenging behaviour, may prolong treatment (i.e., take longer to achieve treatment gains) or treatment effects may be muted in comparison to what other approaches could achieve (i.e., less overall reduction in challenging behaviour; Gover et al., 2019; Hanley et al., 2005). Extended treatment duration can mean extremely harmful behaviours continue to occur at the peril of the individual with intellectual and developmental disabilities and their caregivers. This consequence may be further amplified by the current emphasis in the literature on efficacy research, which has resulted in a relatively substantial gap in understanding on intervention effectiveness for severe challenging behaviour. It is imperative therefore that intervention effectiveness, in addition to efficacy, be thoroughly researched. Further, comprehensive examination of all behaviour analytic strategies should be explored so that treatment gains can be expedited. This means endeavoring to refine reinforcement-based interventions as well as approaches with decelerative properties (i.e., punishment procedures; Lerman & Vorndran, 2002).

Project Approach

Consecutive Controlled Case Series

There are likely many reasons for the general absence of research on intervention effectiveness featuring adults with intellectual and developmental disabilities who engage in severe challenging behaviour. Some possible reasons could include ease of access to child participants compared to adult participants, continued emphasis on the evaluation of reinforcement-based interventions, or perhaps the ongoing subjective use of the term *severe* (Cox et al., 2021). Another important variable that may interfere with conducting (and successfully disseminating) effectiveness research is the continued perpetuation of the "conjoint-set perspective", in which the seven dimensions described by Baer et al. (1968) are held as

necessary components for *good* research (Critchfield & Reed, 2017, p.123). This, arguably, overly strict criterion could perpetuate a shift in focus on specific methods as opposed to informative lines of investigation even though research of societal importance is predominantly outlined in terms of topics or problems (Critchfield & Reed, 2017). One possible solution may be to employ more flexible research designs that are conducive to applied settings. One such example may be a Consecutive Controlled Case Series (CCCS; Hagopian, 2020). Generally, a CCCS may be described as a research design in which a single-case experimental design (SCED) is "employed with each case in a series of consecutively encountered cases that undergo a common procedure" (Hagopian, 2020, p. 599). This means that, with their permission, the results of all participants who experience a common procedure are featured, regardless of their *success*. The CCCS garners internal validity from applying individualized SCED to each participant, which may demonstrate a reliable change in the target behaviour only after the respective program package was implemented in a systematic manner. Given all participants who took part in the treatment program regardless of their outcome are included, this research design may go beyond determining whether the treatment program produced a positive outcome. Instead, it can facilitate questions pertaining to how often the treatment program can produce a positive outcome (i.e., by identifying functional relations that have generality across cases; Hagopian, 2020). The benefits to this design thus include not only exploring efficacy, but also speaking to program generality while maintaining the methodological rigour of an SCED.

To exemplify, a recent study by Fiani and Jessel (2022) used a CCCS to explore the effectiveness of a behavioural assessment and treatment package process comprised of three phases (i.e., practical functional assessment, behavioural treatment, and extension). Assessment and treatment package delivery occurred across two to three clinical visits spanning a duration of

1 hr each across one week for patients referred to an outpatient clinic. Thirteen individuals with challenging behaviour participated. The results of the study found near elimination of challenging behaviour for all patients to socially meaningful levels. The advantage of conducting a CCCS was that the authors were also able to determine that positive outcomes are a probable occurrence by demonstrating clinically significant reductions in challenging behaviour for 11 of the 13 cases.

CCCS has been applied beyond behavior analytic interventions to exploring drug-behaviour interaction in applied contexts. For example, Cox and Virues-Ortega (2022) examined the rate and function of challenging behaviour within and across naturally occurring medication phases when continuous functional analysis and observation were used. Ten participants were consecutively recruited; four of whom met the inclusion criteria (i.e., clinical diagnosis of intellectual and developmental disabilities, current antipsychotic medication prescription, belief of medication adjustments in the foreseeable future by caregivers and/or the psychiatric team at the time of recruitment, and absence of untreated medical conditions). Their analysis allowed for a clinical demonstration and replication of the featured procedures in an ecologically valid setting. That is, establishing the feasibility of implementing a continuous FA and evaluating medication-behaviour and medication-function interactions in a typical setting (Cox & Virues-Ortega, 2022).

Program Evaluations

Applying CCCS logic in the context of a program evaluation can support an independent review initiative aiming to evaluate treatment package effectiveness. In other words, given program evaluations consider ongoing interventions that continue to be in place in real-life applied settings, it can contribute meaningful information pertaining to program inputs,

outcomes, and feasibility (Miller, 2017). Program evaluations can therefore help to identify areas of strength and areas that need improvement with respect to service delivery, which can result in a myriad of benefits. For example, improving program accountability, informing practitioners and other stakeholders of program effectiveness, and promoting adherence to best-practice treatments (Cox et al., 2022). To exemplify, a recent study by Shepley et al. (2021) applied a retrospective CCCS to conduct a program evaluation of a brief family-centred service provision model at a community-based severe behaviour clinic. All families served through the model from 2017 to 2018 were included, regardless of their success. The results found that for families who received treatment, 92% reported improved behaviour and that brief outpatient services may be effective in assessing and treating challenging behaviour in children for families who adhere closely to the clinic's services.

Research Purpose, Rationale and Hypothesis

The purpose of the proposed project was to make a significant contribution to the field by: (1) adding to a relatively understudied area (i.e., severe challenging behaviour in adults with intellectual and developmental disabilities); (2) speaking to program effectiveness (e.g., under which parameters and/or for whom a comprehensive behavioural intervention package featuring core tenants of applied behaviour analysis will work for); and (3) informing other works interested in program evaluations. This was accomplished by conducting a program evaluation on the use of a comprehensive behavioural treatment package encompassing DR, safe extinction, and strategies featuring decelerative properties (e.g., overcorrection, response cost) to produce marked improvements in persons with intellectual and developmental disabilities who engage in severe challenging behaviour (e.g., frequent, repeated ingestion of poisonous and dangerous items; high risk behaviours resulting in incarceration and/or criminal charges). It was

hypothesized that project outcomes would depict the correct application (i.e., acceptable treatment integrity) of the comprehensive treatment package resulting in primarily successful outcomes (i.e., substantial decrease in severe challenging behaviour and commensurate increase in adaptive skills) for treatment home residents (i.e., applied context).

Method

This program evaluation project received ethics clearance from Brock University's Research Ethics Board (file number 22-333) to: (1) invite all past and current clients who experienced the featured treatment package to participate in the program evaluation; (2) compile existing clinical data for consenting participants for secondary use; (3) conduct treatment integrity; (4) collect interobserver agreement (IOA) data for dependent variables and treatment integrity; and (5) administer social validity surveys to clients, caregivers, and case managers.

Participants and Recruitment

The assistant program director circulated invitations to current participants (or their guardians) who had been admitted to the specialized treatment homes. Separate consent from each participant, or their substitute decision maker, was obtained prior to any direct contact between participants and researchers. Thus, participants featured in this project were eight adults diagnosed with intellectual and developmental disabilities who provided consent and are (were) enrolled in the treatment program at the specialized treatment home, regardless of their success. See Table 1 for a detailed description of participant demographics including sex, age, diagnoses, level of intellectual disability, target behaviours, and hypothesized behaviour function. All participants were referred to their respective treatment home after all community resource options had been exhausted. At the time of admission, all participants were considered in-crisis. For example, prior to admission participants had almost daily interaction with emergency

personnel (e.g., police, emergency response, emergency room), repeated placement breakdowns due to dangerous SIB and/or aggressive behaviour (i.e., had been evicted and were without a place to live), and/or were court ordered to treatment due to altercations resulting in charges being laid or assault related convictions (e.g., assault with a weapon). Challenging behaviour severity is exemplified for each participant below.

Lily

Lily is a 28-year-old female with a diagnosis of ASD, mild intellectual disability, generalized anxiety disorder, oppositional defiant disorder, disruptive behaviour disorder – not otherwise specified (NOS), and factitious disorder. She has a longstanding history of severe challenging behaviour, such that prior to admission to the specialized treatment home, she required emergency medical intervention at least twice per week. That is, she would ingest razor blades, poisonous berries, and other harmful substances that required group home staff to call an ambulance or rush her to emergency where she would need gastric suction (i.e., stomach pumped).

Kevin

Kevin is a 29-year-old male diagnosed with mild developmental disability, reactive mood disorder, paraphilia –NOS, and attention deficit hyperactivity disorder (ADHD). As a result of his engagement in severe challenging behaviour for several years, he had been extensively involved with the judicial system and experienced 10 formal charges in one year. That is, prior to admission to the specialized treatment home, Kevin experienced police intervention at least three times per week, such that staff had to call the police almost every other day for several consecutive months due to the severity of his aggression with staff. He would throw feces at staff, ejaculate and throw semen at them, and would also find items in his home to break and

attempt to use as sharp weapons. He was also so self-injurious that he had caused irreparable damage to his body.

Ginny

Ginny is a 23-year-old female with a global developmental delay. Due to severe property destruction and aggressive behaviour towards family members, including several younger siblings, Ginny transitioned into a group home as a young teen. She stayed at a youth group home until its closure after which she transitioned to adult services. As a result of her excessive aggressive behaviour, she was evicted from her adult placements prior to her admission to the featured specialized treatment home.

Riley

Riley is a 20-year-old male with a diagnosis of ASD and moderate intellectual disability. He has an extensive history of severe challenging behaviour that have limited his ability to access the community and socialize with peers and caregivers. For instance, Riley consistently engaged in severe property destruction (e.g., smashing glass, punching through drywall, breaking laptops), physical aggression (e.g., biting family members such that they had to wear long sleeves to protect themselves), disruptive behaviours (e.g., urinating or defecating outside the toilet such that school staff wore goggles and lab coats to protect themselves from his bodily fluids), and elopement that frequently required police and other emergency services. As a result of his challenging behaviour severity, prior to admission to the specialized treatment home, he resided alone in a locked unit that was video monitored by staff and ply-wall reinforced.

Micaela

Micaela is a 26-year-old female diagnosed with borderline personality disorder (BPD) and developmental delay. She has a history of severe aggressive and destructive behaviours that

have resulted in repeated placement breakdowns, hospitalizations, police involvement, and incarceration. To exemplify, prior to admission to the specialized treatment home, she was charged with assault and released on probation under the condition that she abide by treatment programming wherever she resides.

Anthony

Anthony is a 21-year-old male diagnosed with alcohol related neurodevelopmental disorder, ADHD, sensory processing disorder, anxiety disorder – NOS, depressive disorder – NOS, developmental coordination disorder, and a learning disability in reading, writing, and math. He has a history of severe aggressive and destructive behaviours that have resulted in repeated placement breakdowns, hospitalizations, and police involvement. For instance, previous agencies indicated that they could no longer support him due to several incidents involving physical and environmental aggression; one of which required police intervention that resulted in Anthony being charged with two counts of assault and assault of a peace officer. Anthony was suspended from school due to an incident wherein he was arrested and charged with mischief. He was never able to return to school, despite programming aimed at helping him meet the school's prerequisites for him to return to school.

Taylor

Taylor is a 29-year-old female diagnosed with ASD, BPD, schizophrenia, and developmental delay. She has a history of severe aggressive and destructive behaviours that have resulted in placement breakdowns, hospitalizations, and police involvement. To exemplify, at her previous placement, she engaged in severe physical aggression where she injured staff and attempted to injure herself by running onto the streets. As a result of this incident, she was admitted to hospital, charged with three counts of assault, and placed on 12 months of probation

which subsequently led to a short stay at a correction centre as well as at brief shelters and respite placements, which she was ultimately evicted from.

Oliver

Oliver is a 24-year-old male with a diagnosis of unspecified impulse control disorder (sexual urges), major depressive disorder, anxiety disorder –NOS, conduct disorder, ADHD, fetal alcohol spectrum disorder, borderline intellectual functioning as well as global and severe adaptive limitations. He has a history of engaging in severe challenging behaviour, such as property destruction, elopement, and inappropriate sexual behaviours. To exemplify, at his previous placement, police were contacted following an incident where he threatened staff with a hammer and proceeded to engage in environmental destruction. Following this incident, he was admitted to a hospital where he was assessed and found to be a risk to himself and others by the physician, which ultimately resulted in his eviction from his previous placement.

Setting

The featured treatment package was implemented in three specialized treatment homes. More specifically, the treatment homes were designed to treat challenging behaviour, prevent hospital admission, and serve as a transition placement where participants could develop appropriate replacement behaviours and adaptive skills to facilitate successful relocation to community placements. The treatment team at each treatment home was comprised of direct-care staff, clinical staff (i.e., Board-Certified Behaviour Analysts (BCBA), behaviour therapists, and behaviour therapist assistants), a supervising Psychologist who is also a BCBA at the Doctoral level, and a clinical house manager. Each treatment home housed two to three participants who receive continual care 24 hours a day for seven days a week. All three treatment homes consisted

of an open concept layout with individualized programming areas, a kitchen, and a bathroom comprising the shared space as well as separate bedroom units for each participant.

Measurement

Response Measures

Participants engaged in a range of challenging behaviour topographies (i.e., negative target behaviours), and each was operationally defined as per their existing behaviour support plan (BSP). Some examples included verbal aggression, physical aggression, ingesting foreign objects, and SIB. Table 2 provides a sample of negative target behaviours and operational definitions across participants, while Appendix A features a comprehensive list and corresponding operational definitions. Staff collected data on the daily frequency of episodes. For ease of the current program evaluation, daily episodes of challenging behaviour served as the primary dependent variable. That is, an episode was scored if the participant engaged in any instance of negative target behaviour listed on their BSP. This information was collected daily by the frontline team members. Raw data was entered into an excel spreadsheet by behaviour therapist assistants. The data were used to inform clinical decisions as per the supervising Psychologist and BCBA overseeing the case.

The second dependent variable, adaptive behaviour, was also individualized across participants. For each participant, a skill acquisition program (SAP) outlining target adaptive behaviours was developed and updated on a weekly basis by the clinical team, as needed. These behaviours ranged from following two-step instructions and receptive identification to independent public transit use and obtaining and maintaining volunteer and/or paid employment. Table 3 provides a sample of adaptive target behaviours and response requirements, while Appendix B features a comprehensive list. The measurement system and schedule used to collect

data varied for each participant (e.g., a cold probe once a week, a cold probe twice a week, discrete trial training (DTT) four times a week). This information was collected by frontline team members and clinical staff at the treatment homes. For all participants, adaptive behaviour skills progress was measured by percentage program mastery across months in a skills supervision report, which was used to inform clinical decisions.

Interobserver Agreement

IOA data was collected for both dependent variables. That is, for challenging behaviour and adaptive behaviour. Specifically, for challenging behaviour, trained, independent observers: (1) recorded the start and end time of each observation period to facilitate accurate IOA calculations; and (2) stood off to the side of the room and recorded the frequency of challenging behaviour each time it occurred during IOA observation sessions. Total count IOA was used in which reliability between staff recorded data and independent observer data was calculated by dividing the smaller count by the larger count and multiplying that by 100 (Cooper et al., 2020). For adaptive behaviour, trained, independent observers: (1) were onsite when cold probes or DTT were conducted by frontline team or clinical staff; and (2) stood off to the side of the room and recorded data as per the measurement system outlined for the skill being observed. For DTT, trial-by-trial IOA, a conservative method described as a meaningful index for discrete trial data, was used (Cooper et al., 2020). That is, reliability between staff recorded data and the independent observer data was calculated by dividing the number of trials that were in agreement between the two observers by the total number of trials and multiplying that by 100 (Cooper et al., 2020; Ledford & Gast, 2018). Agreement was described as both staff and the independent observer denoting the occurrence of a correct response for a given trial as defined by the program (e.g., tacts correct picture independently without any prompting). For cold probes, researchers

applied occurrence IOA. Specifically, each observation opportunity constituted an occurrence trial. Reliability between original data and observer data was calculated by dividing the number of occurrence trials (i.e., observation opportunities) that were in agreement by the total number of occurrence trials and dividing that by 100 (Ledford & Gast, 2018). Agreement was described as both staff and the independent observer denoting the occurrence of a correct response for the probe as defined by the program (e.g., states correct answer to question about collaborative problem solving).

Participants reside in the residence 24 hours a day for 365 days a year. Programming is applied from when the participants wake up to when they go to sleep across the homes. Given independent observers (myself, and two trained research assistants) could not be onsite for the entire active programming duration which constituted up to 16 hours each day, the team endeavored to collect data for both dependent variables one to two times per week for each participant for up to 20% of sessions (i.e., days). To accomplish this, each day of the week was considered one data collection opportunity and, as such, there were seven data collection opportunities per week per participant. This means that conducting data collection for both dependent variables once per week per participant was equivalent to 14% of sessions (i.e., days), while collecting data for both dependent variables twice per week per participant was equivalent to 29% of sessions (i.e., days). Although, it is recommended that IOA be collected for at least 33% of sessions, due to the extensive nature of the sessions, doing so would require a larger team, and additional individuals onsite more often. More individuals onsite to collect data could have become disruptive to the staff and clients, and possibly affect client progress. Further, adequate space could have become an issue with the addition of more than one independent observer (i.e., additional people occupying a small space). Finally, attempting to capture IOA via

video recording was not feasible given the camera angles made it difficult to discern client or staff behaviour depicted in some of the closed-caption footage being recorded. Further, the closed-caption footage did not have audio, which meant observers would not have been able to accurately capture certain negative target behaviours (e.g., verbal aggression, verbal disruption) or the occurrence of correct responding for adaptive behaviour (e.g., without verbal prompts).

IOA data for challenging behaviour were collected across 8% of sessions for Lily, 9% of sessions for Kevin, 11% of sessions for Ginny, 9% of sessions for Riley, 15% of sessions for Micaela, 8% of sessions for Anthony, 5% of sessions for Taylor, and 7% of sessions for Oliver. Typically, each observation session was approximately 1–1.5 hr in duration and observation sessions were conducted 1–2 days per week for each participant. IOA outcomes for challenging behaviour were as follows: 95% for Lily (range 17%–100%), 100% for Kevin (range 100%), 86% for Ginny (range 33%–100%), 95% for Riley (range 75%–100%), 100% for Micaela (range 100%), 100% for Anthony (range 100%), 100% for Taylor (range 100%), and 100% for Oliver (range 100%).

IOA data for adaptive behaviour were collected across 7% of sessions for Lily, 8% of sessions for Kevin, 6% of sessions for Ginny, 9% of sessions for Riley, 5% of sessions for Anthony, and 6% of sessions for Taylor. Typically, each observation session varied between 10–30 min depending on the skill interval observed and observation sessions were conducted 1–2 days per week for each participant. IOA outcomes for adaptive behaviour were as follows: 94% for Lily (range 63%–100%), 98% for Kevin (range 86%–100%), 100% for Ginny (range 100%), 92% for Riley (range 0%–100%), 97% for Anthony (range 75%–100%), and 95% for Taylor (range 50%–100%). IOA data for adaptive behaviour was not collected for Oliver as he did not have a formal SAP at the time of data collection. IOA data for Micaela was halted at

approximately one month after initiation of data collection due to extenuating circumstances. For these reasons, outcomes for Micaela and Oliver are listed as not applicable.

Comprehensive Treatment Package Procedures

Below is a general summary of the featured comprehensive treatment package that is being implemented for each participant by the staff at the treatment homes. Importantly, the programs are individualized to match each participants' unique profile (i.e., adaptive functioning, preferred items/activities). However, the general process is consistent across participants (as described below).

Pre-admission

The intake process was conducted independent of the program evaluation and could not be a part of the current project. However, in the interest of transparency (i.e., technological; Baer et al., 1968), I will briefly summarize this process, as well as describe corresponding elements enacted. Intake generally involved the agency behaviour therapists completing a comprehensive behaviour assessment report for each participant. This process typically includes interviewing previous support staff and caregivers (Daily Activity Schedule-Questionnaire [DAS-Q]; Linder, 2014), completing a comprehensive file review to obtain relevant past assessment and supports provided, obtaining any existing target behaviour data, and conducting functional behaviour assessment on negative target behaviour to hypothesize behaviour function (e.g., Questions About Behavioural Function [QABF]; Matson & Vollmer, 1995; Functional Assessment Screening Tool [FAST]; Iwata & DeLeon, 2005). Information garnered from this process informed each individualized initial BSP.

Pre-admission Assessments Sample.

DAS-Q (Linder, 2014). The DAS-Q (Linder, 2014) is a questionnaire administered to caregivers to assess the occurrence of participants' activities and participation across days. Caregivers score items pertaining to 10 domains (i.e., completeness, specification, timing, outings, teaching, choice, exercise, variety, chores, and cognitive) on a five-point Likert scale ranging from 0 to 4 (i.e., never to always).

QABF (Matson & Vollmer, 1995). The QABF (Matson & Vollmer, 1995) is a functional assessment tool administered to caregivers to report on categories pertaining to behaviour function. These five categories include: attention, escape, physical, tangible, and non-social; each category consists of five items. Caregivers score items on a four-point Likert scale ranging from 0 to 3 (i.e., never to often).

FAST (Iwata & DeLeon, 2005). The FAST (Iwata & DeLeon, 2005) is a functional assessment tool administered to several individuals who frequently interact with the participant to collect information on behaviour function. The tool speaks to five different functions: attention, activity, escape, sensory stimulation, and pain attenuation. Caregivers score 27 items with forced choice (i.e., yes/no).

Direct Observation, Caregiver and Client Interviews & Comprehensive File Review. Agency BCBAs conduct interviews with the clients (where applicable), caregivers, and other stakeholders, as well as repeat direct observation sessions prior to intake. They also conduct a comprehensive file review.

Intervention

The intervention consisted of three overarching phases: (1) Behavioural Stabilization; (2) Skill Acquisition; and (3) and Generalization and Maintenance. However, participants

experienced individualized behavioural programming changes to facilitate progress through the three overarching phases.

      Phase 1. The first phase, Behavioural Stabilization, is the main focus in this initial step given skills can be difficult to teach when frequent and/or severe negative target behaviours interfere. This is achieved through a combination treatment package that includes differential reinforcement, *safe extinction*, token economy, response cost, overcorrection and response blocking (where applicable). As part of programming, each participant is expected to complete their daily activity schedule (DAS). The DAS is an individualized, highly structured routine consisting of several activity intervals per day. These activities include task sequences, activities of daily living (ADLs), household chores, community outings (where appropriate), physical activity, leisure time, and staff or peer interaction times. Token economy and decelerative strategies (e.g., response cost, overcorrection) are incorporated to discourage negative target behaviours. Points, awarded contingent on the absence of negative target behaviour and ongoing active participation in their DAS, are redeemed for predetermined daily or weekly awards (token economy). If participants do not cooperate with their DAS or reasonable staff instructions (e.g., "It's not time for leisure right now, let's get back to the task at hand"), a least-to-most prompt hierarchy is implemented. For instance, for Oliver, upon task refusal, staff implement active listening (i.e., three prompts spaced 2 min apart followed by a total duration of 5 min to allow the participant to consider their options), then collaborative problem solving is enacted. If these strategies do not facilitate task completion, then staff are expected to provide a gestural prompt followed by full physical prompting. If participants engaged in physical aggression, SIB, environmental destruction, and/or other severe forms of negative target behaviours outlined in their BSP, physical and/or mechanical restraints are implemented to ensure the severe negative

target behaviour does not contact reinforcement (i.e., activity completion is avoided). That is, *safe extinction* is enacted. Restraint durations are minimized, such that specific release criteria are individualized for each participant and strictly adhered to. For example, for Taylor, one limb strap is faded for every 10 correct and independent responses of the task at hand in a prespecified order as outlined in her BSP (i.e., second hand, shoulders, leg, waist). Importantly, one hand is always unrestricted so that the participant can complete tasks.

Phase 2. The second phase, Skill Acquisition, is the focus once sufficient learning has occurred in that negative target behaviours do not result in access to reinforcement. Individualized skill building programs for target adaptive skills are developed, augmented by least-to-most prompting procedures according to the participants' strengths. For example, for Ginny, toileting was identified as a target adaptive skill and an individualized program outlining training steps, as well as a reinforcement hierarchy (i.e., behaviour specific verbal praise, photos on iPad for 2 min, photos on iPad for 4 min, and iPad access for 5 min), was established. The DAS also becomes less structured. For example, Kevin is currently completing an independent DAS, which means that he is required to move through his day on time with few verbal prompts. Importantly, less structure does not mean reduced expectations. That is, participants are expected to maintain high levels of engagement in meaningful daily activities (e.g., studying for the General Educational Development certification). Further, physical and mechanical intervention supports continue to be in place should the participant engage in severe negative target behaviours.

Phase 3. In the third phase, Generalization and Maintenance, the skills that were taught in phase 2 continue to be generalized across a wide variety of individuals, times, and settings. Generalization and Maintenance are the focus, while the phase also involves preparation for

discharge from the treatment setting. For example, during this phase, participants may decide to take part in volunteer or paid work activities outside of the treatment home thereby broadening their opportunities for generalized responding.

Treatment Integrity

A behaviour therapist or BCBA created treatment integrity checklists (TICs) in accordance with each participant's clinical programming (e.g., BSP, Task Guidelines, Reinforcer Guidelines). Relevant modifications to TICs are made on an ongoing basis in accordance with programming changes (e.g., BSP changes). Treatment integrity data are typically collected by either a BCBA, a behaviour therapist assistant, or the house supervisor who observe staff onsite and score each item on the TIC as a correct response or an incorrect response (see Appendix C for a sample TIC). Of note, for the project's purpose, the frequency that treatment integrity is collected by the clinical team does not align with how often research standards dictate these should be completed (see Cooper et al., 2020). Therefore, additional TICs were collected by the research team to better align with these standards. So, clinical TIC data were gathered with the research team augmenting integrity data. That is, the clinical staff enacted treatment integrity as normal, and the research team collected treatment integrity data between one to two times per week for each participant for up to 20% of sessions (i.e., days). To achieve this, like IOA for the dependent variables, each day of the week counts as one TIC opportunity. So, there are seven TIC opportunities per week per participant. This means that conducting one TIC per week per participant was equivalent to 14% of sessions (i.e., days), while conducting two TICs per week per participant was equivalent to 29% of sessions (i.e., days). Treatment integrity was calculated by dividing the total number of correct responses by the total number of items on the checklist and multiplying that by 100 (Cooper et al., 2020).

Given the research team were not onsite for the entire active programming duration, there was the possibility that an opportunity to observe all items on the TIC did not present itself. Therefore, each item on the TIC was scored as a correct or incorrect response through observation *or* through the verbal administration of standardized questions to frontline staff. Standardized questions were created across all TICs for use during verbal administration to generate consistency across all researchers conducting TICs (see Appendix D for a sample TIC with standardized questions and notes). In the instance that an item on the TIC was both observed and verbally administered, the score informed by the observation was recorded. Further, if more than one opportunity to observe an item on the TIC was presented, and one observation yielded an incorrect response by the staff while another depicted a correct response – the incorrect response was scored.

Treatment integrity data were collected across 14% of sessions for Lily, 22% of sessions for Kevin, 23% of sessions for Ginny, 10% of sessions for Riley, 17% of sessions for Micaela, 7% of sessions for Anthony, 10% of sessions for Taylor, and 9% of sessions for Oliver. Typically, each observation session was approximately 1–1.5 hr in duration and observation sessions were conducted 1–2 days per week for each participant. Treatment integrity outcomes were 93% for Lily (range 71%–100%), 86% for Kevin (range 52%–100%), 90% for Ginny (range 60%–100%), 82% for Riley (range 22%–100%), 82% for Micaela (range 50%–100%), 82% for Anthony (range 37%–99%), 88% for Taylor (range 56%–100%), and 80% for Oliver (range 63%–98%). The standards for staff set out by the agency providing services to these clients state that overall treatment integrity scores should yield an average weekly minimum score of 80% before more intensive training and problem solving is to occur by the clinical team and management. Treatment integrity outcomes reflect that these standards are being met.

To lend further confidence to the reliability of the data collected, treatment integrity IOA was conducted once a month for randomly selected participants. That is, two researchers (either two research assistants, or the research lead (BCBA at the Doctoral level) and a research assistant) independently scored each item on the TIC as a correct response or an incorrect response at the same time. Reliability between both researchers was calculated by dividing the number of items on the TIC that were in agreement by the total number of items and multiplying that by 100. Agreement was defined as: (1) both researchers scoring an item on the TIC as a correct response; or (2) both researchers scoring an item on the TIC as an incorrect response.

Of the sessions for which treatment integrity were conducted, treatment integrity IOA data was collected across 7% of sessions for Lily, 5% of sessions for Kevin, 8% of sessions for Ginny, 14% of sessions for Riley, 20% of sessions for Micaela, 11% of sessions for Anthony, 8% of sessions for Taylor, and 16% of sessions for Oliver. Treatment integrity IOA outcomes were 99% for Lily (range 97%–100%), 100% (range 100%) for Kevin, 100% for Ginny (range 100%), 96% for Riley (range 92–-100%), 95% for Micaela (range 90%–100%), 93% for Anthony (range 85%–100%), 96% for Taylor (range 88%–100%), and 92% for Oliver (range 83%–100%).

Error Analysis. The global treatment integrity analysis and outcomes listed above provide an overall quantification of integrity. However, as with any approach that includes aggregating data, it is possible that these overall scores might have masked poorer performance on individual elements of the treatment package (Cook et al., 2015). Poor performance on individual components may differentially affect treatment outcomes (e.g., Carroll et al., 2013). Therefore, conducting an item-by-item error analysis could facilitate an informed understanding

of treatment integrity and may subsequently allow for the clinical team to modify training sessions or implement targeted booster sessions as appropriate.

The number of items and domains on the TIC were as follows for each participant. Lily's TIC contained 52 items across the five domains of preventative strategies, management of behaviour, overnight protocol, bathroom protocol, and data collection. Kevin's TIC consisted of 50 items across the four domains of preventative strategies, reinforcer guidelines and rewards, management of behaviour, and data collection. Ginny's TIC consisted of 65 items across the 12 domains of preventative strategies, teaching task sequence, educative routines, verbal behaviour, management of behaviour, management of behaviour during walks, bathroom protocol, bathroom requests, meal protocol, urine accident after mechanical chair release, back-up staff, and data collection. Riley's TIC totalled 92 items across the seven domains of preventative strategies, daily activity schedule, management of behaviour, bathroom, evening/bedtime, back-up staff, and data collection. Micaela's TIC contained 58 items across the four domains of preventative strategies, reinforcer guidelines and rewards, management of behaviour, and data collection. Anthony's TIC consisted of 71 items across the five domains of preventative strategies, reinforcer guidelines and rewards, management of behaviour, bathroom, and data collection. Taylor's TIC totalled 59 items across the four domains of preventative strategies, reinforcer guidelines and rewards, management of behaviour, and data collection. Finally, Oliver's TIC contained 40 items across the six domains of target behaviours and operational definitions, preventative strategies, reinforcer guidelines and rewards, access protocols, management of behaviour, and data collection.

Design

Data collected for this program evaluation were analyzed within a nonexperimental consecutive case series, which is described as a research design that "involve the application of a defined treatment to a series of cases presenting with similar problems" (Hagopian, 2020, p. 614). The use of a SCED with each case is not *required* within a nonexperimental consecutive case series and, as such, A-B designs are often used (e.g., Lomas Mevers et al., 2018). This means that, with their consent, the results of all participants who experienced the featured treatment package to decrease challenging behaviour and increase adaptive skills at one of the three specialized treatment homes between the dates of October 2017 to July 2023 were reported. The current project may be better described as a *hybrid* nonexperimental consecutive case series given that the program evaluation was initiated in May 2022 and data collection for the project (i.e., IOA, treatment integrity) continued until July 2023. This means that while a significant percentage (i.e., greater than 50%) of the data were compiled during service delivery for each participant, a percentage of the data was also collected following the start of the current project. To exemplify, Lily has been in treatment for approximately 6 years; one of which overlapped with data collection for the program evaluation. This means that approximately 83% of Lily's data were obtained in the course of service delivery prior to the initiation of data collection for this project. For other, more recently, admitted participants this ratio is different. For instance, Oliver has been in treatment for approximately 2 years in which the latter year overlapped with data collection for this project. This means that while half (i.e., 50%) of Oliver's data was obtained in the course of service delivery, the other half was collected after the current program evaluation was initiated.

Data Analysis

Due to the severity of the challenging behaviour engaged in by the participants included in the current project, phase 1 (i.e., Behavioural Stabilization) was implemented immediately upon admission. This approach is not uncommon in the treatment of severe challenging behaviour (Pritchard et al., 2018), even though it can complicate the prediction and verification of resultant intervention outcomes at times. To address this limitation in the current project, baseline and intervention were delineated as early treatment and late treatment, respectively. Such delineations are often observed in cardiology research, in which the first 30 days constitute early treatment and the days spent in treatment after the first 30 days constitute late treatment (O'Hara et al., 1995; Homorodean et el., 2019). For instance, O'Hara et al. (1995) examined mortality and morbidity rates associated with abdominal aortic aneurysm, in which the first 30 days constituted early survival and **all** days after the first 30 days constituted late survival.

In order to generate individualized early treatment phases for each participant featured in the current program evaluation, quantitative criteria described by Schoenfeld et al. (1956) utilized in a program titled *Stability Check* were employed (Costa & Cançado, 2012). The Schoenfeld et al. (1956) criterion is represented by the following formula.

$$\left| \frac{\left[\left(\frac{D_1 + D_2 + D_3}{3} \right) \right] - \left[\left(\frac{D_4 + D_5 + D_6}{3} \right) \right]}{\left(\frac{D_1 + D_2 + D_3 + D_4 + D_5 + D_6}{6} \right)} \right| \times 100\%$$

According to this criterion, the first 7 days on any schedule are not considered in calculating stability. For the next 6 days, the mean of the first 3 days (i.e., D_1, D_2, and D_3, respectively) of the six is compared with that of the last 3 days (i.e., D_4, D_5, and D_6, respectively). If the difference between these means is less than 5% of the six days' mean, the individual is considered to have stabilized and would have been shifted to the next condition. If

the difference between sub means is greater than 5% of the grand mean, another experimental

day is added (i.e., D_7) and similar calculations are made for that day as well as the 5 days

immediately preceding it (D_2, D_3, D_4, D_5, and D_6). The following formula is to be used:

$$\left| \frac{\left[\left(\frac{D_2 + D_3 + D_4}{3}\right)\right] - \left[\left(\frac{D_5 + D_6 + D_7}{3}\right)\right]}{\left(\frac{D_2 + D_3 + D_4 + D_5 + D_6 + D_7}{6}\right)} \right| \times 100\%$$

Such extensions of the experiment and calculations of stability are continued daily until

the prespecified 5% criterion is reached (Schoenfeld et al., 1956).

The Schoenfeld et al. (1956) criterion was employed with the data compiled for each

participant to delineate early treatment (i.e., baseline). Stability was established at 88 days for

Lily, 14 days for Kevin, 26 days for Ginny, 34 days for Riley, 15 days for Micaela, 17 days for

Anthony, and 24 days for Taylor. However, for Oliver, the Schoenfeld et al. (1956) criterion was

unsuccessful in determining stability. Prior research has addressed similar occurrences by

averaging across participants. For instance, Anderson et al. (2023) found missing effect sizes

during data set preparation for regression analysis. After determining that this was due to the

original article not collecting enough data on the same number of behaviours across pairings of

interventionists and learners, they averaged the effect sizes across the interventionist/learner

skills to create a single interventionist/learner effect size that align with the corresponding

interventionist/learner effect size (Anderson et al., 2023). Thus, to preserve Oliver – rather than

excluding this case due to 'missing data', a proxy baseline was delineated at 27 days for Oliver

by averaging the days at which stability was established for the other seven participants.

Modified Brinley Plot

Given featured participants experienced the comprehensive treatment package every day

for up to 6 years, the resultant dataset will consist of an extensive number of sessions (i.e., 591

sessions to 2085 sessions). This would make interpreting the resultant time-series graphs difficult. A modified Brinley plot boasts several advantages over visual inspection of times-series graphs, which are exemplified by the current project (Manolov et al., 2022) First, employing this approach may afford greater efficiency. That is, this plot allows for the representation of results for several comparisons within and across participants and the addition of visual aids (e.g., identity line, grey dashed diagonal line) allow for a quick evaluation of multiple aspects (Manolov et el., 2022). Second, past research has found insufficient agreement between visual analysts inspecting time-series plots (Ninci et al., 2015). In fact, when using time-series graphs, variations in the ratio between the x-axis and y-axis have been found to lead to potential distortion during visual analysis (Kubina et al., 2017). Alternatively, since a modified Brinley plot is square by definition, it is unaffected by graphical features, such as the ratio between the x-axis and y-axis, thereby reducing visual inspector bias during visual analysis (Manolov et al., 2022).

Notably, times-series graphs depicting all baseline and intervention challenging behaviour data across all consecutive applications ($n = 8$) of the comprehensive treatment package may be found in Appendix E presented as a nonconcurrent multiple baseline graph and in Appendix F displayed as individual graphs. However, visual inspection was not enacted and thus did not inform Results interpretation.

Social Validity

Social validity questionnaires were administered to clients, caregivers, and case managers to assess treatment goals, procedures, and effects by consumers (Ferguson et al., 2019). More specifically, two social validity assessment processes were enacted: one for clients and the other for caregivers and case managers.

For clients, neutral personnel (i.e., behaviour therapists who were not affiliated with the program evaluation) informed clients of the opportunity to provide feedback on the services being provided. Clients were notified verbally and in writing of the option to request support from trusted staff to complete the social validity measure. The survey was administered electronically via Qualtrics and was formatted to reduce response effort for clients. This was accomplished in part by: (1) ensuring the survey font was large; (2) bolding important keywords to draw attention to them; (3) including a progress bar at the top of the survey page; and (4) not requiring the completion of all questions in one sitting. The first page of the survey provided clients with information regarding the importance of consumer feedback as well as instructions on how to complete the survey. If clients agreed to participate, they clicked "Yes, I agree" to proceed. However, if they did not agree to participate, they clicked "No, I do not agree" or clicked out of the survey. Given clients are diagnosed with some level of intellectual and developmental disability, if they agreed to participate, a pre-test consisting of five orienting questions was provided at the onset of the survey to ensure the validity of their answers (e.g., What month is it?; Wells & Ruesch, 1972). A percentage correct for the pre-test was generated following completion and a score of 80% (i.e., 4/5) or greater on the pre-test suggested that client responses may be considered valid (Wells & Ruesch, 1972). However, all clients, regardless of their score on the pre-test proceeded to the social validity questions. The social validity survey was separated into three domains: (1) Before Services ($n = 6$ questions); (2) During Services ($n = 5$ questions); and (3) Final Thoughts ($n = 5$ questions; Lambert et al., 2022; Luiselli et al., 2015; Luiselli et al., 2017). Questions were either open-ended, yes/no, or on a five-point Likert scale. The five-point Likert scale constituted the following: Strongly Agree ("I really agree with this"), Agree ("I sort of agree with this"), Disagree ("I sort of disagree with this"), Strongly Disagree ("I

really disagree with this"), and Prefer Not to Answer (Sturgis et al., 2012). See Table 4 for a fulsome description of social validity questionnaire statements, domains, and formats for clients.

For caregivers and case managers, neutral personnel (i.e., behaviour therapists who were not affiliated with the program evaluation) circulated emails informing them of the opportunity to provide feedback on the services received by their clients or loved ones at the specialized treatment home. A link to the Qualtrics social validity survey was included in the email. Response effort to complete the measure was reduced for caregivers and case managers by: (1) including a progress bar at the top of the survey page; and (2) not requiring the completion of all questions in one sitting. Before beginning the survey, like clients, caregivers and case managers were provided with information regarding the importance of consumer feedback as well as instructions on how to complete the survey. Agreement to participant was indicated by clicking "Yes, I agree" to proceed and disagreement to participate was indicated by not clicking the link in the email, clicking "No, I do not agree," or clicking out of the survey. Caregivers and case managers were not presented with the pre-test and instead proceeded directly to the social validity survey questions. The layout and questions of the social validity survey mirrored that of the clients, however, the terminology used within the questions was slightly more advanced (e.g., "The current treatment makes it possible for my client (my loved one) to progress and achieve a better quality of life." instead of "This program has helped me progress and work towards a better life. I feel better now."). See Table 5 for a fulsome description of social validity questionnaire statements, domains, and formats for clients.

Results

Challenging Behaviour Outcomes

A modified Brinley plot provides a graphical display of phase means so that any treatment effect may be clearly observed. That is, a comparison between baseline (phase A) and intervention (phase B) is defined by the phase A mean and the phase B mean (Blampied, 2017; Manolov et al., 2022). A diagonal line, known as an identity line, in which the intercept is zero and the slope is one, is included to represent the equivalence between baseline and intervention (i.e., no difference; Manolov et al., 2022). If the improvement is an increase, then all data points should be above the diagonal line and if the improvement is a reduction, then all data points should be below the diagonal line (Manolov et al., 2022). For the current project, given this is challenging behaviour data, improved outcomes would be depicted by the data points falling below the identity line (Blampied, 2017; Manolov et al., 2022). Data points for Lily, Kevin, Riley, Micaela, Anthony, and Oliver are found below the identity line while the data point for Ginny is depicted above the identity line and the data point for Taylor is on the identity line. Thus, therapeutic outcomes (i.e., a decrease in challenging behaviour) are observed for 75% of participants ($n = 6$), neutral outcomes (i.e., no change) are observed for 13% of participants ($n = 1$), and contra-therapeutic outcomes (i.e., an increase in challenging behaviour) are observed for 13% of participants ($n = 1$).

An additional visual aid in the form of a grey dashed diagonal line can be included on modified Brinley plots. This line represents the desired *magnitude* of change from the baseline level (Manolov et al., 2022). Often, this line is defined using expert judgement as it pertains to applied significance. That is, domain-specific knowledge can be drawn upon to determine when a sufficient departure from challenging behaviour has occurred (Manolov et al., 2022). Existing

literature suggests that an 80% reduction in challenging behaviour is indicative of an effective intervention (Hagopian et al., 1998; Rooker et al., 2013). For the current project, an 80% reduction in challenging behaviour would be depicted by the data points falling on or below the grey dashed diagonal line. Data points for Micaela and Anthony are found below the specified line while the data point for Oliver is depicted on the line and the data points for Lily, Kevin, Ginny, Riley, and Taylor are above the line. Thus, the desired magnitude of change from the baseline level (i.e., 80% reduction) has been achieved for 38% of participants ($n = 3$) and has not *yet* been reached for 63% of participants ($n = 5$).

Adaptive Behaviour Outcomes

Baseline data for adaptive behaviour are not reported. This is because skill acquisition data was not being collected upon intake across participants (see Discussion for further commentary). That is, the earliest recorded skills data was at 16 months after admission for Lily, 15 months after admission for Kevin, 14 months after admission for Ginny, 6 months after admission for Riley, 3 months after admission for Micaela, 4 months after admission for Anthony, and 3 months after admission for Taylor. As noted earlier, Oliver did not have a formal SAP at the time of data compilation for the current program evaluation.

Intervention data for adaptive behaviour (i.e., cumulative number of mastered skills targets per month) is depicted for all consecutive applications ($n = 7$) of the comprehensive treatment package in Figures 2–8. For 100% of the participants ($n = 7$), an accelerating, therapeutic trend was observed across most of the intervention phase. That is, a steep slope indicating an impressive increase in the number of cumulative mastered skills targets was observed for all participants; albeit learning (as evidenced by skills mastered) began at different times. For example, for Lily there were approximately three months wherein no skill mastery

was observed. This was followed by some new skills learned in month 19. Ten months passed wherein skill mastery appeared stagnant. However, at month 30 accelerated skill mastery was observed and continued on a steep upward trend to present (Figure 2). For Kevin, month 30 also appeared to be where skill mastery began to flourish. A modest upward trend appeared at first (month 31–39) followed by a much steeper, sustained upward trend (month 40 to present; Figure 3). For Ginny, month 16 appeared to be the start of a modest upward trend, with a sharp increase in skill mastery from month 20–25 followed by steady increases to present (Figure 4). For Riley, the number of skills mastered appeared stagnant for 2 months before a steady accelerating trend was observed from month eight to present (Figure 5). For Micaela, a zero celerating trend was observed for approximately 5 months wherein no skills were mastered. Following month seven, a modest upward trend was first observed (month eight to 29) after which there appeared to be a steep increase from month 29–35 (Figure 6). For Anthony, a small accelerating trend was observed for the first 7 months indicating that skills mastery continued to consistently occur albeit at low levels. From month 10 to month 11, there was a steep accelerating trend with the number of skills mastered increasing from 26 to 188. A gradual upward trend was then observed from month 12 to present, suggesting that Anthony continued to reliably master skills targets (Figure 7). For Taylor, a stable increasing trend was observed starting in the fourth month and continued to present (Figure 8).

Treatment Integrity Error Analysis Outcomes

The results of the treatment integrity item-by-item error analysis are depicted for all participants ($n = 8$) in Figures 9-16. For Lily, percentage of error are generally low across all items ranging from 0%–33% with more error observed on items 24 (27%), 31 (27%), 32 (33%) from the management of behaviour domain, and item 43 from the overnight protocol domain

(30%). For Kevin, low to moderate error percentages were observed across all items (range 0%–58%). Items wherein more errors were made stemmed from the management of behaviour domain, namely items 26 (55%), 27, (45%), and 28 (58%). For Ginny errors were relatively low across items, apart from item 53 (49%), which resulted in a range of 0% to 49%. This item was from the urine accident after mechanical chair release domain. For Riley, percent error on TIC items appeared highly variable (0%–81%). Notably, most errors appeared to coincide with items 2 (58%), 5 (81%), and 10 (69%), which corresponded with the preventative strategies domain. Modest percentage errors were observed on items 9 (35%; preventative strategies), 26 (35%; daily activity schedule), 27 (39%; daily activity schedule), 41 (42%; daily activity schedule), 47 (35%; management of behaviour), 48 (35%; management of behaviour), 72 (35%; bathroom), and 79 (35%; evening/bedtime). For Micaela, error levels were quite variable ranging from 0% to 86%. More errors appeared to coincide with items 2 (71%) in the preventative strategies domain, as well as items 54 (71%), and 55 (71%). Modest error percentages coincided with items 29 (42%), 31 (42%), 32 (57%), 36 (42%), 43 (57%), 44 (71%), 45 (57%), 48 (42%), 50 (85%), 51 (57%), and 53 (43%). For Anthony, like Riley and Micaela, percent error on items varied from low to high (range, 0%–84%). Most errors appeared to coincide with the preventative strategies domain, namely items 4 (68%), 7 (84%), and 9 (72%). Errors also appeared most often in the management of behaviour domain, specifically items 60 (60%), 62 (56%), and 63 (64%). For Taylor, outcomes also depict quite variable error patterns (range 0%-88%). Higher percentage error appeared to correspond with items 2 (69%), 6 (88%), and 8 (69%) in the preventative strategies domain; as well as items 38 (63%), 39 (50%), and 40 (56%) in the management of behaviour domain. Finally, for Oliver error patterns were also variable, ranging from 0%–88%. The highest error percentages appeared to coincide with the management of

behaviour domain. Specifically, items 27 (59%), 29 (53%), 30 (71%), and 33 (88%). Item 14, which was in the preventative strategies domain, also coincided with more error (70%).

Social Validity Outcomes

Seventy-five percent of participants elected to complete the social validity survey ($n = 6$), while seven caregivers/case managers elected to complete the survey, which were distributed at the cessation of data collection for the current program evaluation.

Clients

Social validity results for clients are presented in Table 6. The pre-test score of 80% (i.e., correct responding on 4/5 questions) or greater outlined a priori was achieved by 75% of clients ($n = 5$) who completed the survey. Their answers were thus considered valid for inclusion. Among this sample, 80% (4/5) of participants indicated that the comprehensive treatment package was the first treatment they had ever received. Forty percent (2/5) of participant respondents indicated that before admission to the specialized treatment home it was important to them to decrease their challenging behaviour. Sixty percent (3/5) reported that it was important to them to increase their adaptive skills. After experiencing the comprehensive treatment package, 60% (3/5) of participants found that the treatment is necessary for safety. They also felt it was acceptable for behaviour support, and has been effective in increasing their adaptive skills. Eighty percent (4/5) of participant respondents reported that the treatment is effective in decreasing their challenging behaviour and 100% (5/5) felt that it was instrumental in helping them achieve a better quality of life. In providing final thoughts, 60% (3/5) of participants felt that they are supported and are receiving the best treatment at the specialized treatment home. Eighty percent (4/5) of participants reported that, overall, the comprehensive treatment package was helpful. More specifically, participants described features of the

comprehensive treatment package that they felt were especially helpful. These included: (1) the focus on time management, productivity, and skills required for an effective lifestyle; (2) staff encouragement and reinforcement; (3) staff and the clinical team; (4) being able to visually see helpful processes (e.g., dynamic process); and (5) feeling less alone and sad about engaging in severe challenging behaviours. Suggestions for improvement were only provided by two participants. One expressed the need for opportunities to request washroom breaks during skills intervals while the other disclosed that they would like to receive marijuana and "more freedom but not too much".

Caregivers and Case Managers

Social validity results for caregivers and case managers are presented in Table 7. Among the caregivers and case managers who completed the survey, 57% (4/7) indicated that their clients (or loved ones) received another treatment prior to the featured comprehensive treatment package. Eighty-six percent (6/7) reported that these initial treatments were not helpful. Before starting the current services, it was important to 100% (7/7) of respondents to decrease their client's (loved one's) challenging behaviour and increase their client's (loved one's) adaptive skills. After experiencing the comprehensive treatment package, 100% (7/7) of caregivers and case managers agree that the treatment is: (1) necessary to ensure their client's (loved one's) safety, (2) acceptable for behaviour support; (3) effective in decreasing their client's (loved one's) challenging behaviour and increasing their adaptive skills; and (4) conducive in facilitating a better quality of life for their client (loved one). In providing final thoughts, 86% (6/7) of respondents indicated that they feel as if their client (loved one) is receiving the best treatment at the specialized treatment home. One hundred percent (7/7) reported that they feel confident in the supports, safety, and overall helpfulness of the current placement. More

specifically, caregivers and case managers described elements of the comprehensive treatment package that were particularly helpful. These included: (1) a clear routine and predictable environment; (2) the use of a daily activity schedule to provide structure specific to the client's interests and abilities; (3) consistent expectations and responses from staff; (4) support in developing healthy and appropriate self-care activities; (5) consistency and communication between all teams; (6) consistency in staff training levels; and (7) physical restraints. Suggestions for improvement include reducing staff turnover rates, increasing participation in community activities, providing additional assistance in addressing emotional triggers and coping with trauma, and adding counselling to check in on feelings.

Discussion

The purpose of the current project was to conduct a program evaluation on the use of a comprehensive behavioural treatment package encompassing differential reinforcement, safe extinction, and strategies featuring decelerative properties (e.g., overcorrection, response cost). This intervention aimed to produce marked improvements in adults with intellectual and developmental disabilities who engage in severe challenging behaviour. It was hypothesized that project outcomes would depict the correct application (i.e., treatment integrity) of the comprehensive treatment package resulting in primarily successful outcomes (i.e., substantial reduction in severe challenging behaviour and corresponding increase in adaptive skills) for participants (i.e., treatment home residents). The results suggest evidence in favour of this hypothesis. That is, project outcomes depicted the correct application of the comprehensive treatment package such that treatment integrity was greater than 80% thereby meeting the standards set out by the agency. Importantly, the treatment integrity outcomes also met the threshold for acceptable integrity performance outlined in the research literature (80%; e.g.,

Cook et al., 2015; Dart et al., 2017; Hagermoser Sanetti & Kratochwill, 2008). Project outcomes also suggest the application of the comprehensive intervention resulted in a substantial decrease in challenging behaviour from baseline to intervention for most participants. Finally, all participants appeared to be steadily learning adaptive skills, as evidenced by an impressive number of mastered target skills.

A Demonstration of Effectiveness

Regarding effectiveness, the modified CCCS allowed the researcher to explore which parameters and/or for whom the featured comprehensive behavioural treatment package would work for, as well as to tentatively demonstrate the generality, clinical utility, and acceptability of the treatment (Ghaemmaghami et al., 2021). Specific to the current project, a demonstration of generality may have been accomplished in two ways. First, the researcher evaluated the application of the comprehensive behavioural treatment package with "relatively *heterogeneous* populations, in typical settings where these treatments occur" (Ghaemmaghami et al., 2021, p. 124; Singal et al., 2014). So, although all participants were adults with intellectual and developmental disabilities, other characteristics suggested heterogeneity in sex (i.e., four males and four females), diagnoses (e.g., ASD, ADHD, BPD, generalized anxiety disorder, schizophrenia), level of intellectual disability (e.g., mild, moderate), and function of behaviour (e.g., automatic, multiply controlled). Second, the use of a hybrid nonexperimental consecutive case series afforded determining the proportion of cases in which the comprehensive treatment package produced a positive outcome (Hagopian, 2020). Specifically, the results indicated that therapeutic outcomes (i.e., challenging behaviour reduction) were observed for 75% of participants ($n = 6$). These six participants (i.e., Lily, Kevin, Riley, Micaela, Anthony, Oliver) have comorbid diagnoses with mild to moderate intellectual disability as well as multiply

controlled function of behaviour. Neutral outcomes (i.e., no change) were observed for 13% of participants (*n* = 1; Taylor), although the data point for Oliver bordered the no change line. Notably, Taylor and Oliver are the latter two cases in the series of consecutively encountered cases and therefore have experienced the comprehensive treatment package for only 2–3 years compared to 3–6 years by all other participants. Of note, Oliver's outcomes also suggest he just missed the 80% *cut-off*, as evidenced by the datapoint falling directly on the *magnitude* line. Interestingly, for Taylor and Oliver, severe challenging behaviour frequency at baseline were among the lowest observed across participants. An 80% reduction of an already low value would leave very little room for even a single incident to occur without missing this criterion. For example, if baseline rates of responding were 10, less than two episodes could occur across treatment – otherwise the 80% threshold would not be met. By contrast, if 100 episodes occurred at baseline participants could engage in up to 20 episodes and still meet the 80% criterion. Admittedly, proportionally they are the same. However, practically it seems to leave little room for participant *misstep*. In reviewing participant data, Oliver engaged in one episode of challenging behavior during baseline (27 days). The challenging behaviour occurred on day 20 in baseline. In considering this, over the next 564 days (treatment phase), Oliver only engaged in challenging behavior five times. If baseline trends had continued, one might have expected challenging behavior to occur approximately 28 more times over the next 564 days. However, this was not the case. So, although Oliver appeared to have bordered the no change line, his performance over treatment sessions (564) might suggest clinically important changes in challenging behaviour were observed.

Contra-therapeutic outcomes (i.e., an increase in challenging behaviour) may have been observed for 13% of participants (*n* = 1). Interestingly, this participant (i.e., Ginny) is unique in

that she is only diagnosed with global developmental delay (as opposed to comorbid diagnoses as with all other participants) and her challenging behaviour was found to be maintained solely by automatic function (as opposed to being multiply controlled as with all other participants). Moreover, this participant's behaviours have been correlated with biological processes. That is, she reportedly engages in more challenging behaviour approximately three to four days before her menses. With respect to medical management, she has been prescribed birth control pills (e.g., Alesse 28 20mcg/100mcg, Marvelon 21 0.15mg/0.03mg) by a physician during some portions of the treatment phase. Further, modified behavioural programming is implemented during her menses to better accommodate her. Early research suggests cyclical patterns observed in the context of severe challenging behaviour frequency, such as SIB, may correspond to changing peripheral and central endorphin and pain threshold during the menstrual cycle (Taylor et al., 1993). It may be important for future research to explore the topic of menses and severe challenging behaviour in this clinical population for an updated understanding on the interaction between the two (de Winter et al., 2011)

Collectively, these findings could suggest that the featured comprehensive treatment package may be effective for adults with intellectual and developmental disabilities who engage in truly severe challenging behaviour and have: (1) comorbid diagnoses; (2) mild to moderate intellectual disability; (3) multiply controlled function of behaviour; and (4) who have experienced the treatment for at least three years.

Treatment Acceptability

The inclusion of a social validity measure allowed the researcher to examine the acceptability of the comprehensive treatment package by clients, caregivers, and case managers. Overall, of the clients who completed the questionnaire, most felt that they are receiving the best

treatment at the specialized treatment home and all clients reported that they felt the treatment program was *instrumental* in improving their quality of life. Similarly, of the caregivers and case managers who completed the questionnaire, most indicated that they feel as if their client (or loved one) is receiving the best treatment at the specialized treatment home and all respondents felt that the comprehensive treatment packaged has helped their client (loved one). These outcomes lend support for the acceptability of the featured treatment by clients and stakeholders thereby bolstering its effectiveness.

A Demonstration of Program Evaluation Processes

Another project objective aimed to showcase the barriers and corresponding processes to overcome them so that others interested in executing a program evaluation may be better able to do this. First, setting access and subsequent project coordination between the researcher, participants, and agency staff was enabled and supported by the program's Clinical Director. This support ultimately initiated and ensured the project could be executed. Transparency, respect, prioritizing shared goals, and bi-weekly communications (i.e., virtual meetings) were vital contributors to seeing the project to completion (Rycroft-Malone et al., 2015).

A second barrier to evaluating clinical research outcomes is not being able to collect IOA with staff data (Chartier et al., 2020). Although participants featured in the current program evaluation were in treatment 24 hours a day for 365 days a year, IOA with staff data were collected by adjusting accordingly for the applied context. That is, each day of the week was considered one data collection opportunity such that conducting data collection for both dependent variables once per week per participant was equivalent to 14% of sessions (i.e., days), while collecting data for both dependent variables twice per week per participant was equivalent to 29% of sessions (i.e., days).

A third barrier commonly associated with treating severe challenging behaviour is that intervention is implemented immediately upon admission (Pritchard et al., 2018). To address the absence of a *true* baseline, the current project demonstrated the viability of using innovative tools such as Stability Check based on the Schoenfeld et al. (1956) criterion to delineate individualized *proxy* baselines (Costa & Cançado, 2012). Moreover, it may not be uncommon for clients with severe challenging behaviour to experience treatment for months and up to years at a time within an applied setting. To inspect data on a time series graph and generate meaningful interpretations, one would likely need to average participant data so that each data point reflects either weekly or bi-weekly data. Averaging data can misrepresent participant outcomes, and is influenced by variability (Ninci et al., 2015). The advantage of including a modified Brinley plot with days' and years' worth of data as showcased in the current program evaluation is that each data point informs the final data point on the graph thereby eliminating any influence of variability that may result from averaging (Kubina et al., 2017; Manolov et al., 2022; Ninci et al., 2015). It, therefore, may represent a viable alternative for future researchers who are interested in interpreting participant outcomes for large, cumulative data sets.

Project Limitations & Program Strengths and Areas for Improvement

In generating a thorough discussion, project limitations are described. In addition, because the purpose of a program evaluation is to generate an empirical answer to how well a program is doing so that it may improve service delivery (Miller, 2017), the researcher also shares some possible areas for improvement, as well as treatment program strengths informed by the data collected.

First, the researcher conducted a *hybrid* nonexperimental consecutive case series in the context of a program evaluation. That is, all participants who experienced the featured

comprehensive treatment package between October 2017 to July 2023 were included. However, as mentioned above, data collection for the current project (i.e., IOA, treatment integrity) meant greater than or equal to 50% of the data were compiled during service delivery for each participant *before* the start of data collection for the current project. Importantly, in quantitative statistics, researchers are rarely able to collect data from an entire population (Emmert-Streib & Dehmer, 2019). Instead, they rely on sampling as a mechanism for understanding phenomena (Emmert-Streib & Dehmer, 2019). So, although the researcher was only able to sample a portion of time, it may be reasonable to apply this *sampling* logic to the current project. It follows that the IOA, and treatment fidelity sample obtained may be representative of what is happening on an ongoing basis. Further, because the IOA and treatment fidelity data were collected over the course of a year, internal validity threats such as adaptation may have been offset. To elaborate, adaptation pertains to participants behaving differently than they naturally would at the start of data collection due to the novel conditions (Ledford & Gast, 2018). Thus, it is possible that at the onset of researchers being onsite, participants (and staff) may not have behaved as they naturally would. However, this was likely offset throughout the progression of data collection from May 2022 to July 2023 and therefore, may not have unduly affected outcomes. Further, researchers took great care to be as unobtrusive as possible during data collection (e.g., by standing outside of doorways or at the back of a room away from a participant's line of sight) in order to counter any reactive effects (Ledford & Gast, 2018).

Second, although the researcher aimed to collect up to 29% of IOA data for both challenging behaviour and adaptive behaviour, this was not achieved for many participants due to the uncontrolled applied setting (i.e., treatment home). That is, between May 2022 to July 2023 there were several factors that restricted onsite observation for up to weeks at a time

including: (1) clients testing positive for COVID-19 resulting in treatment home lockdowns; (2) clients getting sick; and (3) frontline staff not recording data. Moreover, specific to adaptive behaviour – the SAP was structured in a way that did not permit IOA to be collected each time a skills interval was observed. This is because for certain participants, unless they were at a "checkpoint" (i.e., probe data), staff were only required to sign off that the interval was completed and thereby did not collect data on participant correct responding. To clarify, skills programming for certain participants were structured such that they moved through several lessons (e.g., reading a passage with staff, watching a YouTube video) at their own pace during skills intervals. After a lesson had commenced staff initialed whether the lesson was completed. After several lessons, participants reached a checkpoint where they were tested on the content of the previous lessons. During these checkpoints, clinical staff recorded whether the participant responded accurately and mastery criteria (e.g., 100% accuracy over two consecutive days) for these checkpoints were required to be met before the participant was cleared to move on with subsequent lessons (see Appendix G for a sample skills lessons and corresponding checkpoint). Attempts were made to coordinate observation during checkpoints. However, given the nature of the applied setting (e.g., clinical staff schedule to observe checkpoints, clients being late to intervals, schedules being shifted to accommodate special activities), checkpoints often occurred earlier or later than scheduled. In addition, IOA data could not be collected for Oliver as a formal SAP was not enacted for this participant at the time of data collection for the current program evaluation.

Another limitation was the fact that the design was not a *controlled* consecutive case series because only A-B designs were used, and with these designs full experimental control cannot be demonstrated. However, the current project does showcase replication of the A-B

design across six of the eight participants, which may lend credibility to the outcomes observed (i.e., clinically important challenging behavior reduction; Lomas Mevers et al., 2018). Further, the absence of a baseline for adaptive behaviour presents multiple threats to internal validity (e.g., data instability; Ledford & Gast, 2018). However, when visually analyzing adaptive behaviour graphs (Figures 2–8), for all participants, a zero celerating trend occurred for at least up to 3 months after the agency initiated skills data collection before an accelerating trend was observed. Given this trend was replicated across all participants, it could be argued that all data before the agency initiated skills data collection may have also been zero celerating signaling no adaptive behaviour skill mastery. Moving forward, it may be important for the agency to enact skill mastery data collection upon admission to verify the patterns observed in the current project. Another consideration may be to digitize and keep an online record of daily skills data (e.g., DTT, probe) to facilitate data monitoring. Doing so would also better align with how challenging behavior data is entered and thus, permit monitoring by clinical team members who may not be onsite every day given competing demands (e.g., other treatment home clients they oversee).

Fourth, as described above, treatment effectiveness can be evaluated by examining its implementation in uncontrolled settings where such treatments typically occur (Singal et al., 2014). This criterion was met; however, the comprehensive treatment package was enacted at *specialized* treatment homes. That is, this setting is specialized in that it consists of a vast number of resources (e.g., staffing, specialized training in physical intervention techniques) that may not be representative of what is available at other applied, community settings, thus limiting generalizability.

Finally, the results indicated that the desired magnitude of change from the baseline level has only been achieved for 38% of participants ($n = 3$) and has not *yet* been reached for 63% of participants ($n = 5$). However, it is important to note that proxy baseline rates of challenging behaviour were low for 63% of participants ($n = 5$) in part due to the severity of challenging behaviour (e.g., aggression with weapons, eating extremely dangerous substances) engaged in by these participants (i.e., Kevin, Micaela, Anthony, Taylor, Oliver). Given the comprehensive treatment package was implemented at specialized treatment homes in which there are environmental restrictions in place, such as locked units and sanitized spaces, it is possible that some negative target behaviours may be harder to engage in. Thus, the proxy baseline may have underrepresented how often these behaviours occurred at *true* baseline (i.e., previous placement settings) thereby making it difficult to achieve the desired 80% magnitude of change from the baseline level. From a systems-level perspective, it may be important for previous support settings where participant resided to enact better data collection, so that a more accurate magnitude of change from baseline may be observed.

Treatment Integrity Error Analysis

There may be some areas for improvement as informed by the treatment integrity error analysis. Notably, for participants who engaged in higher frequencies of challenging behaviour (i.e., Lily, Ginny, Riley), percentage error was generally lower and tended to coincide with items outside of the management of behaviour domain (e.g., overnight protocol, urine accident after mechanical chair release, daily activity schedule, bathroom, and evening/bedtime). By contrast, for participants who engaged in challenging behaviour less often (i.e., Kevin, Micaela, Anthony, Taylor, Oliver), percentage error was generally higher in the management of behaviour domain. It is possible that because these participants engaged in low frequencies of challenging

behaviour, staff were unaccustomed to executing (or describing) certain management of behaviour items and were therefore not able to recall the protocol readily.

Interestingly, higher error percentages were also observed in the preventative strategies domain across several participants (i.e., Riley, Micaela, Anthony, Taylor, and Oliver). For most participants, these items involved wearing protective equipment, keeping the environment free of any objects that may be thrown or used as a weapon, and locking bedroom doors. Given most of these participants engage in low frequencies of challenging behaviour, staff may not have felt it necessary to wear all protective gear (e.g., bump cap, Kevlar sleeves, *and* gloves) or remove all unnecessary objects from the environments. It may be important for the house manager or other supervisory staff to provide regular/daily feedback (i.e., verbal reminders, praise statements) so that staff become better at enacting these items.

Regarding other items that coincided with a high percentage of errors, it may be important for the clinical team to automate error analyses so that they can review the errors informed by all the TICs collected in a month. Collectively, these data would advise the targeted trainings that could be conducted once per month on the items associated with high percentage errors.

Importantly, treatment integrity error analysis outcomes also showcased program areas of strength. For instance, consistent low percentages of error were observed for participants who engaged in higher frequencies of challenging behaviour (i.e., Lily, Ginny). This indicates that staff who regularly implemented management of behaviour domain components did so with high levels of integrity. Additionally, across all eight participants, minimal error percentages were observed for all items in the data collection domain (0%–37%). These items specifically pertained to whether data was recorded: (1) as required; (2) to its entirety with no information

left blank or missing; and (3) accurately. Thus, the high integrity levels for these items are particularly impressive as it lends confidence to the accurateness and completeness of the data reported.

Social Validity

Recommendations by caregivers and case managers on the social validity questionnaire included reducing staff turnover rates, increasing participation in community activities, providing additional assistance in addressing emotional triggers and coping with trauma, and adding counselling to check in on feelings. Although suggestions around reducing staff turnover may be tricky to address. That is, standard staff turnover rates in the disabilities sector (not to mention clients who engage in severe challenging behaviour) is approximately 43% (National Core Indicators, 2022). However, it may be important for the agency to communicate their staff retention efforts to stakeholders. Outlining these efforts clearly to stakeholder may be a way to acknowledge this recommendation, maintain transparency, and ultimately foster rapport between parties. In addition, the recommendation to incorporate additional counselling may be costly. However, pending budgetary approval and taking into consideration whether it is clinically appropriate for a participant (e.g., participant with significant cognitive impairment may not benefit from standard counselling; Byrne & O'Mahony, 2020; Unwin et al., 2016), this suggestion could be considered for inclusion during Phase 3 (Generalization and Maintenance) when participants are preparing to transition into the community.

Respondents, both clients and caregivers, also made note of several program strengths. More specifically, 75% (3/4) of participants who answered the open-ended question pertaining to program aspects they found most helpful stated that staff and the clinical team were instrumental in supporting their success. Further, one participant noted that they found programming to be

directly useful to skills (e.g., time management, productivity) they believe is important for an effective lifestyle. Of the five caregivers and case managers who answered the same open-ended question, 80% (4/5) reported that the structure and consistency offered by the program (e.g., daily activity schedule, staff expectations, staff training, communication between all teams) played a substantial role in the success of their client (loved one).

Future Research

Future research could explore addressing some of this project's limitations, as well as build on the featured processes. For instance, given proxy baselines were generated retrospectively for all participants, future research may look to conducting a Monte Carlo analysis for each participant to compare clinical data for challenging behaviour to simulated samples of behavioural data in order to discern the likelihood of the clinically obtained challenging behaviour data occurring due to chance (Friedel et al., 2021). That is, a Monte Carlo analysis for each participant would quantify the likelihood of the independent variable (i.e., comprehensive treatment package) being related to the variable of interest (i.e., challenging behaviour; Friedel et al., 2021). The question of interest, here, may be as follows: Was behaviour elevated in baseline relative to intervention? This analysis could be conducted using readily available Applications (see Friedel et al., 2021; https://shiny.georgiasouthern.edu/BA_Monte_Carlo/). Importantly, enacting this approach in the current study may have been premature given existing literature has not yet established a standard for this analysis within the context of CCCS and/or program evaluations. That is, researchers will first need to uncover the common significance outcomes (*p values*) observed, as informed by existing program evaluation data. Future CCCS and/or program evaluation studies that enact Monte Carlo simulations can then use this information to advise a standard against

which to compare their outcomes. Doing so would lend further support to the believability of their results.

Another area for future research pertains to the social validity questionnaire developed for the current program evaluation. Presently, there are no standard methods for reporting the reliability and validity of social validity measures and, as such, these measures generally lack psychometric rigor (Anderson et al., 2021). Future studies could therefore look to establish test-retest reliability, predictive validity, and construct validity of the social validity questionnaire used in the current program evaluation (Anderson et al., 2021). Doing so may not only increase the confidence of the acceptability results obtained here, but may also encourage and facilitate the administration of social validity questionnaires to *clients* as well as stakeholders. This is particularly important as a review of social validity literature found that social validity measures were not frequently reported (i.e., 12%; Ferguson et al., 2019), goals selected for treatment were least likely to be assessed (i.e., 12%; Ferguson et al., 2019), and direct recipients of behaviour-change intervention were underrepresented (i.e., less than 3% of applications; Hanley, 2010).

Conclusion

More broadly, the aim of the current program evaluation was three-fold: (1) to add to a relatively understudied area (i.e., severe challenging behaviour in adults with intellectual and developmental disabilities); (2) to speak to program effectiveness (e.g., under which parameters and/or for whom a comprehensive behavioural intervention package featuring core tenants of applied behaviour analysis will work for); and (3) to inform other works interested in program evaluations. Through successful collaboration with the partnering agency, the researcher was able to recruit participants from an understudied population and uncover trends pertaining to program effectiveness using a systematic program evaluation. The results of this project will

hopefully inform treatment implementation for individuals with similar profiles (i.e., severe

challenging behaviour) and encourage continued collaboration with community partners to

evaluate and disseminate programming outcomes.

References

Agosta, J. M., Close, D. W., Hops, H., & Rusch, F. R. (1980). Treatment of self-injurious behaviour through overcorrection procedures. *Research and Practice for Persons with Severe Disabilities*, *5*(1), 5-12. https://doi.org/10.1177/15407968000500102

Ahearn, W. H., Clark, K. M., & MacDonald, R. P. F. (2007). Assessing and treating vocal stereotypy in children with Autism. *Journal of Applied Behavior Analysis*, *40*(2), 263-275. https://doi.org/10.1901/jaba.2007.30-06

Anderson, B. M., Kozluk, A., Morgan, M. C. M., MacDonald, M. A., Friedel, J., & Cox, A. D. (2023). Exploring characteristics of interventionist training associated with improved learner outcomes: A meta-analysis. *Journal of Behavioral Education*. https://doi.org/10.1007/s10864-022-09504-2

Anderson, C. M., & Long, E. S. (2002). Use of a structured descriptive assessment methodology to identify variables affecting problem behavior. *Journal of Applied Behavior Analysis*, *35*(2), 137-154. https://doi.org/10.1901/jaba.2002.35-137

Anderson, J., & Le, D. D. (2011). Abatement of intractable vocal stereotypy using an overcorrection procedure. *Behavioral Interventions*, *26*(2), 134-146. https://doi.org/10.1002/bin.326

Anderson, R., Taylor, S., Taylor, T., & Virues-Ortega, J. (2021). Thematic and textual analysis methods for developing social validity questionnaires in applied behavior analysis. *Behavioral Interventions*, *37*(3), 732-753. https://doi.org/10.1002/bin.1832

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

Bagwell, A., Barnett, M., & Falcomata, T. S. (2022). Response cost and time-out from

reinforcement. In J. B. Leaf, J. H., Cihon, J. L. Ferguson, & M. J. Weiss (Eds.),

Handbook of applied behavior analysis interventions for autism (1st ed., pp. 479-496).

Springer. https://doi.org/10.1007/978-3-030-96478-8

Bartlett, S. M., Rapp, J. T., Krueger, T. K., & Henrickson, M. L. (2011). The use of response

cost to treat spitting by a child with autism. *Behavioral Interventions*, *26*(1), 76-83.

https://doi.org/10.1002/bin.322

Behavior Analyst Certification Board. (2020). *Ethics code for behavior analysts*.

https://www.bacb.com/wp-content/uploads/2022/01/Ethics-Code-for-Behavior-Analysts-

220316-2.pdf

Blampied, N. M. (2017). Analyzing therapeutic change using modified Brinley plots: History,

construction, and interpretation. *Behavior Therapy*, *48*(1), 115-127.

https://doi.org/10.1016/j.beth.2016.09.002

Bonner, A. C., & Borrero, J. C. (2019). Differential reinforcement of low rate schedules reduce

severe problem behavior. *Behavior Modification*, *42*(5), 747-764.

https://doi.org/10.1177/0145445517731723

Bowring, D. L., Totsika, V., Hastings, R. P., Toogood, S., & Griffith, G. M. (2017). Challenging

behaviours in adults with an intellectual disability: A total population study and

exploration of risk indices. *British Journal of Clinical Psychology, 56*(1), 16-32.

https://doi.org/10.1111/bjc.12118

Busch, L. P. A., Saini, V., Zorzos, C., & Duyile, L. (2018). Treatment of life-threathening pica

with 5-year follow-up. *Advances in Neurodevelopmental Disorders*, *2*(4), 335-343.

https://doi.org/10.1007/s41252-018-0053-9

Butterill, D., Lin, E., Durbin, J., Lunsky, Y., Urbanoski, K., & Soberman, H. (2009). From

  hospital to home: The transitioning of alternate level of care and long-stay mental health

  clients. *Health Systems Research and Consulting Unit, Center for Addiction and Mental*

  Health.

Byrne, G., & O'Mahony, T. (2020). Acceptance and commitment therapy (ACT) for adults with

  intellectual disabilities and/or autism spectrum (ASC): A systematic review. *Journal of*

  Contextual Behavioral Science, *18*, 247-255. https://doi.org/10.1016/j.jcbs.2020.10.001

Carr, E. G., & Durand, V. M. (1985). Reducing behavior problem through functional

  communication training. *Journal of Applied Behavior Analysis*, *18*(2), 111-126.

  https://doi.org/10.1901/jaba.1985.18-111

Carr, J. E., Severtson, J., & Lepper, T. (2009). Noncontingent reinforcement is an empirically

  supported treatment for problem behavior exhibited by individuals with developmental

  disabilities. *Research in Developmental Disabilities*, *30*(1), 44-57.

  https://doi.org/10.1016/j.ridd.2008.03.002

Carroll, R. A., Kodak, T., Fisher, W. W. (2013). An evaluation of programmed treatment-

  integrity errors during discrete-trial instruction. *Journal of Applied Behavior Analysis*,

  46(2), 379-394. https://doi.org/10.1002/jaba.49

Cengher, M., Ramazon, N. H., & Strohmeier, C. W. (2020). Using extinction to increase

  behavior: Capitalizing on extinction-induced response variability to establish mands with

  autoclitic frames. *The Analysis of Verbal Behavior*, *36*(1), 102-114.

  https://doi.org/10.1007/s40616-019-00188-w

Chartier, K., & Feldman, M. (2015). Outcome evaluation of a specialized treatment home for

adults with dual diagnosis and challenging behaviour. *Journal on Developmental Disabilities*, *21*(2), 69-82.

Chazin, K. T., Velez, M. S., & Ledford, J. (2022). Reducing escape without escape extinction: A systematic review and meta-analysis of escape-based intervention. *Journal of Behavioral Education*, *31*(1), 186-215. https://doi.org/10.1007/s10864-021-09453-2

Chowdury, M., & Benson, B. A. (2011). Use of differential reinforcement to reduce behavior problems in adults with intellectual disabilities: A methodological review. *Research in Developmental Disabilities*, *32*(2), 383-394. https://doi.org/10.1016/j.ridd.2010.11.015

Conyers, C., Miltenberger, R., Maki, A., Barenz, R., Jurgens, M., Sailer, A., Haugen, M., & Kopp, B. (2004). A comparison of response cost and differential reinforcement of other behavior to reduce disruptive behavior in a preschool classroom. *Journal of Applied Behavior Analysis*, *38*(3), 405-409. https://doi.org/10.1901/jaba.2005.7104

Cook, J. E., Subramaniam, S., Brunson, L. Y., Larson, N. A., Poe, S. G., & St. Peter, C. C. (2015). Global measures of treatment integrity may mask important errors in discrete-trial training. *Behavior Analysis in Practice*, *8*(1), 37-47. https://doi.org/10.1007/s40617-014-0039-7

Cooper, J. O., Heron, T. E., & Heward, W. L. (2020). *Applied behavior analysis* (3rd ed.). Pearson Education.

Costa, C. E., & Cançado, C. R. X. (2012). Stability Check: A program for calculating the stability of behavior. *Mexican Journal of Behavior Analysis*, *38*(1), 61-71.

Cox, A., Leung, J., Anderson, B. M., & Morgan, M-C. (2021). Examining research trends in the

treatment of adults with problem behaviour and intellectual and developmental

disabilities: A brief report. *Behavioural Development, 26*(1), 14-28.

https://doi.org/10.1037/bdb0000100

Cox, A. D., Pontone, M., & Gunnarsson, K. F. (2022). Program evaluation of in-patient

treatment units for adults with acquired brain injury and challenging behaviour. *Brain*

Injury, 1-11. https://doi.orh/10.1080/02699052.2022.2033840

Cox, A. D., & Virues-Ortega, J. (2022). Long-term functional stability of problem behavior

exposed to psychotropic medications. *Journal of Applied Behavior Analysis*, *55*(1), 214-

229. https://doi.org/10.1002/jaba.873

Critchfield, T. S., & Reed, D. D. (2017). The fuzzy concept of applied behavior analysis

research. *The Behavior Analyst*, *40*(1), 123-159. https://doi.org/10.1007/s40614-017-

0093-x

Dart, E. H., Radley, K. C., Furlow, C. M., & Murphy, A. N. (2017). Using behavioral skills

training to teach high school students to implement discrete trial training. *Behavior*

Analysis: Research and Practice, *17*(3), 237-249. https://doi.org/10.1037/bar0000075

Dart, E. H., Radley, K. C., Mason, B. A., & Allen, J. P. (2018). Addressing escape-maintained

behavior for students with developmental disabilities: A systematic review of school-

based interventions. *Psychology in Schools*, *55*(3), 295-304.

https://doi.org/10.1002/pits.22108

de Winter, C. F., Jansen, A. A. C., & Evenhuis, H. M. (2011). Physical conditions and

challenging behaviour in people with intellectual disability: A systematic review. *Journal*

of Intellectual Disability Research, *55*(7), 675-698. https://doi.org/10.1111/j.1365-

2788.2011.01390.x

DeRosa, N. M., Novak, M. D., Morley, A. J., & Roane, H. S. (2019). Comparing response

blocking and response interruption/redirection on levels of motor stereotypy: Effects of

data analysis procedures. *Journal of Applied Behavior Analysis*, *52*(4), 1021-1033.

https://doi.org/10.1002/jaba.644

Emerson, E., Kiernan, C., Alborz, A., Reeves, D., Mason, H., Swarbrick, R., Mason, L., &

Hatton, C. (2001). The prevalence of challenging behaviours: A total population study.

Research in Developmental Disabilities, *22*(1), 77-93. https://doi.org/10.1016/S0891-

4222(00)00061-5

Emmert-Streib, F., & Dehmer, M. (2019). Understanding statistical hypothesis testing: The logic

of statistical inference. *Machine Learning and Knowledge Extraction*, *1*(3), 945-961.

https://doi.org/10.3390/make1030054

Fahmie, T. A., & Iwata, B. A. (2011). Topographical and functional properties of precursors to

severe problem behavior. *Journal of Applied Behavior Analysis*, *44*(4), 993-997.

https://doi.org/10.1901/jaba.2011.44-993

Falcomata, T. S., Roane, H. S., Hovanetz, A. N., Kettering, T. L., & Keeney, K. M. (2004). An

evaluation of response cost in the treatment on inappropriate vocalizations maintained by

automatic reinforcement. *Journal of Applied Behavior Analysis*, *37*(1), 83-87.

https://doi.org/10.1901/jaba.2004.37-83

Falligant, J. M., & Pence, S. T. (2020). Interventions for inappropriate sexual behavior in

individuals with intellectual and developmental disabilities: A brief review. *Journal of

Applied Behavior Analysis*, *53*(3), 1316-1320. https://doi.org/10.1002/jaba.716

Ferguson, J. L., Cihon, J. H., Leaf, J. B., Van Meter, S. M., McEachin, J., & Leaf, R. (2019).

Assessment of social validity trends in the journal of applied behavior analysis. *European Journal of Behavior Analysis*, *20*(1), 146-157. https://doi.org/10.1080/15021149.2018.1534771

Fiani, T., & Jessel, J. (2022). Practical functional assessment and behavioral treatment of challenging behavior for clinically based outpatient services: A consecutive case series evaluation. *Education and Treatment of Children*. https://doi.org/10.1007/s43494-022-00071-9

Fisher, W. W., DeLeon, I. G., Rodriguez-Catter, V., & Keeney, K. M. (2004). Enhancing the effects of extinction on attention-maintained behavior through noncontingent delivery of attention or stimuli identified via a competing stimulus assessment. *Journal of Applied Behavior Analysis*, *37*(2), https://doi.org/10.1901/jaba.2004.37-171

Fisher, W. W., Grace, N. C., & Murphy, C. (1996). Further analysis of the relationship between self-injury and self-restraint. *Journal of Applied Behavior Analysis*, *29*(1), 103-106. https://doi.org/10.1901/jaba.1996.29-103

Friedel, J. E., Cox, A., Galizio, A., Swisher, M., Small, M. L., & Perez, S. (2021). Monte Carlo analyses for single-case experimental designs: An untapped resource for applied behavioral researchers and practitioners. *Perspectives on Behavior Science*, *45*(1), 209-237. https://doi.org/10.1007/s40614-021-00318-7

Foxx, R. M., & Azrin, N. H. (1972). Restitution: A method of eliminating aggressive-disruptive behaviour of retarded and brain damaged patients. *Behaviour Research and Therapy*, *10*(1), 15-27. https://doi.org/10.1016/0005-7967(72)90003-4

Foxx, R. M., & Garito, J. (2007). The long term successful treatment of the very severe

behaviors of a preadolescent with autism. *Behavioral Interventions*, *22*(1), 69-82.

https://doi.org/10.1002/bin.232

Geiger, K. B., Carr, J. E., & LeBlanc, L. A. (2010). Function-based treatments for escape-maintained problem behavior: A treatment-selection model for practicing behavior analysts. *Behavior Analysis in Practice*, *3*(1), 22-32. https://doi.org/10.1007/BF03391755

Ghaemmaghami, M., Hanley, G. P., & Jessel, J. (2021). Functional communication training: From efficacy to effectiveness. *Journal of Applied Behavior Analysis, 54*(1), 122-143. https://doi.org/10.1002/jaba.762

Gover, H. C., Fahmie, T. A., & McKeown, C. A. (2019). A review of environmental enrichment as treatment for problem behavior maintained by automatic reinforcement. *Journal of Applied Behavior Analysis*, *52*(1), 299-314. https://doi.org/10.1002/jaba.508

Hagermoser Sanetti, L. M., & Kratochwill, T. R. (2008). Treatment integrity in behavioral consultation: Measurement, promotion, and outcomes. *International Journal of Behavioral Consultation and Therapy*, *4*(1), 95-114. https://doi.org/10.1037/h0100835

Hagopian, L. P. (2020). The consecutive controlled case series: Design, data-analytics, and reporting methods supporting the study of generality. *Journal of Applied Behavior Analysis*, *53*(2), 596-619. https://doi.org/10.1002/jaba.691

Hagopian, L. P., Fisher, W. W., Sullivan, M. T., Acquisto, J., & LeBlanc, L. A. (1998). Effectiveness of functional communication training with and without extinction and punishment: A summary of 21 inpatient cases. *Journal of Applied Behavior Analysis*, *31*(2), 211-235. https://doi.org/10.1901/jaba.1998.31-211

Halpern, L. F., & Andrasik, F. (1986). The immediate and long-term effectiveness of

overcorrection in treating self-injurious behavior in a mentally retarded adult. *Applied Research in Mental Retardation*, *7*(1), 59-65. https://doi.org/10.1016/0270-3092(86)90018-4

Hanley, G. P. (2010). Towards effective and preferred programming: A case for the objective measurement of social validity with recipients of behavior-change programs. *Behavior Analysis in Practice*, *3*(1), 13-21. https://doi.org/10.1007/BF03391754

Hanley, G. P., Piazza, C. C., Fisher, W. W., & Maglieri, K. A. (2005). On the effectiveness of and preference for punishment and extinction components of function-based interventions. *Journal of Applied Behavior Analysis*, *38*(1), 51-65. https://doi.1901/jaba.2005.6-04

Homorodean, C., Iancu, A. C., Lecuta, D., Balanescu, S., Dregoesc, I. M., Spinu, M., Ober, M., Tataru, D., Olinic, M., Bindea, D., & Olinic, D. (2019). New predictors of early and late outcomes after primary percutaneous coronary intervention in patients with ST-segment elevation myocardial infarction and unprotected left main coronary artery culprit lesion. *Journal of Interventional Cardiology*, *2019*(2), 1-9. https://doi.org/10.1155/2019/8238972

Iwata, B., & DeLeon, I. (2005). *The functional analysis screening tool*. The Florida Center on Self-Injury, University of Florida.

Iwata, B., Pace, G. M., Cowdery, G. E., & Miltenberger, R. (1994). What makes extinction work: An analysis of procedural form and function. *Journal of Applied Behavior Analysis*, *27*(1), 131-144. https://doi.org/10.1901/jaba.1994.27-131

Iwata, B. A., Pace, G. M., Kalsher, M. J., Cowdery, G. E., & Cataldo, M. F. (1990).

Experimental analysis and extinction of self-injurious escape behavior. *Journal of Applied Behavior Analysis*, *23*(1), 11-27. https://doi.org/10.1901/jaba.1990.23-11

Kliebert, M. L., Tiger, J. H., & Toussaint, K. A. (2011). An approach to identifying the conditions under which response interruption will reduce automatically reinforced problem behavior. *Behavior Analysis in Practice*, *4*(1), 17-26. https://doi.org/10.1007/BF03391771

Kubina, R. M., Kostewicz, D. E., Brennan, K. M., & King, S. A. (2017). A critical review of line graphs in behavior analytic journals. *Educational Psychology Review*, *29*(3), 583-598. https://doi.org/10.1007/s10648-015-9339-x

Kurtz, P. F., Leoni, M., & Hagopian, L. P. (2021). Behavioral approaches to assessment and early intervention for severe problem behavior in intellectual and developmental disabilities. *Pediatric Clinics of North America, 67*(3), 499-511. https://doi.org/10.1016/j.pcl.2020.02.005

Lalli, J. S., Casey, S., Goh, H., & Merlino, J. (1994). Treatment of escape-maintained aberrant behavior with escape extinction and predictable routines. *Journal of Applied Behavior Analysis*, *27*(4), 705-714. https://doi.org/10.1901/jaba.1994.27-705

Lambert, J. M., Sandstrom, A. L., Hodapp, R. M., Copeland, B. A., Paranczak, J. L., Macdonald, M. J., & Houchins-Juarez, N. J. (2022). Revisiting the social validity of services rendered through a university-based practicum addressing challenging behavior. *Journal of Applied Behavior Analysis*, *55*(4), 1220-1238. https://doi.org/10.1002/jaba.939

Ledford, J. R., & Gast, D. L. (2018). *Single case research methodology* (3rd ed.). Routledge.

Lerman, D. C., & Iwata, B. A. (1995). Prevalence of the extinction burst and its attenuation

during treatment. *Journal of Applied Behavior Analysis*, *28*(1), 93-94.

https://doi.org/10.1901/jaba.1995.28-93

Lerman, D. C., & Vorndran, C. M. (2002). On the status of knowledge for using punishment

implications for treating behavior disorders. *Journal of Applied Behavior Analysis*, *35*(4),

431-464. https://doi.org/10.1901/jaba.2002.35-431

Lindberg, J. S., Iwata, B. A., Roscoe, E. M., Worsdell, A. S., & Hanley, G. P. (2003). Treatment

efficacy of noncontingent reinforcement during brief and extended application. *Journal*

of Applied Behavior Analysis, *36*(1), 1-19. https://doi.org/10.1901/jaba.2003.36-1

Linder, B. (2014). *Daily Activity Schedule-Quality Rating Scale*.

Linscheid, T. R., Copeland, A. P., Jacobstein, D. M., & Smith, J. L. (1981). Overcorrection

treatment for nighttime self-injurious behavior in two normal children. *Journal of*

Pediatric Psychology, *6*(1), 29-35. https://doi.org/10.1093/jpepsy/6.1.29

Lomas Mevers, J., Muething, C., Call, N. A., Scheithauer, M., & Hewett, S. (2018). A

consecutive case series analysis of a behavioral intervention for enuresis in children with

developmental disabilities. *Developmental Neurorehabilitation*, *21*(5), 336-344.

https://doi.org/10.1080/17518423.2018.1462269

Lowe, K., Allen, D., Jones, E., Brophy, S., Moore, K., & James, W. (2007). Challenging

behaviours: Prevalence and topographies. *Journal of Intellectual Disability Research,*

51(8), 625-636. https://doi.org/10.1111/j.1365-2788.2006.00948.x

Luiselli, J. K., Sperry, J. M., & Draper, C. (2015). Social validity assessment of physical restraint

intervention by care providers of adults with intellectual and developmental disabilities.

Behavior Analysis in Practice, *8*(2), 170-175. https://doi.org/10.1007/s40617-015-00820z

Luiselli, J. K., Sperry, J. M., Draper, C., & Richards, C. (2017). Parent-guardian evaluation of

physical restraint among adults with intellectual disability: A social validity assessment. *Advances in Neurodevelopmental Disorders*, *1*(2), 73-78. https://doi.org/10.1007/s41252-017-0013-9

Lunsky, Y., Khuu W., Tadrous, M., Vigod, S., Cobigo, V., & Gomes, T. (2018). Antipsychotic use with and without comorbid psychiatric diagnosis among adults with intellectual disabilities. *Canadian Journal of Psychiatry, 63*(6), 361-369. https://doi.org/10.1177/0706743717727240

Lydon, S., Healy, O., O'Reilly, M. & McCoy, A. (2013). A systematic review and evaluation of response redirection as a treatment for challenging behavior in individuals with developmental disabilities. *Research in Developmental Disabilities*, *34*(10), 3148-3158. https://doi.org/10.1016/j.ridd.2013.06.010

Manente, C. J., & LaRue, R. H. (2017). Treatment of self-injurious behavior using differential punishment of high rates of behavior (DPH). *Behavioral Interventions*, *32*(3), 262-271. https://doi.org/10.1002/bin.1477

Manolov, R., & Tanious, R. (2022). Assessing consistency in single-case data features using modified brinley plots. *Behavior Modification, 46*(3), 581-627. https://doi.org/10.1177/0145445520982969

Manolov, R., Tanious, R., & Fernandez-Castilla, B. (2022). A proposal for the assessment of replication of effects in single-case experimental designs. *Journal of Applied Behavior Analysis*, 9999, 1-28. https://doi.org/10.1002/jaba.923

McAdam, D., & Knapp, V. M. (2021). Overcorrection. In F. R. Volkmar (Ed.), *Encyclopedia of autism spectrum disorders* (2nd ed., pp. 3250-3253). Springer. https://doi.org/10.1007/978-3-319-91280-6

McConnachie, G., & Carr, E. G. (1997). The effects of child behavior problems on the maintenance of intervention fidelity. *Behavior Modification*, *21*(2), 123-158. https://doi.org/10.1177/01454455970212001

Melanson, I. J., & Fahmie, T. A. (2023). Functional analysis of problem behavior: A 40-year review. *Journal of Applied Behavior Analysis*, *56*(2), 262-281. https://doi.org/10.1002/jaba.983

Miller, K. L. (2017). The use of evaluation in treatment programs for children with autism. *Behavior Analysis in Practice*, *10*(1), 35-44. https://doi.org.10.1007/s40617-016-0130-3

Miltenberger, R. G., & Fuqua, R. W. (1981). Overcorrection: A review and critical analysis. *The Behavior Analyst*, *4*(2), 123-141. https://doi.org/10.1007/BF03391860

Ministry of Community and Social Services (MCSS). (2008). *O. Reg. 299/10: Quality Assurance Measures under Services and Supports to Promote the Social Inclusion of Persons with Developmental Disabilities Act, 2008, c. 14.* https://www.ontario.ca/laws/regulation/100299

Matson, J. L., Stephens, R. M., & Smith, C. (1978). Treatment of self-injurious behaviour with overcorrection. *Journal of Mental Deficiency Research*, *22*(3), 175-178. https://doi.org/10.1111/j.2788.1978.tb00974.x

Matson, J. L., & Vollmer, T. R. (1995). *The questions about behavioral function (QABF) user's guide.* Scientific Publishers.

National Core Indicators (2022). *National core indicators intellectual and developmental disabilities 2020 staff stability survey report.* https://legacy.nationalcoreindicators.org/upload/core-indicators/2020StaffStabilitySurveyReport_FINAL.pdf

Newcomb, E. T., & Hagopian, L. P. (2018). Treatment of severe problem behaviour in children with autism spectrum disorder and intellectual disabilities. *International Review of Psychiatry*, *30*(1), 96-109. https://doi.org/10.1080/09540261.2018.1435513

Ninci, J., Vannet, K. J., Willson, V., & Zhang, N. (2015). Interrater agreement between visual analysts of single-case data: A meta-analysis. *Behavior Modification*, *39*(4), 510-541. https://doi.org/10.1177/0145445515581327

O'Hara, P. J., Hertzer, N. R., Krajewski, L. P., Tan, M. Xiong, X., & Beven, E. G. (1995). Ten-year experience with abdominal aortic aneurysm repair in octogenarians: Early results and late outcomes. *Journal of Vascular Surgery*, *21*(5), 830-837. https://doi.org/10.1016/s0741-5214

Oropeza, M. E., Fritz, J. N., Nissen, M. A., Terrel, A. S., & Phillips, L. A. (2018). Effects of therapist-worn protective equipment during functional analysis of aggression. *Journal of Applied Behavior Analysis, 51*(3), https://doi.org/10.1002/jaba.457

Peters, L. C., & Thompson, R. H. (2013). Some indirect effects of positive practice overcorrection. *Journal of Applied Behavior Analysis*, *46*(3), 613-625. https://doi.org/10.1002/jaba.63

Petscher, E. S., Rey, C., & Bailey, J. S. (2009). A review of empirical support for differential reinforcement of alternative behavior. *Research in Developmental Disabilities, 30*(3), 409-425. https://doi.org/10.1016/j.ridd.2008.08.008

Phillips, C. L., Iannaccone, J. A., Rooker, G. W., & Hagopian, L. P. (2017). Noncontingent reinforcement for the treatment of severe problem behavior: An analysis of 27 consecutive applications. *Journal of Applied Behavior Analysis, 50*(2), 357-376. https://doi.org/10.1002/jaba.376

Poppes, P., van der Putten, A., & Vlaskamp, C. (2010). Frequency and severity of challenging behaviour in people with profound intellectual and multiple disabilities. *Research in Developmental Disabilities*, *31*(6), 1269-1275. https://doi.org/10.1016/j.ridd.2010.07.017

Pritchard, J. (2013). Response cost. In F. Volkmar (Ed.), *Encyclopedia of autism spectrum disorders*, (2nd ed., 2573-2574). Springer. https://doi.org/10.1007/978-1-4419-1698-3_1308

Pritchard, D., Penney, H., & Mace, F. C. (2017). The ACHIEVE! Program: A point and level system for reducing severe problem behavior. *Behavioral Interventions*, *33*(1), 41-55. https://doi.org/10.1002/bin.1506

Rapport, M. D., & Begolli, G. (2007). Response cost. In M. Herson, J. Rosqvist, A. M. Gross, R. S. Drabman, G. Sugai, & R. Horner (Eds.), *Encyclopedia of behavior modification and cognitive behavior therapy* (pp. 1003-1006). Sage Publications. https://doi.org/10.4135/9781412950534

Reid, D. H., Parsons, M. B., Phillips, J. F., Green, C. W. (1993). Reduction of self-injurious hand mouthing using response blocking. *Journal of Applied Behavior Analysis*, *26*(1), 139-140. https://doi.org/10.1901/jaba.1993.26-139

Rooker, G. W., Jessel, J., Kurtz, P. F., & Hagopian, L. P. (2013). Functional communication training with and without alternative reinforcement and punishment: An analysis of 58 application. *Journal of Applied Behavior Analysis*, *46*(4). https://doi.org/10.1002/jaba.76

Rycroft-Malone, J., Burton, C., Wilkinson, J. E., & Harvey, G. (2015). Collective action for knowledge mobilisation: A realist evaluation of the collaborations for leadership in applied health research and care. *Health Services and Delivery Research*, *3*(44), 1-166. https://doi.org/10.3310/hsdr03440

Saini, V., Fisher, W. W., Pisman, M. D. (2017). Persistence during and resurgence following noncontingent reinforcement implemented with and without extinction. *Journal of Applied Behavior Analysis*, *50*(2), 377-392. https://doi.org/10.1002/jaba.380

Salameh, J., & Linder, B. (2016, May 27-31). An assessment protocol and outcome data for "safe extinction" with adults with severe behavioural disorder and developmental disabilities. In B. Linder (Chair), *Special ethical issues in intrusive programming* [Symposium]. Association for Behavior Analysis International 42nd Annual Convention, Chicago, IL, United States.

Schoenfeld, W. N., Cumming, W. W., & Hearst, E. (1956). On the classification of reinforcement schedules. *Proceedings of the National Academy of Sciences of the United States of America*, *42*(8), 5630570. https://doi.org/10.1073/pnas.42.8.563

Şenkal, A., Toper, O., & Kiyak, Ü. E. (2023). Noncontingent reinforcement in decreasing problem behaviors: A systematic review. *Kalem Uluslararasi Egitim ve Insan Bilimleri Dergisi*, *13*(2), 407-444. https://doi.org/10.23863/kalem.2023.255

Shepley, C., Shepley, S. B., Allday, R. A., Tyner-Wilson, M., & Larrow, D. (2021). Evaluation of a brief family-centred service provision model for treating children's severe behaviour: A retrospective consecutive case series analysis. *Behaviour Analysis in Practice*, *14*(1), 86-96. https://doi.org/10.1007/s40617-020-00487-y

Sheppard-Jones, K., Thompson Prout, H., Kleinert, H., Taylor, S. J. (2005). Quality of life dimensions for adults with developmental disabilities: A comparative study. *Mental Retardation, 43*(4), 281-291. https://doi.org/10.1352/0047-6765(2005)43[291:QOLDFA]2.0.CO;2

Singal, A. G., Higgins, P. D. R., & Waljee, A. K. (2014). A primer on effectiveness and efficacy

trials. *Clinical and Translational Gastroenterology*, *5*(1), e45.

https://doi.org/10.1038/ctg.2013.13

Steinhauser, H. M. K., Ahearn, W. H., Foster, R. A., Jacobs, M., Doggett, C. G., & Goad, M. S.

(2021). Examining stereotypy in naturalistic contexts: Differential reinforcement and

context-specific redirection. *Journal of Applied Behavior Analysis*, *54*(4), 1420-1436.

https://doi.org/10.1002/jaba.847

Sturgis, P., Roberts, C., & Smith, P. (2012). Middle alternatives revisited: How the neither/nor

response acts as a way of saying "I don't know"?. *Sociological Methods and Research*,

43(1), 15-38. https://doi.org/10.1177/0049124112452527

Taylor, D. V., Rush, D., Hetrick, W. P., & Sandman, C. A. (1993). Self-injurious behavior within

the menstrual cycle of women with mental retardation. *American Journal of Mental*

Retardation, *97*(6), 659-664.

Tereshko, L., & Sottolano, D. (2017). The effects of an escape extinction procedure using

protective equipment on self-injurious behavior. *Behavioral Interventions*, *32*(2), 152-

159. https://doi.org/10.1002/bin.1475

Unwin, G., Tsimopoulou, I., Stenfert Kroese, B., & Azmi, S. (2016). Effectiveness of cognitive

behavioural therapy (CBT) programmes for anxiety or depression in adults with

intellectual disabilities: A review of the literature. *Research in Developmental*

Disabilities, *51-52*, 60-75. https://doi.org/10.1016/j.ridd.2015.12.010

Valdovinos, M. G. (2019). Psychotropic medication in intellectual and developmental

disabilities: Patterns of use and recommendations for monitoring effects. *Current*

Developmental Disorders Reports, *6*(4), 195-201. https://doi.org/10.1007/s40474-019-

00179-5

Vollmer, T. R., Marcus, B. A., & Ringdahl, J. E. (1995). Noncontingent escape as treatment for self-injurious behavior maintained by negative reinforcement. *Journal of Applied Behavior Analysis*, *28*(1), 15-26. https://doi.org/jaba.1995.28-15

Wells, F. L., & Ruesch, J. (1972). *The mental examiners' handbook comp.* Psychological Corporation.

Wilder, D. A., Bevacqua, J. A., Hodges, A. C., Ertel, H., & Luong, N. (2019). Comparison of function-based, nonfunction-based, and combined treatments for escape-maintained aggression in a child with attention deficit hyperactivity disorder: A case study. *Clinical Case Studies*, *18*(6). https://doi.org/10.1177/1534650119872270

Williams, D. E., Kirkpatrick-Sanchez, S., Enzinna, C., Dunn, J., & Borden-Karasack, D. (2009). The clinical management and prevention of pica: A retrospective follow-up of 41 individuals with intellectual disabilities and pica. *Journal of Applied Research in Intellectual Disabilities*, *22*(2), 210-215. https://doi.org/10.1111/j.1468-3148.2008.00490.x

Woods, T. S. (1982). Reducing severe aggressive and self-injurious behaviour: A nonintrusive, home based approach. *Behavioral Disorder*, *7*(3), 180-188. https://doi.org/10.1177/019874298200400311

Table 1

Participant Demographics

Application	Name	Sex	Age	Diagnoses	Level of ID	Target Behaviours	Hypothesized Behaviour Function
1	Lily	F	28	ASD, GAD, ODD, DBD-NOS, FD	Mild	NC, VA, VD, ED, SIB, elopement, flopping, PA, ingesting foreign objects	Attention, escape, access to tangibles
2	Kevin	M	29	RMD, P-NOS, ADHD, SB, Type 2 Diabetes	Mild	RAF, NC, agitation, ED, SIB, PA, suicidal ideologies, eloping, social fantasizing/play	Escape, attention, access to preferred items and activities
3	Ginny	F	23	GDD	-	NC, VA, ED, elopement, PA, fixation on items, SIB	Automatic
4	Riley	M	20	ASD	Moderate	NC, perseverating and whining, VA, elopement, SIB, PA, ED	Escape, access to negative social attention and tangibles
5	Micaela	F	26	BPD, DD	-	NC, VA, VD, PA, ED, SIB, elopement, requesting emergency services	Access to tangibles, attention, escape
6	Anthony	M	21	ARND, ADHD, ADHD, SPD, AD-NOS, DD-NOS, DCD	-	NC, VA, VC, MB, ISB, SIB, ED, PA, elopement, inappropriate internet/electronic usage	Escape, Access to Preferred Social Interactions and tangibles
7	Taylor	F	28	ASD, BPD, Schizophrenia, DD	-	NC, VA, AB, perseveration, ED, SIB, SB, elopement	Access to tangibles, attention, escape
8	Oliver	M	24	Unspecified ICD, MDD-SEM, AD-NOS, CD-CO, ADHD, FASD	-	NC, VC, VA, intrusive thoughts and fixations, inappropriate social behaviour, deceptive statements, PA, ED, elopement,	Escape, attention, access to tangibles

Application	Name	Sex	Age	Diagnoses	Level of ID	Target Behaviours	Hypothesized Behaviour Function
						compulsive spending, food stealing	

Note. ID = intellectual disability, ASD = autism spectrum disorder, GAD = generalized anxiety disorder, ODD = oppositional defiant disorder, DBD-NOS = disruptive behaviour disorder – not otherwise specified, FD = factitious disorder, RMD = reactive mood disorder, P-NOS = paraphilia – not otherwise specified, ADHD = attention deficit hyperactivity disorder, SB = spina bifida, GDD = global developmental disability, BPD = borderline personality disorder, DD = developmental disability, ARND = alcohol related neurodevelopmental disorder, SPD = sensory processing disorder, AD-NOS = anxiety disorder – not otherwise specified, DD-NOS = depressive disorder – not otherwise specified, ICD = impulse control disorder, MDD-SEM = major depressive disorder – single episode and moderate, CD-CO = conduct disorder – childhood-onset, FASD = fetal alcohol spectrum disorder. NC = noncompliance, VA = verbal aggression, VD = verbal disruption, ED = environmental destruction, SIB = self-injurious behaviour, PA = physical aggression, RAF = requesting, arguing, and fabricating, VC = verbal complaining, MB = mischief behaviour, ISB = inappropriate and sexualized behaviours, AB = antagonizing behaviour, SB = sexualized behaviour. Dashed cells indicate that data were not obtained or reported.

Table 2

Sample Negative Target Behaviours and Operational Definitions

Target behaviour	Operational definition
Verbal aggression	Any instance of yelling, swearing, insulting or criticizing staff, threatening, accusing others, and loud complaining (i.e., speaking above conversational level). This does not include crying or complaining at a conversational or quiet level unless swearing or threatening/attacking staff is observed.
Physical aggression	Any attempted or accomplished instance of biting, pinching, scratching, flailing at others, punching, hitting, kicking, hair pulling, spitting directly at/onto another person, and throwing items at others.
Ingesting foreign objects	Any witnessed or disclosed instance of intentionally seeking out and swallowing non-food items for self-harm, including unsuccessful attempts. Items may include, but are not limited to bobby pins, nails, screws, hooks, light bulb pieces, pop cans, license plates, mailboxes, eaves troughs, batteries, coil springs, razor blades, toilet bowl cleaning, and other hazardous liquids.
Self-injurious behaviour	Any attempted or accomplished instance of causing harm to self. This may include, but is not limited to biting, head-banging on hard surfaces, body-slamming, wrist-cutting with sharp and broken items, and self-asphyxiation.
Environmental destruction	Any attempted or accomplished instance of destroying items in the environment. This includes punching/kicking holes in walls, destroying furniture, and picking at baseboards, electric sockets, vents, etc.
Elopement	Any instance of leaving designated areas without permission, accompaniment, or staff knowledge.
Requesting, arguing, and fabricating (RAF)	Any instance in which: a) a request is made over and above what the current behaviour support plan/daily activity schedule permits (i.e., wanting to cook during skills interval as opposed to prescribed cooking interval, wanting an extra cigarette, asking staff to bring coffee from Tim Horton's, etc.); b) arguing with staff when told 'no' or when advised to discuss it during 'daily concerns' time; and (c) saying that a BT or other person approved the request earlier (at a different time) when they did not.
Noncompliance	Any instance of not initiating a task within 30 seconds of instruction, which may be accompanied by vocalizations of refusal.

Table 3

Sample Adaptive Target Behaviours and Response Requirements

Target behaviour	Correct response	Incorrect response
Completes single chores independently	A correct response is defined as completing a step as outlined (e.g., open lid of the garbage, remove full garbage bag, hold both sides of the garbage bag, cross both sides of the garbage bag, loop the ends of the bag threading one side of the garbage bag through the hole, etc.) independently and without any prompts.	An incorrect response is defined as requiring more than one gestural prompt or a higher level of prompting (i.e., verbal, partial, or full physical) for a step.
Reading	A correct response is defined as reading a target word (e.g., chamber, nobleman, utterly, etc.) accurately, independently, and without any prompts.	An incorrect response is defined as reading a target word inaccurately or not knowing how to say a target word.
Spelling	A correct response is defined as spelling a target word (e.g., search, check, quick, etc.) accurately and within five seconds of the instruction being delivered (i.e., "Spell ___).	An incorrect response is defined as spelling a target word inaccurately or after five seconds of the instruction being delivered.
Maintains a conversation	A correct response is defined as responding within 10 seconds and independently meeting each of the following criteria: (1) comments on the staff's questions; (2) uses a filler word; and (3) asks staff a WH-question specific to the topic being discussed.	An incorrect response is defined as not responding within 10 seconds or erroring on one or more of the following criteria: (1) comments on the staff's questions; (2) uses a filler word; and (3) asks staff a WH-question specific to the topic being discussed.
Performs basic Microsoft Word functions	A correct response is defined as accurately and independently performing the action (e.g., creating a bulleted, numbered, or lettered list).	An incorrect response is defined as requiring a reminder on how to perform the action.
Yoga	A correct response is defined as watching the YouTube video and trying to complete the moves for the entire duration of the interval.	An incorrect response is defined as not trying to watch the YouTube video and not trying to complete the moves for the entire duration of the interval.

Table 4

Social Validity Statements, Domains, and Formats for Clients

Statements	Domain	Format
S1: The services at SMG are the first behavioural services I have ever received.	Before services	Y/N
S2: Other services (like behavioural interventions) before coming to SMG have not been helpful.	Before services	Y/N
S3: Other services (like behavioural interventions) before SMG were helpful, but did not reduce my challenging behaviour. For example, breaking things, physically hurting other people, swearing at people, getting arrested.	Before services	Likert scale
S4: Other services like behavioural interventions) before SMG were helpful in reducing my challenging behaviour, but it came back.	Before services	Likert scale
S5: Before SMG, it was important to me to stop engaging in challenging behaviour.	Before services	Likert scale
S6: Before SMG, it was important to me to learn new things and learn new skills.	Before services	Likert scale
S7: The services at SMG are necessary to make sure I do not get hurt and do not hurt others.	During services	Likert scale
S8: All of the items in my behaviour support plan are helpful.	During services	Likert scale
S9: My behaviour support plan is helping me decrease how often I try to hurt myself or others. That is, engage in challenging behaviour.	During services	Likert scale
S10: My skill building program is helping me learn new skills.	During services	Likert scale
S11: SMG has helped me progress and work towards a better life. I feel better now.	During services	Likert scale
S12: I feel like I am getting the best service here.	Final thoughts	Y/N
S13: I feel supported here.	Final thoughts	Y/N
S14: I feel safe where I am currently living.	Final thoughts	Y/N
S15: Did the program help you?	Final thoughts	Y/N
S16: Can you list one (or more) things that you feel have helped you the most?	Final thoughts	Open-ended
S17: If you could change one thing about the services you are getting at SMG, what would it be?	Final thoughts	Open-ended

Note. Y/N = Yes or No.

Table 5

Social Validity Statements, Domains, and Formats for Caregivers and Case Managers

Statements	Domain	Format
S1: The current treatment was the first treatment my client (my loved one) has ever received.	Before services	Y/N
S2: In my opinion, other behavioural services my client (my loved one) experienced were not really helpful.	Before services	Y/N
S3: In my opinion, previous treatments were helpful but did not address my client's (my loved one's) challenging behaviour.	Before services	Likert scale
S4: In my opinion, previous treatments successfully addressed my client's (my loved one's) challenging behaviour, but the challenging behaviour came back.	Before services	Likert scale
S5: Before starting the current treatment, it was important to me to decrease my client's (my loved one's) challenging behaviour.	Before services	Likert scale
S6: Before starting the current treatment, it was important to me to increase my client's (my loved one's) adaptive skills.	Before services	Likert scale
S7: The current treatment is necessary to ensure my client's (my loved one's) safety.	During services	Likert scale
S8: The current treatment is acceptable for behaviour support.	During services	Likert scale
S9: The current treatment seems to be effective in decreasing my client's (my loved one's) challenging behaviour.	During services	Likert scale
S10: The current treatment seems to be effective in increasing my client's (my loved one's) adaptive skills.	During services	Likert scale
S11: The current treatment makes it possible for my client (my loved one) to progress and achieve a better quality of life.	During services	Likert scale
S12: I feel like my client (my loved one) is getting the best service here.	Final thoughts	Y/N
S13: I feel like my client (my loved one) is well supported in their current placement.	Final thoughts	Y/N
S14: I feel like my client (my loved one) is safe where they are currently living.	Final thoughts	Y/N
S15: Did the treatment help your client (your loved one)?	Final thoughts	Y/N
S16: Can you list one (or more) things that you feel have helped your client (your loved one) the most?	Final thoughts	Open-ended
S17: If you could change one thing about the supports that your client (your loved one) is receiving, what would it be?	Final thoughts	Open-ended

Note. Y/N = Yes or No.

Table 6

Social Validity Results for Clients

Statement	Strongly agree		Agree		Disagree		Strongly disagree		Prefer not to answer		Yes		No		Open-ended	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
S1											4	80	1	20		
S2											1	20	4	80		
S3	2	40	2	40	0	0	1	20	0	0						
S4	1	20	0	0	1	20	3	60	0	0						
S5	2	40	0	0	1	20	2	40	0	0						
S6	1	20	2	40	1	20	1	20	0	0						
S7	2	40	1	20	1	20	1	20	0	0						
S8	2	40	1	20	2	40	0	0	0	0						
S9	2	40	2	40	1	20	0	0	0	0						
S10	3	60	0	0	2	40	0	0	0	0						
S11	2	40	3	60	0	0	0	0	0	0						
S12											3	60	2	40		
S13											3	60	2	40		
S14											3	60	2	40		
S15											4	80	1	20		
S16															4	80
S17															5	100

Note. $N = 5$. Blank cells indicate that data are not applicable. See Table 4 for a full description.

Table 7

Social Validity Results for Caregivers and Case Managers

Statement	Strongly agree		Agree		Disagree		Strongly disagree		Prefer not to answer		Yes		No		Open-ended	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
S1											3	43	4	57		
S2											6	86	1	14		
S3	2	29	4	57	1	14	0	0	0	0						
S4	0	0	5	71	2	29	0	0	0	0						
S5	7	100	0	0	0	0	0	0	0	0						
S6	7	100	0	0	0	0	0	0	0	0						
S7	3	43	4	57	0	0	0	0	0	0						
S8	4	57	3	43	0	0	0	0	0	0						
S9	3	43	4	57	0	0	0	0	0	0						
S10	1	14	6	86	0	0	0	0	0	0						
S11	4	57	3	43	0	0	0	0	0	0						
S12											6	86	1	14		
S13											7	100	0	0		
S14											7	100	0	0		
S15											7	100	0	0		
S16															5	71
S17															4	57

Note. $N = 7$. Blank cells indicate that data are not applicable. See Table 5 for a full description.

Figure 1

Modified Brinley Plot for Challenging Behaviour across Baseline and Intervention
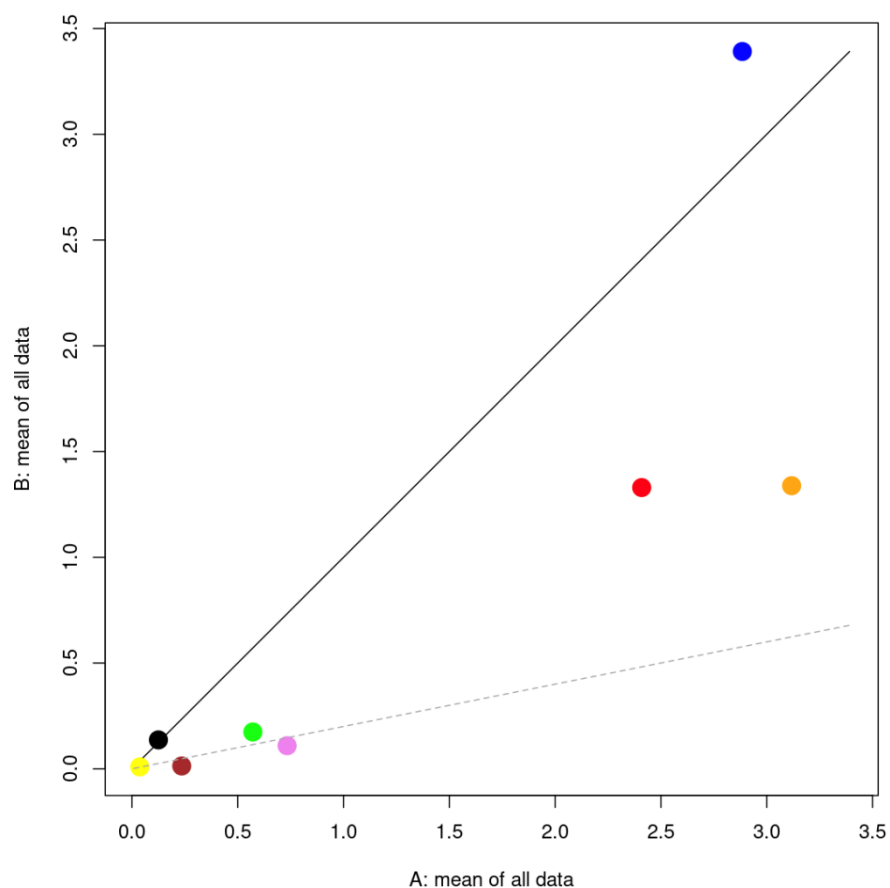
Note. This figure depicts challenging behaviour data across baseline and intervention for all participants on a modified Brinley plot. The x-axis represents the baseline mean and the y-axis represents the intervention mean. The black solid diagonal line signifies the equivalence between baseline and intervention and the grey dashed diagonal line indicates the desired magnitude of change from the baseline level. The data points are designated as follows: red = Lily, green = Kevin, blue = Ginny, orange = Riley, pink = Micaela, brown = Anthony, black = Taylor, and yellow = Oliver.
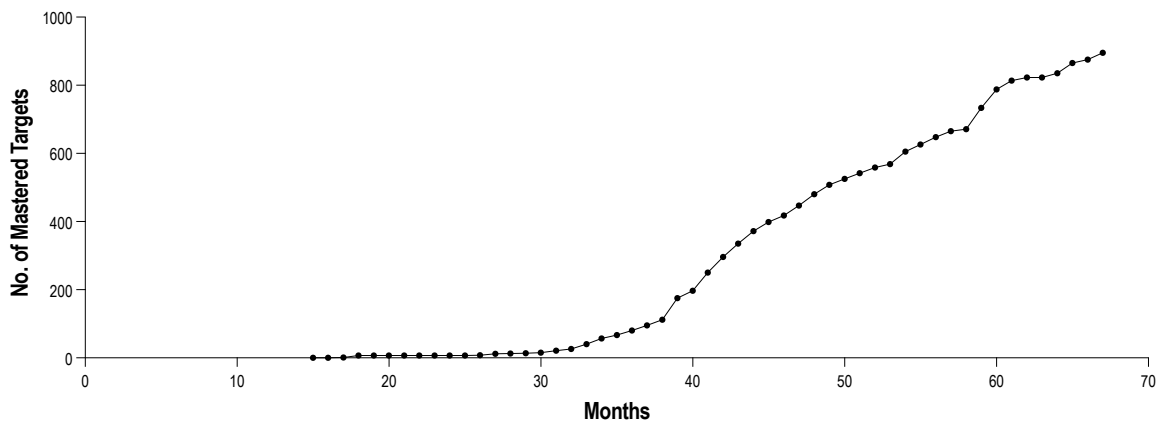
Figure 2

Intervention Data for Lily's Adaptive Behaviour across Months
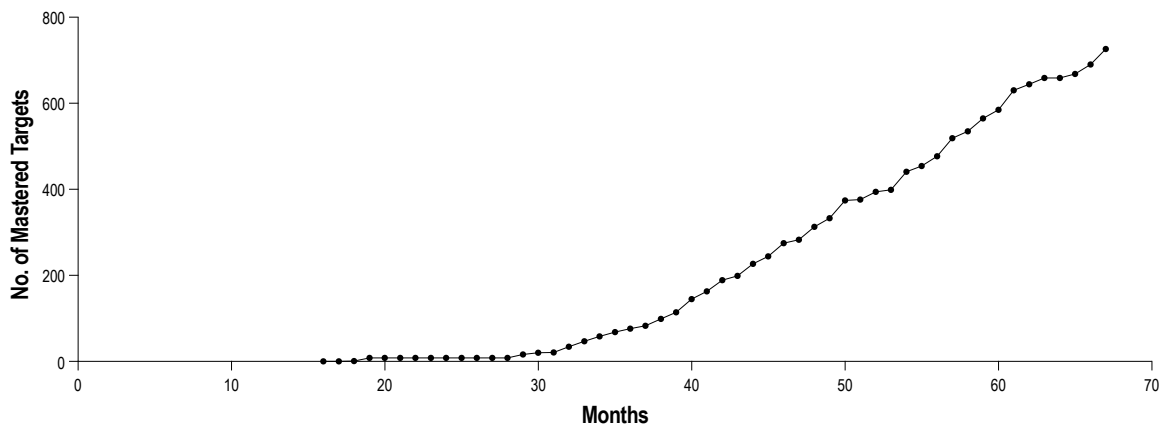
Note. This graph depicts intervention data for Lily's adaptive behaviour. The y-axis represents

the number of mastered skills targets and the x-axis represents months in treatment. Data

collection for adaptive behaviour commenced at 16 months after admission.
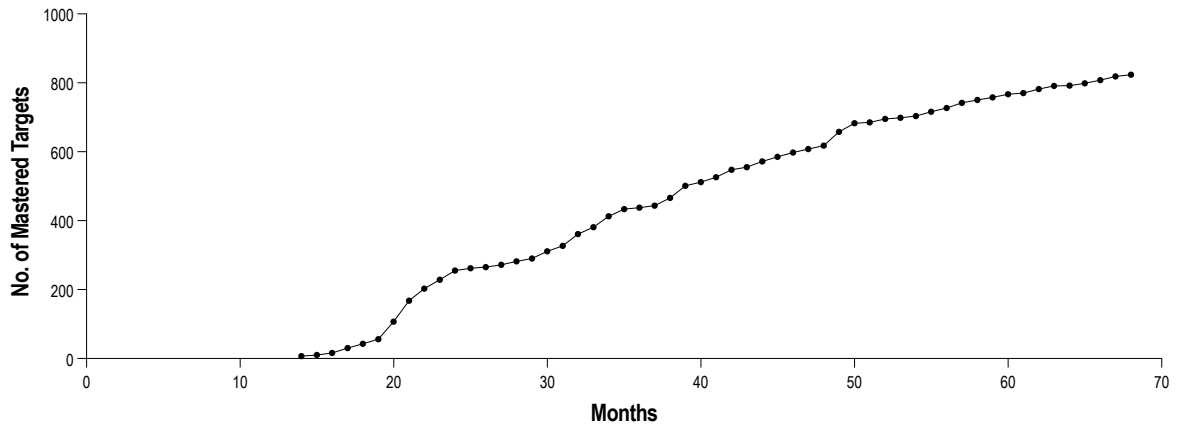
Figure 3

Intervention Data for Kevin's Adaptive Behaviour across Months

Note. This graph depicts intervention data for Kevin's adaptive behaviour. The y-axis represents the number of mastered skills targets and the x-axis represents months in treatment. Data collection for adaptive behaviour commenced at 15 months after admission.
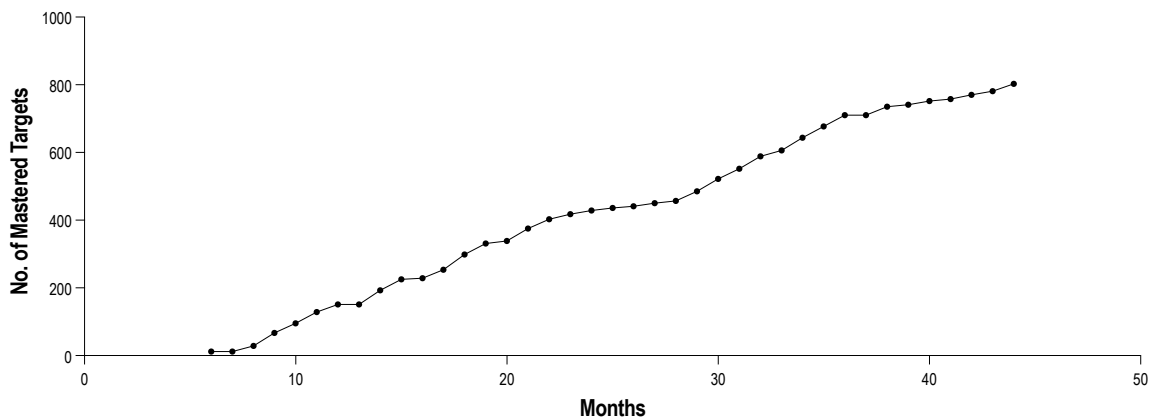
Figure 4

Intervention Data for Ginny's Adaptive Behaviour across Months

Note. This graph depicts intervention data for Ginny's adaptive behaviour. The y-axis represents the number of mastered skills targets and the x-axis represents months in treatment. Data collection for adaptive behaviour commenced at 14 months after admission.
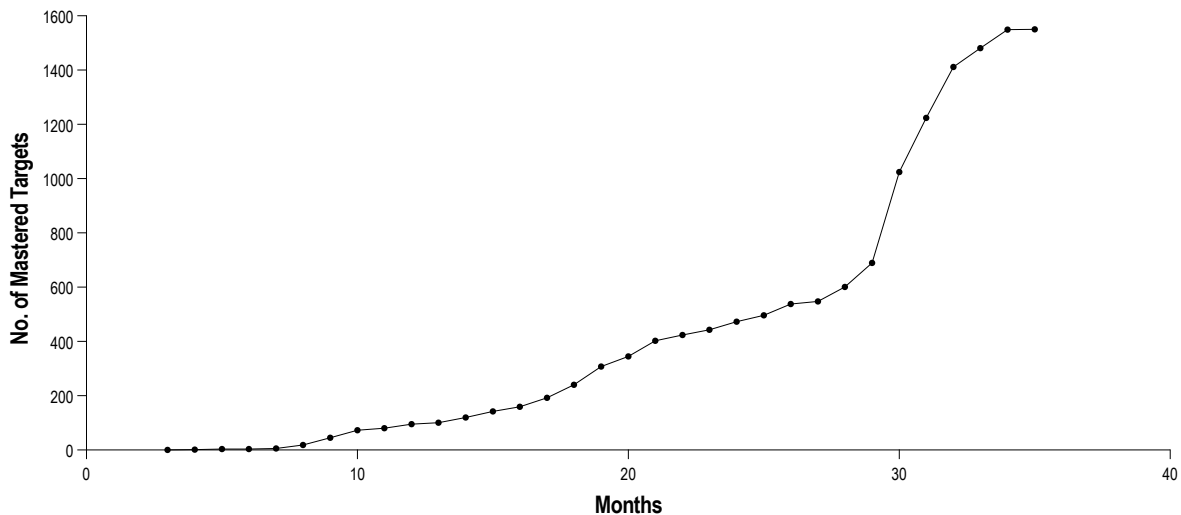
Figure 5

Intervention Data for Riley's Adaptive Behaviour across Months

Note. This graph depicts intervention data for Riley's adaptive behaviour. The y-axis represents the number of mastered skills targets and the x-axis represents months in treatment. Data collection for adaptive behaviour commenced at 6 months after admission.
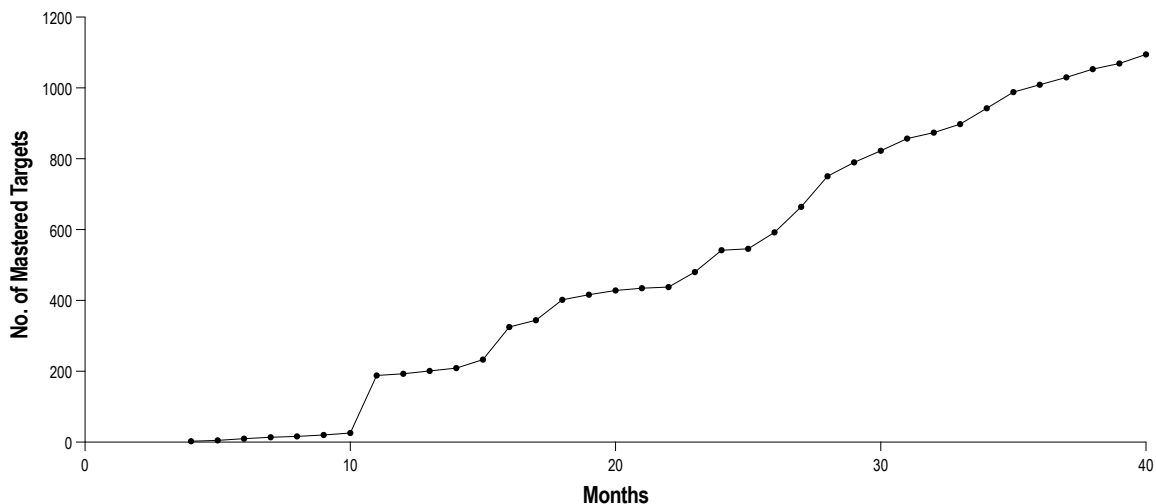
Figure 6

Intervention Data for Micaela's Adaptive Behaviour across Months

Note. This graph depicts intervention data for Micaela's adaptive behaviour. The y-axis

represents the number of mastered skills targets and the x-axis represents months in treatment.

Data collection for adaptive behaviour commenced at 3 months after admission.
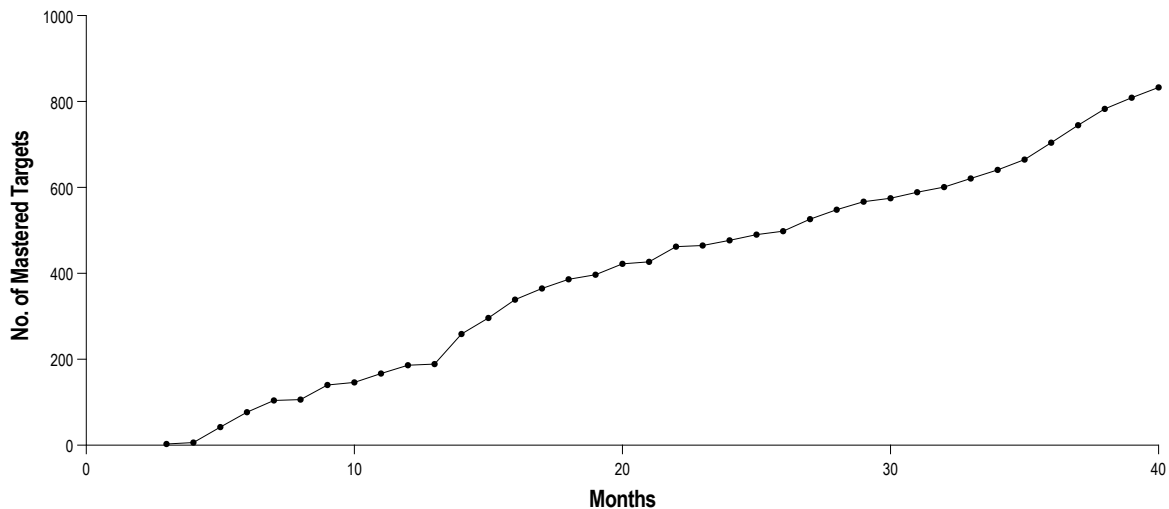
Figure 7

Intervention Data for Anthony's Adaptive Behaviour across Months

Note. This graph depicts intervention data for Anthony's adaptive behaviour. The y-axis

represents the number of mastered skills targets and the x-axis represents months in treatment.

Data collection for adaptive behaviour commenced at 4 months after admission.

Figure 8

Intervention Data for Taylor's Adaptive Behaviour across Months

Note. This graph depicts intervention data for Taylor's adaptive behaviour. The y-axis represents the number of mastered skills targets and the x-axis represents months in treatment. Data collection for adaptive behaviour commenced at 4 months after admission.
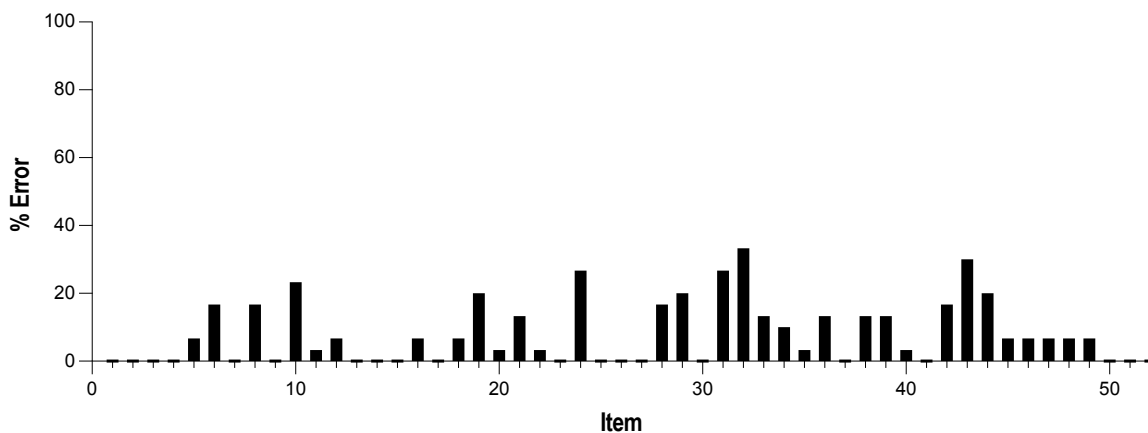
Figure 9

Percentage of Errors across Items for Lily's Treatment Integrity Checklist

Note. This graph depicts the results of the item-by-item treatment integrity checklist (TIC) error analysis for Lily. The x-axis represents the number of items on the TIC and the y-axis represents the percentage of error. Lily's TIC consisted of five domains: preventative strategies (items 1–12), management of behaviour (items 13–41), overnight protocol (items 42–45), bathroom protocol (items 46–49), and data collection (items 50–52).

Figure 10

Percentage of Errors across Items for Kevin's Treatment Integrity Checklist

Note. This graph depicts the results of the item-by-item treatment integrity checklist (TIC) error analysis for Kevin. The x-axis represents the number of items on the TIC and the y-axis represents the percentage of error. Kevin's TIC consisted of four domains: preventative strategies (items 1–11), reinforcer guidelines and rewards (items 12–25), management of behaviour (items 26–47), and data collection (items 48–50).
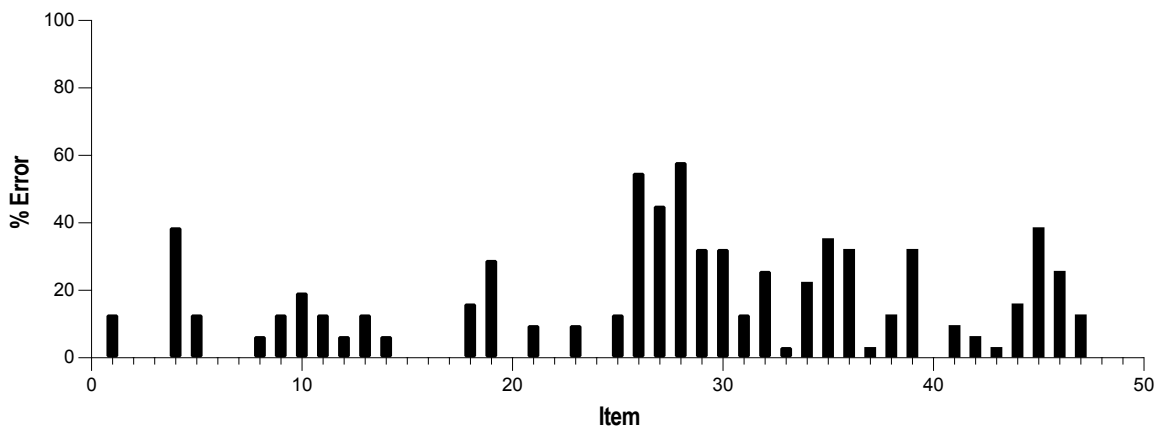
Figure 11

Percentage of Errors across Items for Ginny's Treatment Integrity Checklist

Note. This graph depicts the results of the item-by-item treatment integrity checklist (TIC) error analysis for Ginny. The x-axis represents the number of items on the TIC and the y-axis represents the percentage of error. Ginny's TIC consisted of 12 domains: preventative strategies (items 1–12), teaching task sequence (items 13–17), educative routines (items 18–20), verbal behaviour (items 21–25), management of behaviour (items 26–38), management of behaviour during walks (items 39–40), bathroom protocol (items 41–48), bathroom requests (item 49), meal protocol (items 50–52), urine accident after mechanical chair release (items 53–55), back-up staff (items 56–62), and data collection (items 63–65).
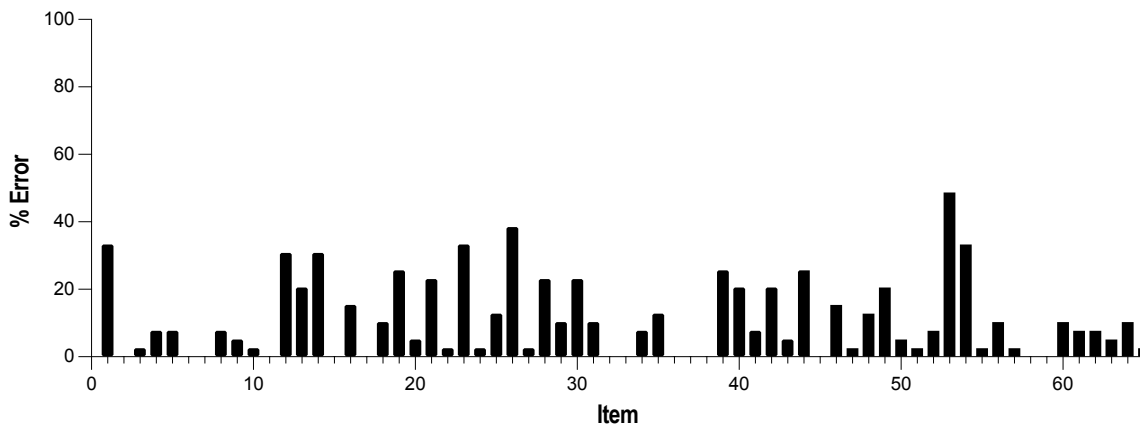
Figure 12

Percentage of Errors across Items for Riley's Treatment Integrity Checklist

Note. This graph depicts the results of the item-by-item treatment integrity checklist (TIC) error analysis for Riley. The x-axis represents the number of items on the TIC and the y-axis represents the percentage of error. Riley's TIC consisted of seven domains: preventative strategies (items 1–23), daily activity schedule (items 24–44), management of behaviour (items 45–67), bathroom (items 68–75), evening/bedtime (items 71–82), back-up staff (items 83–89), and data collection (items 90–92).
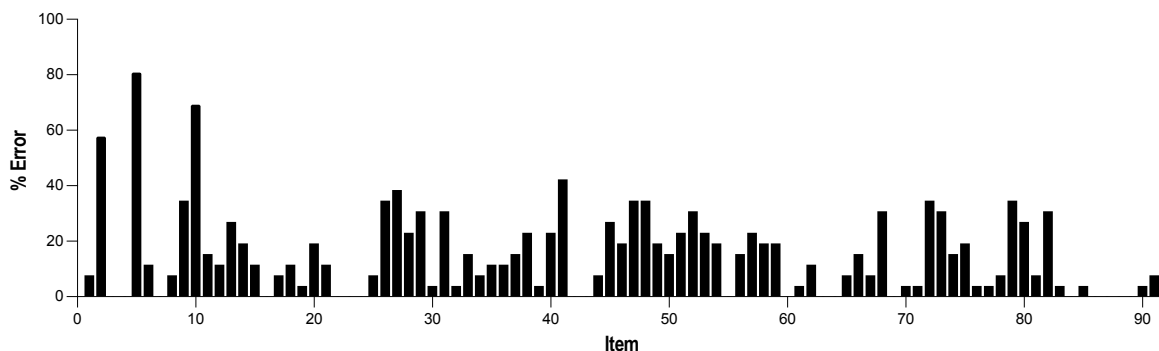
Figure 13

Percentage of Errors across Items for Micaela's Treatment Integrity Checklist

Note. This graph depicts the results of the item-by-item treatment integrity checklist (TIC) error

analysis for Micaela. The x-axis represents the number of items on the TIC and the y-axis

represents the percentage of error. Micaela's TIC consisted of four domains: preventative

strategies (items 1–13), reinforcer guidelines and rewards (items 14–28), management of

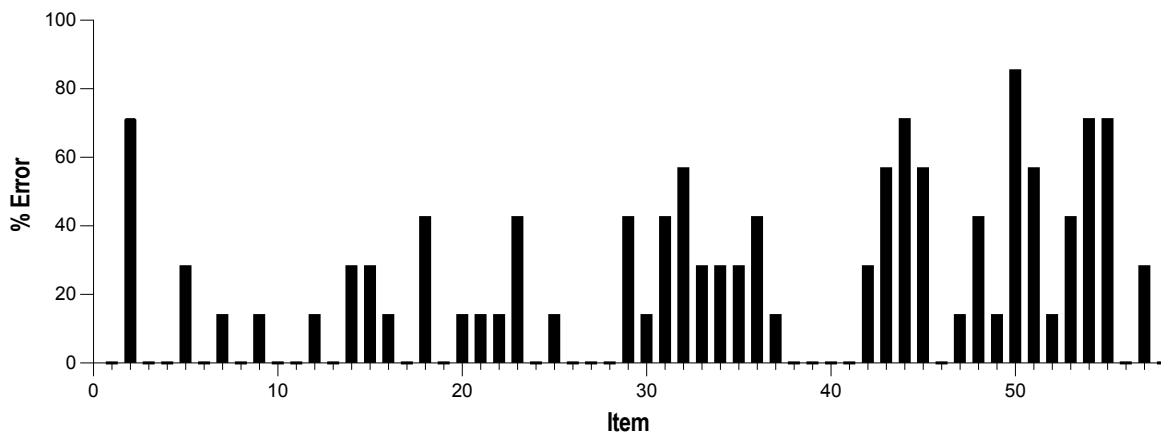behaviour (items 29–55), and data collection (items 56–58).

Figure 14

Percentage of Errors across Items for Anthony's Treatment Integrity Checklist

Note. This graph depicts the results of the item-by-item treatment integrity checklist (TIC) error analysis for Anthony. The x-axis represents the number of items on the TIC and the y-axis represents the percentage of error. Anthony's TIC consisted of five domains: preventative strategies (items 1–16), reinforcer guidelines and rewards (items 17–43), management of behaviour (items 44–64), bathroom (items 65–68), and data collection (items 69–71).

Figure 15

Percentage of Errors across Items for Taylor's Treatment Integrity Checklist

Note. This graph depicts the results of the item-by-item treatment integrity checklist (TIC) error analysis for Taylor. The x-axis represents the number of items on the TIC and the y-axis represents the percentage of error. Taylor's TIC consisted of four domains: preventative strategies (items 1–18), reinforcer guidelines and rewards (items 19–36), management of behaviour (items 37–56), and data collection (items 57–59).
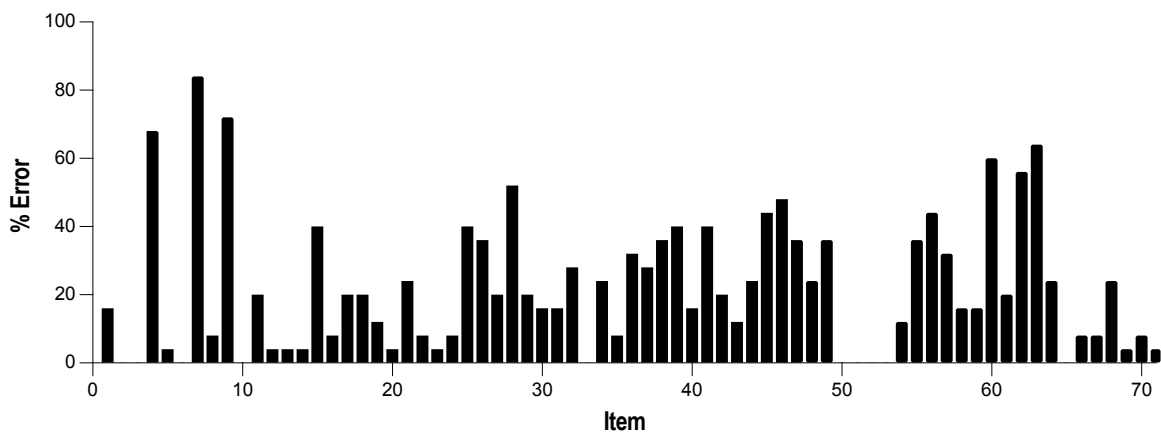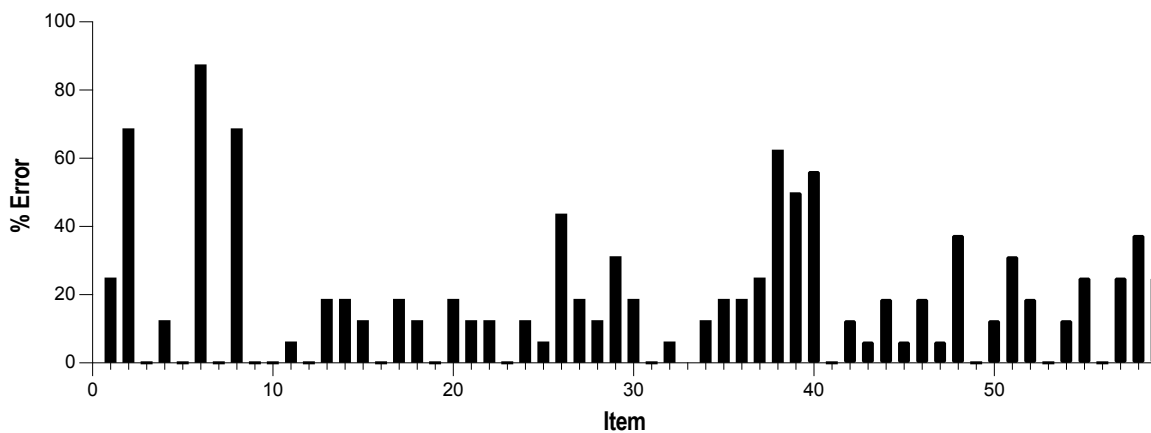
Figure 16

Percentage of Errors across Items for Oliver's Treatment Integrity Checklist
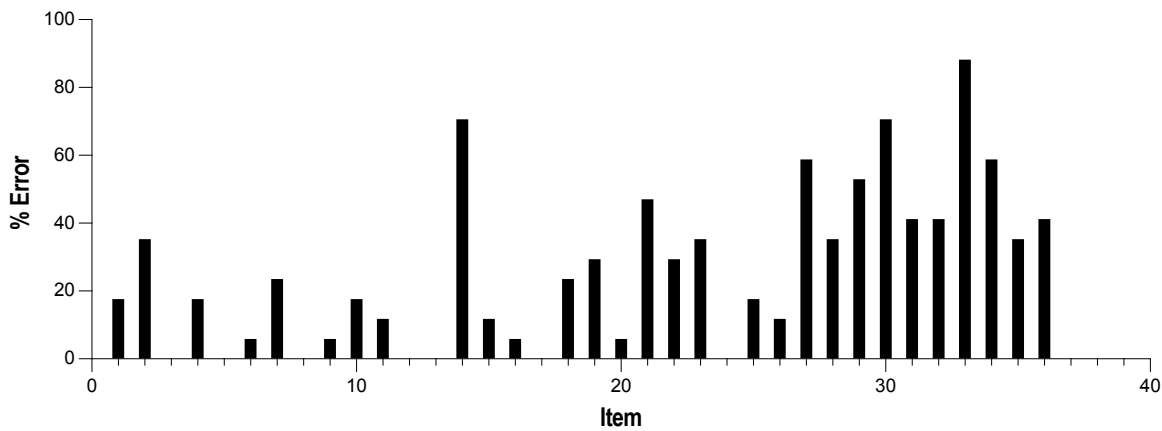
Note. This graph depicts the results of the item-by-item treatment integrity checklist (TIC) error analysis for Oliver. The x-axis represents the number of items on the TIC and the y-axis represents the percentage of error. Oliver's TIC consisted of six domains: target behaviours and operational definitions (items 1–2), preventative strategies (items 3–14), reinforcer guidelines and rewards (items 15–18), access protocols (items 19–25), management of behaviour (items 26–36), and data collection (items 37–40).

Appendix A

Negative Target Behaviours and Operational Definitions

Table A1

Negative Target Behaviours and Operational Definitions for Lily

Target behaviour	Operational definition
Noncompliance	Any instance of refusing to complete requests by verbally responding "no", "leave me alone", by ignoring requests, or by responding to direct with escalating behaviour.
Verbal aggression	Any instance of yelling, swearing, insulting or criticizing staff, threatening, accusing others, and loud complaining (speaking above a conversational level). This does not include crying or complaining at a conversational or quiet level unless Lily is swearing or threatening/attacking staff.
Verbal disruption	Any instance of non-contextual statements or comments. Examples of such comments are, "I was put in jail", "There's a dead body in the bathtub", "I was put in the dryer". "The mop water is dirty and smells like a dead body", "I wipe chocolate on my pants" (when she is not eating chocolate) etc. Lily will make these statements using a loud, aggressive, conversational, or audible whispering tone. If the statement/comment is inappropriate and is not relevant to what she is currently completing, this will be considered 1 frequency of verbal disruption. If she makes appropriate statements that are contextual and relate to the current task, then this will not be considered a verbal disruption. For example, "The bathtub is filthy" (while she is cleaning the bathtub), or "The mop water is dirty" (while she's handling the mop bucket), etc. *Please note: If staff cannot clearly hear what Lily has said standing or sitting 5ft (arm's length) away from her, this will not be considered a VD as it is not clearly audible to the listener.*
Environmental destruction	Any attempt or actual instance of destroying items in her environment. This includes punching/kicking holes in walls, destroying furniture, and picking at baseboards, electric sockets, vents, etc.
Self-injurious behaviour	Any attempt or accomplished instance of causing harm to self. This may include but is not limited to biting, head-banging on hard surfaces, body-slamming, wrist-cutting with sharp and broken items, and self-asphyxiation.
Elopement	Any instance of leaving designated areas without permission, accompaniment, or staff knowledge.
Flopping	Any instance of dropping to the ground and becoming limp to resist full physical prompting or compliance with a demand.
Physical aggression	Any attempt or accomplished instance of pushing, grabbing, scratching, hair-pulling, kicking, and punching others. Biting attempts

Target Behaviour	Operational Definition
	may occur if/when she is prevented/unable to grab, scratch, pull hair, kick, or punch others.
Ingesting foreign objects	Any witnessed or disclosed instances of intentionally seeking out and swallowing non-food items for self-harm, including unsuccessful attempts. Items may include but are not limited to bobby pins, nails, screws, hooks, lightbulb pieces, pop cans, license plates, mailboxes, eaves troughs, batteries, coil springs, razor blades, toilet bowl cleaning, and other hazardous liquids.

Table A2

Negative Target Behaviours and Operational Definitions for Kevin

Target behaviour	Operational definition
Requesting, arguing, and fabricating (RAF)	Any instance in which: a) a request is made over and above what his current behaviour support plan/daily activity schedule permits (i.e., wanting to cook during skills interval as opposed to prescribed cooking interval, wanting an extra cigarette, asking staff to bring coffee from Tim Horton's, etc.); b) arguing with staff when told 'no' or when advised to discuss it during 'daily concerns' time; and (c) saying that a BT or other person approved the request earlier (at a different time) when they did not.
Noncompliance	Any instance of not following direction from staff such as not washing his hands before using the catheter. He will often flatly refuse to follow through with expectations or state that it is his decision to make.
Agitation	Any instance of verbal aggression in the form of swearing, raising his voice, making derogatory comments, challenging staff, and invading the personal space of others.
Threats towards others	Any instance of verbal or gestural threats towards others.
Property damage	Any instance of ripping, breaking, defacing, kicking, or throwing items (not in the direction of others).
Self-injurious behaviour	Any attempt or accomplished instance of engaging in self-harm behaviour. This may include but is not limited to head-banging, self-cutting, self-striking, ingestion, inserting items into his belly button, picking wounds, and self-biting.
Physical aggression	Any attempt or accomplished instance of Kevin hitting, kicking, or throwing items in the direction of others or actin gout threats towards others.
Suicidal ideologies	Any instance of talking about committing suicide or wanting to call the hospital because of these thoughts.
Eloping	Any instance of leaving the home and/or staff's presence in the community without staff knowledge or approval. This has occurred most often with new staff.
Social fantasizing/play	Any instance of Kevin verbally sharing future plans and goals involving achievements that are unattainable or unrealistic. These can include topics like big purchases (e.g., trailer, motorcycle) or unrealistic or dangerous activities (e.g., skating lessons/becoming an Olympic skater, biking across Canada).

Table A3

Negative Target Behaviours and Operational Definitions for Ginny

Target behaviour	Operational definition
Noncompliance	Any instance of Ginny refusing to follow her daily schedule. If a verbal instruction is delivered and Ginny does not respond, staff will implement the prompt hierarchy. If staff have to implement a full physical prompt, this will be considered one instance of noncompliant behaviour. If Ginny states "no" after a verbal instruction, walks away, or drops to the ground after a verbal instruction has been delivered, this will be considered one instance of noncompliant behaviour. Please note the following steps within the prompt hierarchy will not be considered noncompliant behaviour: high probability response, re-delivering verbal instruction, gestural prompt, and partial prompt.
Verbal aggression	Any instance of Ginny yelling, swearing, and screaming "ouch" or "ow". A new instance of verbal aggression is recorded any time there is a 2-3 s break between any of the behaviours listed above.
Environmental destruction	Any attempt or accomplished instance of throwing items, flipping tables, tossing chairs, and punching objects in the environment, such as doors, mirrors, or windows. This does not include slapping the table or counter.
Elopement	Any attempt to leave the home or from staff if she sees something she desires. She will also run across roads.
Physical aggression	Any attempt (e.g., finger flicking) or accomplished instance of Ginny biting, pinching, scratching, flailing at others, punching, hitting, kicking, hair pulling, spitting directly at/onto another person, and throwing items at others.
Fixation on items	Any instance of repetitive requesting an item that she desires thereby requiring staff to utilize gestural, partial, and/or full physical prompting to redirect her back to the task at hand. This means that each time staff must make their way through the prompt hierarchy, they will record one instance of fixation. Fixations arise to gain access to items that are currently not obtained/granted access to.
Self-injurious behaviour	Any attempt or accomplished instance of causing harm to self. This may include but is not limited to biting, head-banging on hard surface, hitting any part of her body with an open or closed fist.

Table A4

Negative Target Behaviours and Operational Definitions for Riley

Target behaviour	Operational definition
Noncompliance	Any instance of not initiating a task within 30 s of instructions/answering a question, which may be accompanied by vocalizations of refusal, stating the task is too hard/he cannot do it, and may also include dropping/flopping to the floor.
Perseverating and whining	Any instance of fixating over a topic, continuously asking questions about it (i.e., why it has not happened or when it will happen), and refusing to focus on the task at hand. His thoughts, conversations, questions, and verbal language will be all about the topic he is fixated on.
Verbal aggression	Any instance of yelling, screaming, swearing, or threatening others. Examples include: "I want to call the police" "Stop talking", "Don't do that", "I want to go to another place", etc. This also includes all instances of verbally antagonizing, taunting, or purposefully making malicious sounds/statements to provoke others in the environment.
Elopement	Any instance of leaving the environment without staff's knowledge or permission.
Self-injurious behaviour	Any instance of hitting a body part with an open or closed fist or biting any part of the body.
Physical aggression	Any attempt or accomplished instance of kicking, punching, hitting, pushing/charging, spitting, biting, throwing objects at others, or using an object to hit someone. This may or may not be done with intense force.
Environmental destruction	Any attempt or accomplished instance of throwing, ripping, or breaking items in the environment using objects or any part of his body. This has included breaking electronics, damaging walls or windows, slamming doors or other objects that open and close, and flipping furniture. This also includes touching his helmet or the straps of the mechanical chair.

Table A5

Negative Target Behaviours and Operational Definitions for Micaela

Target behaviour	Operational definition
Noncompliance	Any instance of not initiating a task within 30 s of instructions, which may be accompanied by vocalizations of refusal.
Verbal aggression	Any instance of yelling or screaming, swearing at others, name calling, or making verbal threats towards self or others. This includes racial slurs and inappropriate sexual comments directed toward staff and/or other members. One aggressive word/statement directed towards another individual will be counted as one instance of verbal aggression.
Verbal disruption	Any insult, criticism, general threats towards environment, or accusation made towards or about someone/something in the environment. This includes complaints and/or criticisms, such as "this is too hard", "jail was better than this", "you do not know what I have been through", and various other statements made in a loud conversational tone to antagonize others within the environment. This behaviour is often associated with on-task behaviour as Micaela will engage in verbal disruption while completing the task at hand.
Physical aggression	Any attempt or accomplished instance of hitting, kicking, pinching, pushing or punching others, making threatening gestures, as well as pulling hair or clothing, spitting at others, or throwing objects in the direction of another person.
Environmental destruction	Any attempt or accomplished instance of ripping, breaking, destroying, spitting at/towards objects, punching or kicking holes in walls, kicking doors, repeatedly slamming doors, throwing objects such as tables and televisions, or pulling items off the wall.
Self-injurious behaviour	Any attempt or accomplished instance of biting self, head banging, cutting any part of her body with another object. This may include glass, sharp plastic, or any other sharp object.
Elopement	Any instance of leaving the house, designated area of the environment, or community location without staff knowledge or permission.
Requesting Emergency Services	Any request or demand from Micaela to access police or go to the hospital for non-emergency reasons or other reasons that do not exist.

Table A6

Negative Target Behaviours and Operational Definitions for Anthony

Target behaviour	Operational definition
Noncompliance	Any instance of not following through with the daily activity schedule or other demands within 30 s or any verbal refusal, such as "no" when a demand is placed.
Verbal aggression	Any instance of screaming, swearing, name calling, threatening to harm others, and loud/disrupting blubbering.
Verbal complaining and lying	Any instance of arguing about rules and persistent attempts to negotiate after rules/expectations have been outlined. This also includes attempting to or successfully withholding or modifying the truth across individuals and manipulating facts/rules to elicit a desired outcome.
Mischief behaviour	Any instance of taking objects that do not belong to him, hiding objects that do not belong to him, and/or hiding objects that he is not permitted to have on him or in the area.
Inappropriate and sexualized behaviours	Any instance of making sexual references towards items and people in the environment, attempting to or successfully touching, poking, cupping, grabbing, nudging, and groping other people's body parts. In addition, he will make inappropriate gestures and remarks towards another individual that is sexually suggestive.
Inappropriate internet/electronic usage	Any instance of using his electronics to gain access to inappropriate, unapproved, or illegal context. This also includes instances where he hides electronics from staff and/or fails to give them up during designated intervals.
Self-injurious behaviour	Any instance of slapping self, biting self, headbanging against floor, wall, table, chair, and/or another's arms while in containment. The latter will only be defined as self-injury if Anthony continues to hit his head against staff's arms because their arm is positioned to prevent his head from making contact with the surface he is attempting to hit himself on.
Environmental destruction	Any instance of aggressively tearing/ripping up paper, setting fires, or throwing/breaking items in the environment using objects or any part of his body. This includes throwing heavy objects at the walls or vehicles and flipping furniture.
Physical aggression	Any attempt or accomplished instance of intense kicking, punching, hitting, pushing/shoving, biting, spitting, throwing objects at others, or using object as weapons.
Elopement	Any instance of leaving the environment without staff's knowledge or permission. Community elopements involve leaving the designated/approved route to wander around his community, refusing to return to the program when asked to by staff.

Table A7

Negative Target Behaviours and Operational Definitions for Taylor

Target behaviour	Operational definition
Noncompliance	Any instance of not initiating a task within 30 s of instruction which may be accompanied by vocalizations of refusal.
Verbal aggression	Any instance of yelling or screaming, swearing at others, name calling, rude comments, or making verbal threats. Additionally, verbal aggression is defined as using an argumentative or aggressive tone to express negative commentary and/or criticism of someone or something.
Antagonizing behaviour	Any instance of deliberately engaging in actions that cause others to feel upset, hostile, angry, or uncomfortable. For instance, giving people the middle finger, purposefully snapping her fingers in someone's ear, or licking their food.
Perseveration	Any instance of continuous rumination or fixation of a specific topic to obtain a desired response. This means that Taylor will continuously ask questions about the same topic looking for a specific response. Sometimes, if Taylor does not receive the response she is looking for/wants, then she will ask the same question in a different way or to another individual.
Physical aggression	Any attempt or accomplished instance of hitting, kicking, biting, pushing or punching others as well as pulling hair or clothing or throwing objects in the direction of another person.
Environmental destruction	Any attempt or accomplished instance of breaking objects in the environment, punching or kicking holes in walls, slamming doors, throwing objects, or pulling items off the wall.
Self-injurious behaviour	Any attempt or accomplished instance of cutting any part of her body with another object, biting or scratching self, slapping her face, head-banging, and running into traffic.
Sexualized behaviour	Any instance of fixating or repeatedly speaking about sexual themes, touching self and others, and engaging in behaviours of self-stimulation for pleasure.
Elopement	Any instance of flopping, leaving the house, designated area of the environment or community location without staff knowledge or permission.

Table A8

Negative Target Behaviours and Operational Definitions for Oliver

Target behaviour	Operational definition
Noncompliance	Any instance of not following through with the daily activity schedule or other demands within 30 s or any verbal refusal such as "no" when a demand is placed.
Verbal complaining	Any instance of arguing about rules and persistent attempts to negotiate after rules/expectations have been outlined. This must occur for at least 60 consecutive seconds to be identified as verbal complaining.
Verbal statements of intrusive thoughts and fixations	Any instance of Oliver experiencing unwanted thoughts or images that he finds distressing or disturbing or becoming preoccupied with a particular person or object and cannot be redirected away from this after 5 min. This includes thoughts of wanting to harm self or others, sexual obsessions (thoughts related to sexual intercourse, sexual desires, or fantasies), or religious thinking (has negative beliefs regarding his faith or religious beliefs).
Inappropriate social behaviour	Any instance of making inappropriate comments or jokes that are offensive regarding sex, gender, or race.
Deceptive statements	Any attempt or actual instances of Oliver withholding the truth.
Physical aggression	Any attempt or accomplished instance of hitting, kicking biting, pushing, or shoving others. This also includes any attempt or instance of him holding up an item while verbalizing that he is going to use it as a weapon or him throwing an object in the direction of others.
Mild environmental destruction	Any attempt or successful instance of environmental destruction of a mild degree (e.g., ripping a page out of his bible).
Severe environmental destruction	Any attempt or successful instance of a severe degree (e.g., punching a wall, slamming the door, smashing a plate, throwing his chair, flipping the table, etc.).
Compulsive spending	Any attempt or accomplished instance of him using his electronics (e.g., Xbox or personal computer) to purchase several preferred items (e.g., video games or movies) in excess (e.g., spending over $1000) in a short period of time and without the clinical team's approval.
Food stealing	Any attempt or accomplished instance of Oliver taking more food than he is permitted to eat as per his mealtime guidelines and doing so while staff attention is diverted from him.

Appendix B

Adaptive Behaviour Domains and Programs

Participant	Domain	Program
Lily	Activities of daily living	Grocery shopping, first aid, public transportation
	Academic	Reading comprehension, time management, learns new information
	Vocational	Learns to type, basic computer skills, Introduction to PowerPoint
	Replacement skills	Labels and expresses emotions
	Social	Maintains a conversation
Kevin	Activities of daily living	Engages in physical activity, public transportation, collaborative problem solving
	Academic	Learns new information, learns grammar
	Learning skills	Independent study unit
	Expressive	Uses Duolingo to learn Spanish
	Social	Plans a social event
Ginny	Activities of daily living	Toileting, completes chores independently, personal care, makes a salad, unloads dishwasher
	Recreational/leisure	Community outings, engages appropriately during leisure
	Replacement skills	Tolerates delayed/denied access
	Spontaneous language	Mands for items and actions
	Vocational	Independent task
	Gross motor	Physical activity
	Expressive	Articulates speech sounds
	Community	Labels community locations
Riley	Activities of daily living	Daily activity schedule, completes grocery inventory, first aid

Participant	Domain	Program
	Academic	Learns new information, introduction to emergency vehicles
	Math	Math concepts (grade two)
	Writing	Writes appropriately
	Reading	Follows written instructions
	Replacement skills	Anger management, labels and expresses emotions, recounts weekly events
	Expressive	Identifies and labels objects in Portuguese
	Vocational	Uses email
Micaela	Activities of daily living	Decorates a cake
	Academic	First Nations, plant experiment, creative writing
	Spelling	Spells functional words
	Learning skills	Oral/visual presentation skills, learns new information, staying on task,
	Collaborative problem solving	Collaborative problem solving
	Vocational	Introduction to Microsoft Word, writes a resume, volunteer application, introduction to hairstyling
Anthony	Academic	Learns new information, STEM projects, First Nations
	Replacement skills	Learns and understands behaviour support plan
	Collaborative problem solving	Collaborative problem solving
	Vocational	Typing academy, job/volunteer application, intro to Microsoft Excel, teaches in a dyad, mechanical and electrical systems
	Community	Boy scouts
Taylor	Activities of daily living	Independent daily activity schedule
	Writing	Creative writing

Participant	Domain	Program
	Replacement skills	Collaborative problem solving
	Learning skills	Learns new information
	Expressive	Uses Duolingo to learn French
	Domestic	Baking, plans a meal
	Vocational	Introduction to Microsoft Word, basic computer skills, writes an email, maintains a schedule, learns to type
	Community	Public transportation, volunteering
	Adaptive behaviour	Anger management

Note. STEM = Science, Technology, Engineering, and Mathematics.

Appendix C

Sample Treatment Integrity Checklist

RE: [name]   For BSP Dated:
                                                             Updated:

Name of Staff Evaluated:
Name of Evaluator:
Date(s) and Duration(s) of Observation(s):

Reason for TIC completion: Random/Preliminary/BSP Revisions/Follow-Up

Procedure/Domain	Verbal (V) or Observed (O)?	Proficient	Needs Improvement	Follow-Up Required (Y/N)
Preventative Strategies				
1. Signed [name]'s BSP, guidelines, and protocols.				
2. [name] is woken up at time stated on DAS each morning.				
3. Current activity matches activity on DAS.				
4. Table is cleared and environment is free of any objects that may be used as weapon to harm himself or staff.				
5. [name]'s helmet is in the environment he is currently working in and is not on the floor.				
6. [name] is not permitted to be in the kitchen unless it pertains to programming and DAS interval.				
7. Staff communicate with [name] clearly and concisely.				
8. [name] can choose to wear women's clothing when dressing in the morning, however, cannot change once dressed.				
9. [name] is permitted to wear makeup but must fit it into his morning hygiene time and cannot cash in for it during his leisure. He cannot reapply during the day outside of those times.				

10. Staff refrain from seeking out clinical approval for activities. Staff follow guidelines as written and remind [name] that any changes to rules/expectations can be discussed with the clinical team later, but currently, he must comply with set expectation. Staff seek out clarification discreetly from clinical.				
Reinforcer Guidelines and Rewards				
1. [name] receives points and a checkmark on his DAS for all activities completed on-time in the absence of behaviours.				
2. Staff allow [name] to complete his DAS independently and only intervene with an 'X' if he runs over schedule (over two minutes) or if prompted to complete his DAS following over two minutes.				
3. [name] has access to a timer/clock.				
4. Staff can identify the proper amount of points [name] has earned in his day thus-far.				
5. When [name] finishes an activity early, staff ensure [name] is engaging in one of the activities from the R+ Activities List or has started his next activity interval early.				
6. Exchanges points for equal number of minutes with reinforcer items/activities on leisure time.				
7. Pornography is only accessible during his last evening leisure. He does not have to cash in for it.				
8. Skills, Chores, Cooking must continue until end of DAS interval.				
9. Partial points cannot be earned.				
10. Staff follow smoking schedule as outlined on DAS and [name] cannot have more than one smoke as specified on the DAS.				

11. Staff do not provide [name] with food or beverages outside of DAS/R+ guidelines or low blood sugar protocols. Staff do not buy for or sell items to [name].				
12. [name]'s closet should be locked at night and during the day.				
13. [name]'s bedroom is free of weapons.				
Management of Behaviour				
1. If [name] is agitated, staff reduce their verbal interactions with him.				
2. Staff ignore inappropriate comments of behaviour and model appropriate behaviour. Staff do not use confrontational language or make defensive/argumentative statements.				
3. Agitation and Threats Towards Others During Evening Leisure Interval (7:45pm-10:00pm): Will result in [name] losing access to his electronics for the remainder of that evening leisure interval. He can still cash in for other leisure activities that do not consist of electronic use.				
4. Any agitation or threats after 9:30pm will result in evening DAS. Please reference the evening DAS section in this BSP.				
5. RAFing: Staff deliver expectations once. If [name] continues, staff advise him to bring it up during concerns time once. If [name]continues, staff redirect back to task and ignore.				
6. Staff record duration of RAFing on ABC sheet.				
7. Task Refusal/Early Escalation: Staff will prompt three times, two minutes apart followed by				

five minutes to consider options.				
8. Task Refusal: Only if [name] is willing, staff will engage in CPS to solve problem.				
9. Task Refusal: If CPS fails, staff use gestural then HOH to prompt [name] through activity.				
10. Continued Escalation: Staff do not use active listening. Staff remind [name] of R+ and points system and maintain safe distance.				
11. PA/ED/SIB: Staff immediately implement SMG approved hold.				
12. Staff transfer [name] to mechanical chair/cot (if possible) and implement HOH prompting.				
13. HOH: Staff resort to wiping table until compliance is achieved for a duration of one minute before HOH is released.				
14. If [name] attempts to engage in self-injurious behaviour, physical aggression, or environmental destruction, the helmet is applied.				
15. Only the lead staff communicates with [name] and active listening is not used during HOH.				
16. Grounding Protocol/Re-Integration: If [name] attempts to elope or successfully elopes from the home and/or escalates within the community, then he will be placed on a temporary grounding procedure.				
17. Grounding Protocol/Re-integration: What happens during the first week of the grounding procedure? [name] cannot leave the treatment home for a week.				
18. Grounding Protocol/Re-integration: There are 4 re-integration phases after the first week of [name]'s grounding.				

Staff should track the re-integration phases on [name]'s visual calendar.				
19. Bedtime Routine: If SIB while in bed, staff give three verbal warning spaced two minutes apart that cot and pinels will be applied. Staff call in another treatment home upon first verbal warning.				
20. Bedtime Routine: If SIB after third reminder, cot and pinels are applied for remainder of the night.				
21. Bedtime Routine: If pinels/cot are applied, staff complete circulation checks every 30 minutes.				
Data Collection				
1. All data is recorded as required (ABC and DAS).				
2. Completes data to its entirety with no information left blank/missing.				
3. Data is recorded accurately.				
TIC Score: __ / __ (*100) = __%				

Appendix D

Sample Treatment Integrity Checklist with Standardized Questions and Notes

RE: [name]   For BSP Dated:
                                                       Updated:

Name of Staff Evaluated:
Name of Evaluator:
Date(s) and Duration(s) of Observation(s):

Reason for TIC completion: Random/Preliminary/BSP Revisions/Follow-Up

Procedure/Domain	Verbal (V) or Observed (O)?	Proficient	Needs Improvement	Follow-Up Required (Y/N)
Preventative Strategies				
1. Staff have signed [name]'s BSP, guidelines, and protocols.				
2. Staff are wearing the appropriate PPE (bump caps and Kevlar sleeves).				
Can you talk about topics such as religion, relationships, politics, etc.? 3. Staff maintain professional boundaries, not talking about topics such as religion, relationships, politics, etc.				
What time is [name] woken up? 4. [name] is woken up at 8:30AM. **Note:** *[name] is required to wake up independently (Guidelines, 01/05/2022).*				
5. Current activity matches activity on DAS.				
6. Environment is free of any objects that may be used as weapon to harm herself or staff.				
7. [name]'s helmet is in the environment she is currently working in and moved to each new environment with her.				
8. [name]'s bedroom door is locked.				
When can [name] use the kitchen? 9. [name] uses the kitchen while programming (mealtime, preferred drink).				

How should you communicate with [name]? 10. Staff communicate with [name] clearly and concisely. They use short statements that do not use too many big words/abstract concepts.				
Can you discuss personal information with [name] and what type of relationship should you maintain with her? 11. Staff refrain from discussing personal information with [name]. They maintain a professional relationship, identifying themselves as treatment staff.				
Can you provide [name] with food, gifts, nicknames, and/or gestures that blur professional boundaries? 12. Staff do not provide [name] with food, gifts, nicknames, and/or gestures that blur professional boundaries.				
When are clinical team members allowed to engage with [name]? 13. Staff only allow clinical staff to engage with [name] during her "talk times."				
What are the criteria for [name] to earn a preferred drink? 14. [name] can only have a preferred drink when and if she demonstrates six consecutive hours of zero instances of antagonizing behaviour, verbal aggression, physical aggression, environmental destruction, self-injury, and sexualized behaviours.				
What are the criteria for [name] to earn a 'Tea Party', when does a 'Tea Party' occur, and what drink can she choose? 15. If [name] earns 24/26 checkmarks on her DAS for that given day, [name] is able to earn a 'Tea Party', which				

occurs at 8:00PM and she can choose any preferred drink. **Note:** *It is also correct if they say she will earn a "Tea Party" if she completes 90% or more on her DAS (BSP, 03/30/2022).*				
Which target behaviours will result in tokens being re-set and when can [name] begin to re-earn tokens? 16. If [name] engages in antagonizing behaviour, verbal aggression, physical aggression, environmental destruction, self-injury, and sexualized behaviours, the tokens will be reset and she may being to re-earn tokens from the end time of the behaviour on every hour. **Note:** *It is also correct if they do not say she may begin to re-earn tokens from the end time of the behaviour on every hour (Guidelines, 01/05/2022).*				
Reinforcer Guidelines and Rewards				
Which activities can [name] receive points and a checkmark for on her DAS? 1. [name] receives points and a checkmark on her DAS for all activities completed on-time in the absence of negative behaviours.				
What is the criterion for [name] to receive a checkmark on her DAS? What happens if she does not meet the criterion? 2. [name] will only receive a checkmark on her DAS if she signs her DAS independently within 30-seconds of completing an interval. If she does not, staff will remind her to do so and will tell her she received an "X" for the interval.				
3. [name] has access to a timer/clock.				

How many points has [name] earned in her day so far? 4. Staff can identify the proper amount of points [name] has earned in his day thus-far.				
What is the criterion for [name] to be provided with 45 additional points? 5. [name] is provided with 45 additional points if she does not engage in antagonizing behaviour, verbal aggression, sexualized behaviour, physical aggression, self-injury, and/pr environmental destruction during the designated time frames.				
What should you do when [name] finishes an activity early? 6. When [name] finishes an activity early, staff ensure that [name] is engaging in a functional activity until the next interval starts. **Note:** *It is also correct if they do not do this or do not say functional activity (BSP, 03/30/2022).*				
How are points exchanged, what can she exchange points for, and when can she exchange points? 7. Exchanges points for equal number of minutes with reinforcer items/activities on leisure time.				
Which activities does [name] have to cash in to do in her bedroom? Which activities doe she not have to cash in for? 8. [name] cashes in to do a leisure interval in her bedroom, however, does not cash in for a nap and/or private time. **Note:** *It is also correct if they specify that [name] can cash in to enter bedroom (i.e., her going into her bedroom and doing whatever she wants in there) and that she does not have to cash in for private time and/or break in room (Guidelines, 01/05/2022).*				

Do skill-based activities have to continue until the end of the DAS interval? What should you do if she finishes the task early? 9. Skill-based activities must continue until the end of DAS interval. If she finishes the task early, staff will prompt her to complete another skill-based activity.				
During an interval, how long should she engage in a chore activity? What should you ensure regarding the chore that [name] is currently completing? 10. Chore activities must continue until the end of the DAS interval. Staff ensure that the chore [name] is completing does not have a checkmark/completion mark on it.				
Describe [name]'s chore interval, including information on what she needs to do to earn her checkmark for the interval. 11. [name] will set a 30-minute timer on her timer. She will choose a chore from the chore list. She will continue to do chores for the entire interval. If there are more than five minutes, she chooses another chore. If there is less than five minutes, she does not need to do another chore.				
Describe [name]'s physical activity interval, including her current physical activity target, who chooses the activity she engages in, and whether it can be repeated. 12. Staff can identify the current physical activity target (exercises for the entire interval, five-minute exercise, 2-minute break). Staff choose the activity [name] is engage in and it cannot be repeated.				
Can partial points be earned?				

13. Partial points cannot be earned.				
What is [name]'s smoking schedule and smoking protocol? 14. Staff follow smoking schedule as outlined on DAS. Staff return the lighter as soon as [name] is finished smoking.				
Can you provide [name] with food or beverages outside of DAS/R+ guidelines and can you buy for or sell items to [name]? 15. Staff do not provide [name] with food or beverages outside of DAS/R+ guidelines. Staff do not buy for or sell items to [name].				
Describe [name]'s meal interval, including information on what she needs to do to earn her checkmark for the meal interval. 16. [name] must eat at least 50% of her meal in order to get a checkmark on her DAS for meal completion. Additionally, [name] washes all her dishes. **Note:** *It is also correct if they do not say 50% (Guidelines, 01/05/2022).*				
On which days does [name] have scheduled outings on her DAS? 17. Community Access: [name] has scheduled outings on her DAS every Mon/Wed/Fri.				
During which intervals can [name] use the washroom? 18. [name] may only use the washroom during meal/leisure/hygiene intervals.				
Management of Behaviour				
What should you do if [name] is agitated? 1. If [name] is agitated, staff reduce their verbal interactions with her.				
What should you do if [name] makes inappropriate comments or negative verbal behaviour? 2. Staff ignore inappropriate comments or negative verbal				

|				
behaviour and model appropriate behaviour.				
What should you do if [name] is persistently asking questions to obtain a specific/desired answer? 3. If/when [name] is persistently asking questions to obtain a specific/desired answer, then staff tell her to ask a BT/BTA on her talk time and then ignore all follow-up questions while redirecting her back to task.				
What should you do if [name] engages in noncompliance or task refusal? 4. Noncompliance/Task Refusal: Staff deliver expectations once. If [name] continues, noncompliance behaviours, they implement prompt hierarchy (first verbal, second verbal, gestural, partial, and full physical). ==Note:== *It is also correct if they do not say that they have to deliver expectations once (BSP, 03/30/2022).*				
What happens if [name] engages in antagonizing behaviours, verbal aggression, and sexualized behaviours? 5. Antagonizing Behaviour, Verbal Aggression, and Sexualized Behaviours: If [name] engages in antagonizing behaviours, verbal aggression, and sexualized behaviours, she will receive an "X" on her DAS, her preferred beverage tokens will be reset, and she will lose her additional 45 points during the designated time frame.				
What happens if [name] continues to engage in antagonizing behaviours, verbal aggression, and sexualized behaviours after you provide her with the warning?				

6. If she continues to engage in antagonizing behaviours, verbal aggression, and sexualized behaviours after staff provided her with the warning, the prompt hierarchy will be implemented from least-to-most (gestural > partial > full physical). She will be prompted to do functional activity if there is no task at hand (mealtimes, leisure, etc.).			
What target behaviours result in the SMG approved physical containment? Under what circumstances will you make the call to transfer her to the mechanical chair? 7. Physical Aggression/Environmental Destruction/Self-injurious Behaviours: Staff immediately implement SMG approved physical containment. If staff cannot safely contain [name] using physical containments, lead staff will make the call to transfer her to the mechanical chair. The mechanical chair is a LAST RESORT.			
When is the helmet applied and when is it removed? 8. If [name] attempts to bite herself or others and/or headbutts, the helmet is applied. Remains on for five minutes after behaviour is over.			
Who communicates with [name] during HOH and is active listening used during HOH? 9. Only the lead staff communicates with [name] and active listening is not used during HOH.			
How would you fade the physical containment and/or mechanical chair straps? 10. Staff fade the physical containment and/or mechanical chair straps following each set			

of 10 calm and independent responses (i.e., in the absence of all negative target behaviours).				
Where should staff be positioned while [name] is using the stairs? 11. Staircase Protocol: One staff remains in from of [name] while another staff remains behind [name]. **Note:** *It is also correct if they say staff will walk two steps ahead of her (BSP, 03/30/2022).*				
What is the staircase protocol if [name] begins to escalate on the stairs? 12. Staircase Protocol: If [name] begins to escalate, staff disengage and wait [name] out. Two at the bottom of the stairs and two at the top of the stairs while providing verbal directives every two-minutes.				
What kind of behaviour does [name] need to engage in prior to entering the bathroom, and for how long? 13. Bathroom Protocol: Prior to entering the bathroom, [name] shows 15 minutes of calm and in-control behaviour. When showing 15 minutes of calm, [name] must do chores. **Note:** *It is also correct if they say the task at hand (interval she is in) may be extended to obtain the calm duration, or she may complete chores if it does not make sense to extend the interval for more time (e.g., lunch, leisure, or social) (BSP, 03/30/2022).*				
How close should you remain to [name] while she is shaving? 14. Bathroom Protocol: While shaving, staff remain within arms-length of [name] at all times.				
What should you do if [name] engages in verbal aggression while in the bathroom, and what				

should do if she continues to engage in verbal aggression? 15. Bathroom Protocol (Verbal Aggression): If [name] engages in verbal aggression, she will be given one verbal reminder, however, if she continues, she will receive an "X" for the interval and her hygiene will be terminated. Staff use gestural > full physical to remove [name] from the washroom.				
Which target behaviours are considered imminent risk in the bathroom? What is the protocol if [name] engages in these target behaviours in the bathroom? 16. Bathroom Protocol (Physical Aggression, Environmental Destruction, and Self-injurious Behaviours): If [name] is at imminent risk while in the washroom, staff will call for assistance. They will enter the washroom, use a corner/wall containment. Helmet will be applied for headbutts and bites. Mechanical chair applied, HOH, 15-minutes of calm before re-entering the washroom.				
Data Collection				
4. All data is recorded as required (ABC and DAS).				
5. Completes data to its entirety with no information left blank/missing.				
6. Data is recorded accurately.				
TIC Score: __ / __ (*100) = __%				

Appendix E

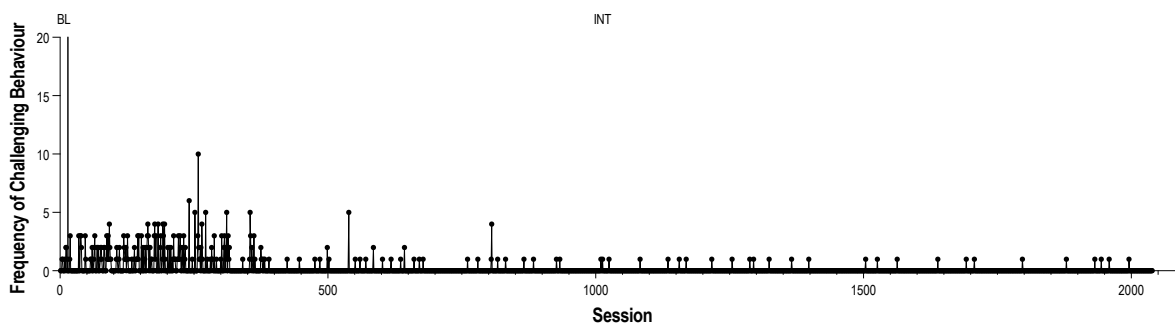**Nonconcurrent Multiple Baseline Graph for Frequency of Challenging Behaviour**
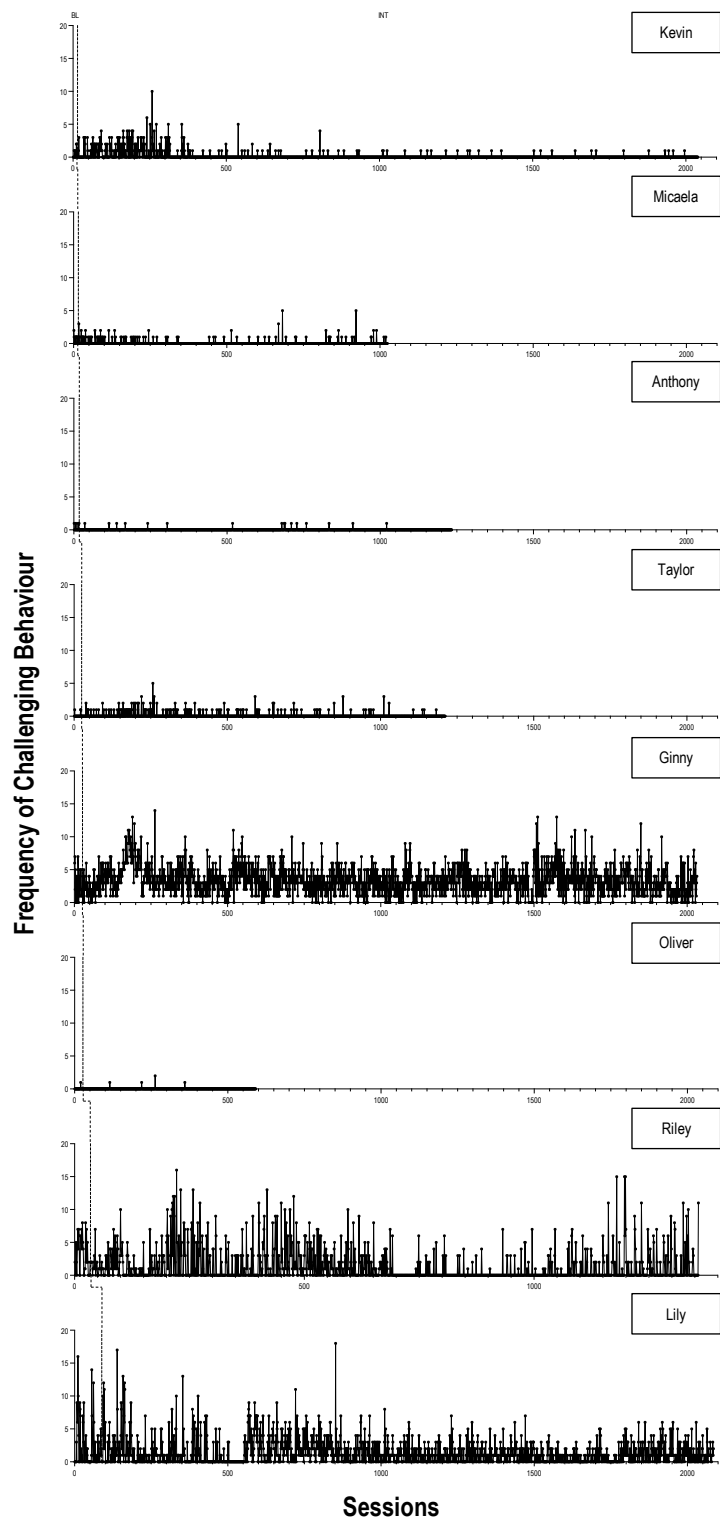
Note. This graph depicts the frequency of challenging behaviour (y-axis) across treatment sessions (x-axis) for all participants within a nonconcurrent multiple baseline design. The dotted line represents the phase change from baseline (BL) to intervention (INT).
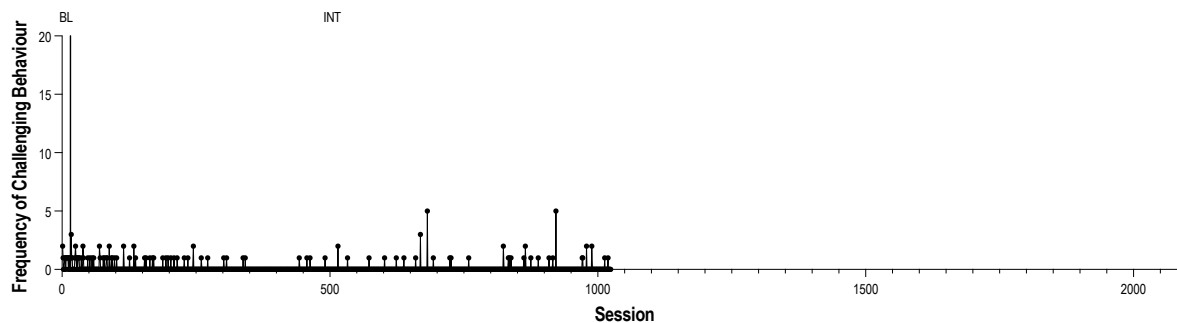
Appendix F

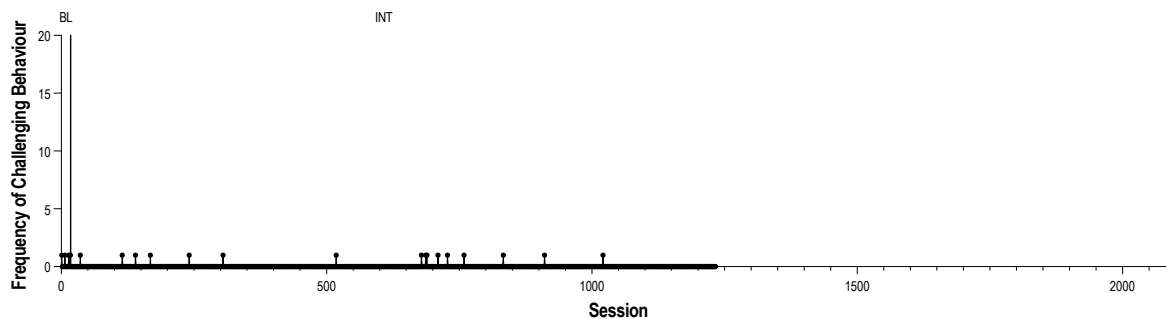**Time-Series Graphs for Frequency of Challenging Behaviour**

Figure F1

Frequency of Challenging Behaviour across Baseline and Intervention for Kevin

Note. This graph depicts baseline and intervention data for Kevin's frequency of challenging behaviour across treatment sessions. The x-axis depicts the session number and the y-axis depicts the frequency of challenging behaviour. The dotted line represents the phase change from baseline to intervention.

Figure F2

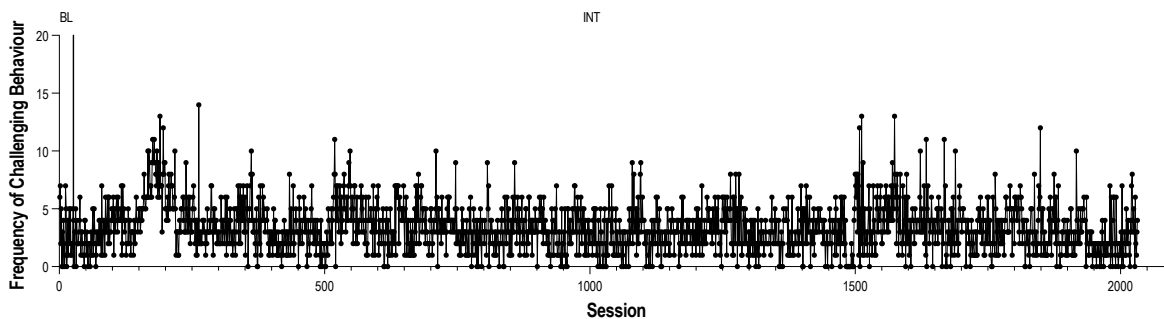*Frequency of Challenging Behaviour across Baseline and Intervention for Micaela*

Note. This graph depicts baseline and intervention data for Micaela's frequency of challenging

behaviour across treatment sessions. The x-axis depicts the session number and the y-axis depicts

the frequency of challenging behaviour. The dotted line represents the phase change from

baseline to intervention

Figure F3

Frequency of Challenging Behaviour across Baseline and Intervention for Anthony

Note. This graph depicts baseline and intervention data for Anthony's frequency of challenging behaviour across treatment sessions. The x-axis depicts the session number and the y-axis depicts the frequency of challenging behaviour. The dotted line represents the phase change from baseline to intervention.

Figure F4

Frequency of Challenging Behaviour across Baseline and Intervention for Taylor
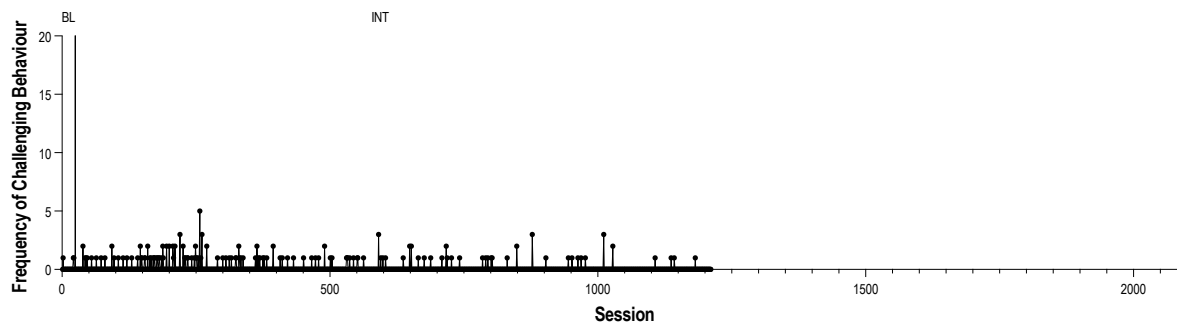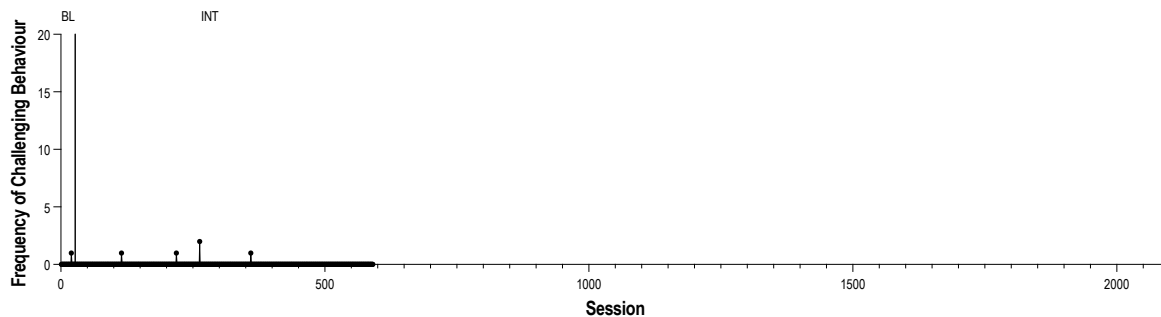

Note. This graph depicts baseline and intervention data for Taylor's frequency of challenging

behaviour across treatment sessions. The x-axis depicts the session number and the y-axis depicts

the frequency of challenging behaviour. The dotted line represents the phase change from

baseline to intervention.

Figure F5

Frequency of Challenging Behaviour across Baseline and Intervention for Ginny
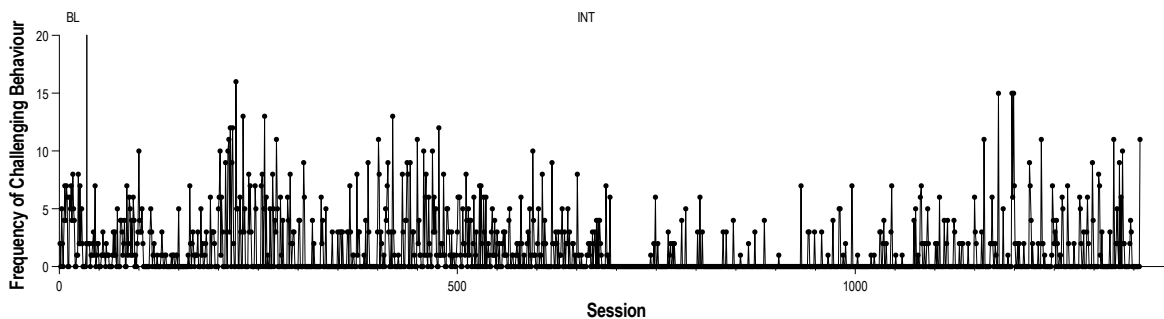

Note. This graph depicts baseline and intervention data for Ginny's frequency of challenging behaviour across treatment sessions. The x-axis depicts the session number and the y-axis depicts the frequency of challenging behaviour. The dotted line represents the phase change from baseline to intervention.

Figure F6

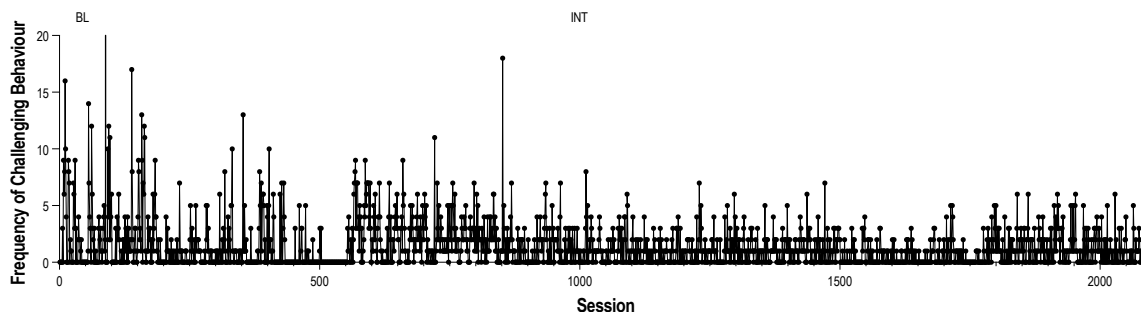*Frequency of Challenging Behaviour across Baseline and Intervention for Oliver*

Note. This graph depicts baseline and intervention data for Oliver's frequency of challenging behaviour across treatment sessions. The x-axis depicts the session number and the y-axis depicts the frequency of challenging behaviour. The dotted line represents the phase change from baseline to intervention.

Figure F7

Frequency of Challenging Behaviour across Baseline and Intervention for Riley

Note. This graph depicts baseline and intervention data for Oliver's frequency of challenging behaviour across treatment sessions. The x-axis depicts the session number and the y-axis depicts the frequency of challenging behaviour. The dotted line represents the phase change from baseline to intervention.

Figure F8

Frequency of Challenging Behaviour across Baseline and Intervention for Lily

Note. This graph depicts baseline and intervention data for Lily's frequency of challenging

behaviour across treatment sessions. The x-axis depicts the session number and the y-axis depicts

the frequency of challenging behaviour. The dotted line represents the phase change from

baseline to intervention.

Appendix G

Sample Skills Lesson and Checkpoint

Aspects of Critical Thinking

<u>**First**</u>
Staff will read the following passage to the client. Staff should ensure that the client is attending. Staff will prompt the client to repeat the nine aspects of critical thinking after the passage has been read. If the client is unable to repeat the nine aspects of critical thinking, staff will re-read them to the client and re-prompt the client to repeat them.

"In this unit of Critical Thinking and Problem-Solving Skills, you will learn about various aspects of critical thinking, including: Observation, Analysis, Interpretation, Reflection, Evaluation, Inference, Explanation, Problem-Solving, and Decision Making."

Staff initial: _____ / Date: _____

<u>**Second**</u>
Staff will prompt the client to play the following video on YouTube. While the video is playing, it's expected that the client is *taking notes* and documenting all of the *most important* information. Staff may assist the client with summarizing topics discussed within the video, if necessary. However, the client should attempt to take notes independently.

YouTube: What is Critical Thinking?
By: Macat

Staff initial: _____ / Date: _____

<u>**Third**</u>
Staff will prompt the client to play the following video on YouTube. While the video is playing, it's expected that the client is *taking notes* and documenting all of the *most important* information. Staff may assist the client with summarizing topics discussed within the video, if necessary. However, the client should attempt to take notes independently.

YouTube: Episode 1.1: What is Critical Thinking?
By: Psychlopaedia.org

Staff initial: _____ / Date: _____

<u>**Fourth**</u>
Staff will provide the client with the following activity. The client will be required to identify which of the following terms are the aspects of critical thinking.

Application: Aspects of Critical Thinking

Staff initial: _____ / Date: _____

- Learning Check Point -

The client is required to list the aspects of critical thinking that they've learnt about thus far in the unit. Staff will probe this twice, on separate, yet consecutive days.

SD: Labels The Aspects Of Critical Thinking
Date: _____ : _____ / Date: _____ : _____ /
Answer: *Observation, Analysis, Interpretation, Reflection, Evaluation, Inference, Explanation, Problem-Solving, Decision Making.*
Mastery Criteria: 100% over 2 consecutive days, otherwise, the section of this unit will be restarted.

Second Attempt (If Applicable)
SD: Labels The Aspects Of Critical Thinking
Date: _____ : _____ / Date: _____ : _____ /
Answer: *Observation, Analysis, Interpretation, Reflection, Evaluation, Inference, Explanation, Problem-Solving, Decision Making.*
Mastery Criteria: 100% over 2 consecutive days, otherwise, the section of this unit will be restarted.

Once completed, please show a clinician for their approval.

BT/BTA: **_____**   ***Date:*** **_____**