Published in final edited form as:

Int J Dev Disabil. 2016 ; 62(3): 183–191. doi:10.1080/20473869.2016.1173316.

Measuring the Complexity of Treatment for Challenging Behavior using the Treatment Intensity Rating Form

Jennifer Zarcone, Louis Hagopian, Jennifer Ninci, Chloe McKay, Andrew Bonner, Christopher Dillon, and Nicole Hausman

Kennedy Krieger Institute and Johns Hopkins School of Medicine

Abstract

Objectives—The goal of this study was to develop and evaluate a tool to measure the complexity and intensity of psychotropic medication interventions, behavioral interventions, and issues related to crisis management for challenging behavior using a standardized rating form.

Method—The Treatment Intensity Rating Form (TIRF) is a 10-item scale with three categories: pharmacological interventions, behavior supports, and protective equipment. In a retrospective review we examined the final treatment recommendations for 74 individuals with self-injurious behavior (SIB) based on psychiatric and behavioral notes and reports. We also compared whether TIRF scores differed across individuals for whom SIB was maintained by social reinforcement (e.g., to access attention or toys/activities, or escape from tasks) versus those for whom SIB was maintained by automatic reinforcement (e.g., occurs independent of social variables, and is presumed to be maintained by sensory reinforcement).

Results—The TIRF was demonstrated to have strong inter-rater reliability (98%) and appears to have good face validity. As hypothesized, individuals with SIB maintained by automatic reinforcement had significantly more medication trials ($p=0.0005$) and required more protective equipment than individuals with SIB maintained by social reinforcement ($p=0.0002$). Antidepressant medication was used more often with individuals with automatically reinforced SIB, although antipsychotics and anticonvulsants were also commonly used across both groups.

Conclusion—Findings provide initial support for the TIRF's reliability, and face validity as a measure the level of complexity of medical and behavioral treatment plans - although additional research is needed to fully evaluate its psychometric properties.

Keywords

Behavioral treatment; Intensity; Complexity; Challenging Behavior; Self-injurious behavior; Medication

The term treatment intensity has been used to characterize behavioral, educational, and medical interventions in terms of their complexity, as well as the level of effort and resources required to carry them out. Given that interventions are often multidimensional,

Address correspondence to Jennifer Zarcone, Ph.D., Neurobehavioral Unit, Kennedy Krieger Institute, 707 N. Broadway, Baltimore, MD 21205. Zarcone@kennedykrieger.org.

Its contents are solely the responsibility of the authors and do not necessarily represent the official views of NICHD.

the intensity of treatment can be measured in several ways. For example, the intensity of medical interventions has been measured based amount and type of intervention needed, such as to successfully eliminate a medical condition (e.g., Engert et al., 2010; Thorpe & Howard, 2006), or the number of interventions used or days in treatment for end of life care (e.g., Barnato, McClellan, Kagay, & Garber, 2004). These studies were conducted to capture the degree to which either less invasive or less costly procedures compare to more invasive (or intensive) treatments in terms of outcome as part of a cost-benefit or epidemiological analysis.

Behavioral services or interventions also can be measured in terms of treatment intensity. One method that has been employed is based on the duration of or amount of exposure to an intervention, often conceptualized as the “dose” of the intervention (e.g., Lovaas, 1987; Goode, Winkler, Lawler, Reeves, Owen, & Eakin, 2011). For example, Lovaas (1987) evaluated long-term outcome data for 38 children with autism who received either a high dose (more than 40 hours per week) or a low dose (less than 10 hours) of one-to-one treatment per week. The experimenters demonstrated that the group that received high dose of behavioral intervention for 2 or more years were more likely to have higher IQ scores and less restrictive educational placements than the group that did not receive as many hours of treatment per week. Similar meta-analyses of research published using Early Intensive Behavioral Interventions have demonstrated similar outcomes in terms of improvements in standardized tests scores and behavior change (Peters-Scheffer, Didden, Korzilius, & Sturmey, 2011). Alternatively, some have described “dose” as the number of teaching or learning opportunities in a time period rather than the duration of treatment. By looking at the number of teaching trials, rather than total number of hours of exposure to a treatment, a better measure of density of intervention per unit of time can be made (Warren, Fey, & Yoder, 2007).

The studies described above focus on how the amount of exposure (i.e., the dose) impacts outcomes achieved with a single method of intervention. However, as noted above, there are other dimensions of behavioral treatments and services that contribute to the complexity and intensity of the intervention including the level of support required to implement an intervention, the number of interventions tested and required, and the restrictiveness of treatment components. Behavioral interventions used to target challenging behavior are typically evaluated solely based on the extent to which they reduce challenging behavior (i.e., their “effectiveness”). Generally, an 80% or greater reduction in the rate of challenging behavior (relative to baseline) has become established as a benchmark indicating a clinically successful outcome. Although measuring the reduction in the rate of challenging behavior provides an individualized, quantitative measure of effectiveness, this method is incomplete in that it fails to consider the complexity of the intervention. The importance of measuring the intensity of behavioral interventions is difficult to overstate for two related reasons. First, behavioral interventions for challenging behavior require continuous, and active application of the intervention on the part of others (e.g., arranging the environment to minimize events that may occasion challenging behavior, reinforcing appropriate behavior, and reacting to each occurrence of challenging behavior in a manner that does not reinforce it). Second, because sustaining these interventions can be effortful, the more complex they are, the more resources are needed (including staff) to implement these interventions with good integrity.

Treatment of challenging behavior in individuals with IDD also often involves the use of medications; therefore there is also a need to consider what type of medications are in use when evaluating the complexity of interventions for challenging behavior. Although medication administration is not as effortful as implementing behavioral interventions, more intensive pharmacological interventions may be more difficult to sustain over time because of adverse effects and increased monitoring requirements. Thus, using only the traditional measure of treatment effectiveness (reduction in challenging behavior), two similarly effective treatments appear equivalent, even though one might be a relatively simple treatment that is easily sustainable, while the other is highly complex treatment requiring intensive resources and medications that require close monitoring. It is easy to see that measuring both effectiveness and intensity of interventions would provide a more comprehensive measure of outcomes.

The purpose of the current study was to develop the Treatment Intensity Rating Form (TIRF) to quantify the complexity of interventions for challenging behavior. In addition, we evaluated interrater agreement and the face validity of the form. We evaluated treatment intensity both in terms of medical/psychiatric treatment (medication) as well as the type and number of components of the child's behavioral interventions. In addition, we applied the TIRF to compare two groups with SIB known to be differentially resistant to treatment. SIB maintained by automatic reinforcement has been shown to be more resistant to behavioral treatment than SIB maintained social variables (Hagopian, Rooker, & Zarcone, 2015). However, differences between these groups have not been quantified in terms of all aspects of treatment complexity, including the use of medications, staffing, and crisis management resources. We hypothesized that the TIRF would detect quantifiable differences in the complexity of interventions across these groups, with those with automatically reinforced SIB having higher TIRF scores.

Method

Participants and Setting

The clinical files of patients treated between the years of 2003 to 2015 on the Neurobehavioral Unit at the Kennedy Krieger Institute who engaged in SIB were identified as eligible for inclusion. The Neurobehavioral Unit (NBU) is a 16-bed inpatient program dedicated to the assessment and treatment of individuals with severe challenging behavior that have not responded to treatment in less intensive settings. The program is nationally known and has served individuals from 38 states across the US. We use a multidisciplinary approach to treatment combining psychiatric care (i.e., psychotropic medications trials) with behavioral interventions based on the principles of applied behavior analysis (Hagopian & Anderson, 2009). In general, the average admission length is 4–5 months for patients on our unit; this varied significantly for the study participants (mean length of stay = 7.2 months, range 1.9 – 20.8 months). For five study participants, the length of stay was over a year due to a variety of residential placement and health complications. To be included in this analysis, the individual's primary targeted behavior was SIB (e.g., head banging, head hitting, body hitting, self-biting, and skin-picking). Some participants also engaged in aggression, property destruction, and other disruptive behaviors.

The retrospective record review was approved by the Institutional Review Board of Johns Hopkins Hospital. Clinical records were reviewed for 74 individuals with SIB (one individual had two admissions several years apart and the TIRF from each admission was counted separately). Participants were between the ages of 5 and 22 years (mean age = 13.9 years). Sixty-eight percent of the participants were male, 32% were female. All patients who met the inclusion criteria were included in this analysis and groups were not matched by age or gender. See Table 1 for a description of the participants.

Rating Scale

The Treatment Intensity Rating Form (TIRF) is a 10-item scale with three subcategories: pharmacological interventions, behavior supports, and protective equipment (see Appendices A and B for individual items on the TIRF as well as the scoring rules). The TIRF was initially developed as a clinical rating scale to characterize the intensity of treatment plans to determine whether a patient's behavior requires extensive community and home-based supports (i.e., staffing ratios, psychiatry oversight). For this study, we also wanted to measure the intensity of treatment for those individuals who engaged in SIB and evaluate whether TIRF scores varied across two groups known to differ in terms of resistance to behavioral treatment based on other indices (Hagopian, Rooker, Bonner, Arevalo, & Zarcone, in press).

All 10 questions on the TIRF were scored using a 0–5 point rating scale by trained master's level staff using each patient's records. A score of 0 on a component indicated less intrusiveness and of lower intensity. As the score increased towards 5, the components were more intrusive and generally of higher intensity. If for any question, the rater was unable to find sufficient information to rate an item, it was scored as Unknown ("U"). All information was collected from clinical documentation used to monitor and summarize a patient's care (e.g., discharge summaries and behavior plans).

There are three subscales for the TIRF. The first subscale is on pharmacological interventions and referred to the number of medication trials conducted during a patient's admission (i.e., the psychotropic medications the patient was admitted on as well as any new medications evaluated) and the number of medications the individual was prescribed at discharge. The second subscale refers to the behavioral supports that are part of the individual's treatment plan including the total number of components of the behavior intervention plan, the number of proactive and reactive procedures, the crisis management procedures recommended, and staff levels required to maintain the individual's safety and implement the plan. The third subscale refers to the number and type of protective equipment needed for the individual and the staff, as well as the level of mechanical restraint needed to maintain the individual's safety. See Appendices A and B for the items and rules for scoring of the TIRF.

Interrater agreement

For 23 of the participants (31%), a second rater scored the TIRF independently using the same medical records as the primary rater. Agreement was calculated by comparing the answers provided by the primary rater with the secondary rater and dividing the lower score

by the higher score for each item (unless only one of them scored a “U” and this was counted as a disagreement). The percentages were summed and divided by 10 (the total number of items) and multiplied by 100 to obtain a percentage. Results indicated very good interrater agreement: mean = 98.03% (range = 86.7 – 100%) across all participants.

Functional Analysis Outcome

As part of the standard of care for the unit, an analog functional analysis was conducted with all participants during their first few weeks of admission to the unit, based on the procedures described by Iwata, Dorsey, Slifer, Bauman & Richman (1994). This procedure involved presenting a variety of test conditions to the participant using a multielement, reversal, or test-control/pairwise single-subject experimental design. Sessions were 5–10 minutes in duration, depending on the severity of the behavior, and were conducted in a therapy room, classroom, or in an area on the main living unit. All SIB, aggression, disruption, and property destruction resulted in the consequences described below; however, biting others, head banging, fecal smearing, disrobing, and eloping were always blocked for safety.

The following are the most common test conditions evaluated. In the *Alone* condition, the individual is in the room alone and without any toys or activities available. There are no consequences for challenging behavior. This condition tested for challenging behavior maintained by automatic reinforcement. For some individuals with SIB, there was a therapist present to block SIB if it became dangerous, but he/she ignored all behavior otherwise. Thus, it was a *No Interaction* condition and similarly tested for automatic reinforcement. In the *Attention* condition the therapist provided attention in the form of a brief statement of concern (e.g., “I don’t like it when you do that”) contingent on challenging behavior. This condition tested for behavior maintained by positive reinforcement in the form of access to adult attention. For some individuals, a *Tangible* condition was also used to evaluate whether positive reinforcement in the form of gaining access to toys or food, maintained challenging behavior. In this condition, the individual had access to his or her most highly preferred toy or food for two minutes prior to session. At the start of the session, the item was removed and 30 seconds access to the item was provided contingent on challenging behavior. In the *Escape* condition, academic and vocational demands were presented to the individual using a three-step prompting procedure (i.e., verbal, gestural, full-physical). Contingent upon a target behavior, the individual received 30 seconds of escape from the task. This condition tested for challenging behavior maintained by negative reinforcement in the form of escape from demands. Finally, a *Control* condition was always included to which all other conditions were compared. For this condition the individual had access to preferred toys and attention was provided every 30 seconds. No consequences were provided contingent on challenging behavior. All conditions were conducted with the participants as part of their treatment plan.

Based on the results of the FA, we categorized all participants by function and identified those who had a social function (i.e., challenging behavior occurred consistently in the attention, tangible, and/or escape condition) and those that had an automatic function (i.e., challenging behavior occurred at a consistently high rate in all conditions or in the Alone/No Interaction condition). Thus, 37 individuals engaged in socially reinforced SIB and 38

engaged in SIB that was maintained by automatic reinforcement. For those individuals with SIB maintained by social reinforcement, 20 (54%) had multiple social functions, six (16%) had an attention function, four (11%) had a tangible function, and six (16%) had an escape function. For one individual, challenging behavior occurred when he was blocked from engaging in rituals or compulsive behavior.

In addition to categorizing by behavioral function, data were collected on the percentage reduction in challenging behavior after treatment was in place based on data collected across the day by 1-to-1 direct care staff. Percentage reduction was calculated by the behavioral team using the mean rate of challenging behavior for each participant during his/her final treatment phase, dividing it by the rate of behavior during the individual's most appropriate baseline phase and subtracting from 100. Thus, if the mean rate of challenging behavior during baseline was 10 responses per hour and was reduced to 2 responses per hour when treatment was implemented, the person would have an 80% reduction in challenging behavior due to treatment.

Statistical Analysis

The two groups were compared using a Mann Whitney t-test for nonparametric data. The means for each item on the TIRF was compared using a significance level of 0.05.

Results

Mean ratings for each item of the TIRF are presented in Table 2 for both study groups. A Mann Whitney t-test for nonparametric data was used to compare the social and the automatic reinforcement group. Overall, the social reinforcement group had lower mean TIRF scores than the individuals in the automatic reinforcement group on all items except Items 5 (number of reactive procedures) and 7 (staffing levels required to maintain safety) in which the mean was the same for both groups. The differences were statistically significant for Items 1 (number of medication trials), 8 (protective equipment recommended for the child), and 10 (use of mechanical restraint) for the automatic reinforcement group.

Data were also collected on the type of medications each participant was prescribed at discharge. The percentage of each group by class of medication (i.e., anticonvulsants, antipsychotics, antidepressants, and benzodiazepines) is shown in Figure 1. Given that individuals with automatically maintained SIB are considered more difficult to treat, we expected them to be taking more medications than the social reinforcement group, but found that those individuals with SIB maintained by social reinforcement were more likely than the automatic reinforcement group to be taking anticonvulsant medication (51.4% and 44.4% for socially- and automatically-maintained SIB, respectively) and antipsychotics medication (75.7% and 68.4% for socially- and automatically-maintained SIB, respectively). The most common antipsychotic prescribed was risperidone, followed by aripiprazole and lithium (which was included in the antipsychotic category for this study). The individuals with automatically-maintained SIB were prescribed more antidepressants (60.5%) than the socially maintained-SIB group (60.5%), and both groups were nearly the same in the percentage taking benzodiazepines (24.3% and 26.3% for socially- and automatically-maintained SIB, respectively).

Mean percentage reduction from baseline due to treatment was also calculated for each group and is noted in the last column of Table 1. As can be seen, the mean percentage reduction for the social reinforcement group was 83.41% while the mean for the automatic reinforcement group was slightly lower (81.31%). These differences are not statistically significant, however. These data are preliminary at best given that interobserver agreement was not assessed across the day when baseline and treatment contingencies were in place on the unit.

Conclusions

Based on an analysis of 74 individuals hospitalized for the assessment and treatment of challenging behavior, the TIRF was shown to have good interrater reliability and face validity. We found important differences on the TIRF between the treatment plans across two groups known to be different with regard to their resistance to treatment using behavioral interventions. Specifically, the treatment plans for the automatic reinforcement group were more complex, in terms of the number of medication trials and the type of protective equipment required to safely manage the individual – relative to the social reinforcement group. Data also support that these treatment plans are generally effective, resulting in an 80% or greater reduction in challenging behavior compared to baseline. Thus, although the plans may have been complex, they were also effective.

Based on our analysis, it appeared that the social reinforcement group had plans that are equally complex as the automatic reinforcement group, but that more supports may be needed, both in terms of medication and the use of protective equipment to address possibly comorbid psychiatric conditions and safety concerns. Additional research with a variety of participants who vary with regard to severity of challenging behavior is needed before we can determine if the TIRF is useful to fully characterize the intensity of behavioral treatments.

Additional modifications may be required for the TIRF to make it useful across a variety of settings and populations. For example, for Item 1, the range of medication trials at each level was chosen based on our clinical experience that 8 trials represented a high “intensity” for an individual, but there are other ways this could have been categorized or the maximum changed for this item. In addition, on Item 7 (staffing levels required to maintain safety), the mean for both groups was 0.43, which is very low and may be related to the way in which Item 7 was written for the patients we serve. Specifically, many of the individuals we support engage in very severe challenging behavior and their treatment plans generally always include the recommendation for 1-to-1 staffing. Thus, the lowest score (“0”) indicated 1-to-1 staffing across all waking hours, when in fact, lower staffing ratios may be recommended for individuals with less severe challenging behavior or who are not dangerous to themselves or others. Thus, Item 7 should be rescaled to accommodate a wider variety of options for supervision.

These results are limited because the population sampled included patients with very severe SIB that required hospitalization for treatment. These patients had some of the most intractable behavior that required a multi-modal treatment approach. To fully evaluate the

TIRF and whether it is a sensitive measure of treatment complexity, additional evaluations with other programs, particularly those serving outpatients or providing school consultation, that serve individuals with less severe challenging behaviors is clearly needed. One might expect significantly lower scores on the medication and protective equipment subscales, but it is unclear how the number and type of treatment components with the behavior plans might change with a group of individuals with less severe behavior. It may be that complex treatments with multiple components are needed to address all factors related to challenging behavior.

The TIRF was developed as a brief measure to objectively quantify the intensity of our treatment plans for patients when they are discharged from our unit. It is difficult to compare to other programs without recognizing that these patients were referred from across the country to participate in our program. Our model of treatment is also unusual in that it is a multidisciplinary approach combining medication evaluation and applied behavior analysis using data-based decision making. This is not possible in many outpatient programs as medical and behavioral treatment is often done independently.

Finally, the TIRF does not capture information on all the different interventions that were tried and failed. Although we did capture this information to some degree with the number and type of medication trials, there was not an item addressing the number of behavioral treatments tried. This could be added to the scale in future revisions. In addition, no data were collected on how the treatment plan was modified after discharge or how effective the plan continued to be once the patient was discharged. Although the TIRF is limited, and the group sampled may be limited, with further data collection and refinement of the scale, it may be a useful tool to capture the degree of complexity and the supports required to accommodate an individual with severely challenging behavior at a specific point in time. Future research might also include conducting the ratings before and after treatment, whether it is in a school, an outpatient program, or an inpatient program.

Acknowledgments

Manuscript preparation was supported by Grant R01 HD076653 from the Eunice K. Shriver National Institute of Child Health and Human Development (NICHD).

The authors would like to acknowledge Alyssa Fisher, Nabil Mezhoudi, Stephanie Gerow, Cara Phillips, and Shari Pincus for their assistance in developing this measure and for data collection.

References

- Barnato AE, Mclellan MB, Kagay CR, Garber AM. Trends in inpatient treatment intensity among Medicare beneficiaries at the end of life. *Health Services Research*. 2004; 39:363–376. [PubMed: 15032959]
- Engert A, Plütschow A, Eich HT, Lohri A, Dörken B, Borchmann P, ... Diehl V. Reduced treatment intensity in patients with early-stage Hodgkin's lymphoma. *New England Journal of Medicine*. 2010; 363:640–652. [PubMed: 20818855]
- Goode AD, Winkler EAH, Lawler SP, Reeves MM, Owen N, Eakin EG. A telephone-delivered physical activity and dietary intervention for type 2 diabetes and hypertension: does intervention dose influence outcomes? *American Journal of Health Promotion*. 2011; 25(4):257–263. [PubMed: 21361811]

- Hagopian, LP.; Anderson, ME. Integrating behavioral and pharmacological interventions for severe problem behavior displayed by children with neurogenetic and developmental disorders. In: Shapiro, BK.; Accardo, PJ., editors. Neurobehavioral disorders: Behavioral issues and their treatment. Brookes; New York: 2010. p. 219-239.
- Hagopian LP, Rooker GW, Bonner A, Arevalo AR, Zarcone JR. Further analysis of subtypes of automatically reinforced self-injurious behavior: A replication and quantitative analysis of published data. *Journal of Applied Behavior Analysis*. in press.
- Hagopian LP, Rooker GW, Zarcone JR. Delineating subtypes of self injurious behavior maintained by automatic reinforcement. *Journal of Applied Behavior Analysis*. 2015; 48:523–543. [PubMed: 26223959]
- Iwata BA, Dorsey MF, Slifer KJ, Bauman KE, Richman GS. Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis*. 1994; 27:197–209. [PubMed: 8063622]
- Lovaas OI. Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology*. 1987; 55:3–9. [PubMed: 3571656]
- Peters-Scheffer N, Didden R, Korzilius H, Sturmey P. A meta-analytic study on the effectiveness of comprehensive ABA-based early intervention programs for children with autism spectrum disorders. *Research in Autism Spectrum Disorders*. 2011; 5:60–69.
- Thorpe KE, Howard DH. The rise in spending among Medicare beneficiaries: The role of chronic disease prevalence and changes in treatment intensity. *Health Affairs*. 2006; 25:w378–w388. [PubMed: 16926180]
- Warren SF, Fey ME, Yoder PJ. Differential treatment intensity research: A missing link to creating optimally effective communication interventions. *Mental Retardation and Developmental Disabilities Research Reviews*. 2007; 13:70–77. [PubMed: 17326112]

Appendix 1

Child's Name:

Rater:

Date:

Treatment Intensity Rating Form

Instructions: Please rate the individual's current pharmacological and behavioral intervention in each specific area described below.

I. PHARMACOLOGICAL INTERVENTION

(All questions refer to medications used to target problem behavior in psychiatry discharge summary)

1. Rating	Number of Medication Trials Include preadmission, addition, or discontinuation of any medication to target behavior. Do not include medications to address medical conditions (e.g., Synthroid), illness (e.g., antibiotics, allergy medication), or sleep medications. Include anticonvulsant medication.
U	Unknown.
0	No medication trials.
1	1–2 medication trials.
2	3–5 medication trials.

3	6 medication trials.
4	7 medication trials.
5	8 or more medication trials.

2. Rating	Number of Medications at Discharge Medications that target behavior only. Do not include medications to address medical conditions (e.g., Synthroid), illness (e.g., antibiotics, allergy medication), or sleep medication. Include anticonvulsant medication.
U	Unknown.
0	Child does not take medications targeting behavior.
1	Child is taking one medication.
2	Child is taking 2 medications.
3	Child is taking 3 medications.
4	Child is taking 4 medications.
5	Child is taking more than 4 medications.

II. BEHAVIOR SUPPORTS

(All questions refer to child's final/current behavior plan)

3. Rating	Behavioral Intervention Plan Total Components (e.g., prompting procedures, extinction/planned ignoring, and protective equipment).
U	Unknown.
0	Child needs limited behavioral intervention (e.g., environmental modifications)
1	Child needs moderate behavioral intervention requiring 1–2 components
2	Child needs significant behavioral intervention requiring 3–5 components.
3	Child needs significant behavioral intervention requiring 6 components.
4	Child needs significant behavioral intervention requiring 7 components.
5	Child needs significant behavioral intervention requiring 8 or more components.

4. Rating	Proactive Procedures Includes prompting procedures and other antecedent manipulations. Do not include protective equipment or mechanical restraint.
U	Unknown.
0	Child needs limited proactive intervention requiring 1 component (e.g., environmental modifications)
1	Child needs moderate proactive intervention requiring 2 components (e.g., environmental modifications and a prompting procedure)
2	Child needs significant proactive intervention requiring 3–5 components.
3	Child needs significant proactive intervention requiring 6 components.
4	Child needs significant proactive intervention requiring 7 components.

5	Child needs significant proactive intervention requiring 8 or more components.
---	--

5. Rating	Reactive Procedures Includes intervention components that are in response to behavior (e.g., response cost, response reduction procedures).
U	Unknown.
0	Child does not need reactive procedures.
1	Extinction or planned ignoring is included in the behavior plan.
2	Child requires response cost or redirection.
3	Child requires brief response blocking or brief, non-exclusionary timeout.
4	Child requires brief, contingent physical holds (e.g., baskethold/facial screen) or exclusionary timeouts that do not last more than 1 minute.
5	Child requires contingent physical holds (e.g., baskethold/facial screen) or exclusionary timeouts that last for more than 1 minute but less than 3 minutes (after 3 minutes, hold should be considered a crisis management procedure).

6. Rating	Crisis Management Procedures Intervention components designed to maintain safety during severe behavioral outbursts of SIB and/or aggression (e.g., holds, time outs, room clear outs).
U	Unknown.
0	Child does not need a crisis management procedure.
1	Child may require minor environmental modifications to maintain safety during outbursts (e.g., moving to a mat, padded area, or safe area).
2	Child requires minor environmental modifications in addition to response blocking (e.g., move to a padded mat and block SIB).
3	Child requires an exclusionary timeout or a room clear out until calm.
4	Child requires multiple staff for holds (e.g., stabilization holds, supine holds, 5-man holds) during outbursts.
5	Child requires a PRN medication (e.g., an IM injection) during outbursts.

7. Rating	Staffing Levels Required to Implement Treatment/Maintain Safety Staffing levels required at discharge.
U	Unknown.
0	1:1 all waking hours across contexts.
1	1:1 all waking hours/2:1 in one other context (e.g., academics, bathroom, meals).
2	1:1 during all waking hours/2:1 during outbursts to maintain safety.
3	2:1 all waking hours.
4	2:1 during all waking hours/3-4:1 during outbursts to maintain safety.
5	2:1 all waking hours/more than 4:1 during outbursts to maintain safety.

III. PROTECTIVE EQUIPMENT

(All questions refer to child's final/current behavior plan)

8. Rating	Protective Equipment-Child Designed to protect an area of the body from injury, but not to restrict movement (e.g., helmet, arm guards, knee pads, Kevlar sleeves).
U	Unknown.
0	Child does not wear protective equipment.
1	Child requires a minor covering/clothing/dressing for wound protection.
2	Child wears 1 form of protective equipment intermittently (e.g., during bursts or as an interruption procedure).
3	Child wears 1 form of protective equipment during all waking hours or more than one form of protective equipment intermittently.
4	Child wears 2 forms of protective equipment during all waking hours.
5	Child wears more than 2 forms of protective equipment during all hours.

9. Rating	Protective Equipment-Staff Designed to protect staff from injury (e.g., helmet, arm guards, Kevlar sleeves).
U	Unknown.
0	Staff do not wear protective equipment when working with child.
1	Staff wear protective clothing only (e.g., hats, jackets, gowns) but no other forms of protective equipment when working with child.
2	Staff wear 1 form of protective equipment when working with child.
3	Staff wear 2 forms of protective equipment when working with child.
4	Staff wear 3 forms of protective equipment when working with child.
5	Staff wear more than 3 forms of protective equipment when working with child.

10. Rating	Mechanical Restraint Designed to prevent self-injury by limiting movement (e.g., arm splints, Posey mitts).
U	Unknown.
0	Child does not wear mechanical restraints.
1	Child wears 1 form of mechanical restraint intermittently (e.g., during bursts or as an interruption procedure) or empty sleeves during all waking hours.
2	Child wears 2 mechanical restraints intermittently (e.g., during bursts or as an interruption procedure).
3	Child wears 1 mechanical restraint during waking hours.
4	Child wears 2 mechanical restraints during waking hours.
5	Child wears at least one mechanical restraint at all times (including in bed).

Appendix B. TIRF Scoring Criteria

All 10 questions on the TIRF were scored using a 0–5 rating scale (0 lowest, 5 highest). If for any item, the information could not be obtained, the item was scored as unknown. All information was collected from documentation used by the inpatient unit to monitor and summarize a patient's admission (i.e., psychiatric discharge summaries, behavior plans, behavioral discharge summaries, and specific unit protocols put in place to ensure the safety of both the staff and the patient).

Question 1: This question referred to the number of medication trials conducted during a patient's admission. Because many patients received multiple medication trials, a range of medications was necessarily at each of the 5 levels. Medications that targeted medical conditions, illness, or sleep, were not included in this count. That is, only psychotropic medications that potentially targeted challenging behavior were evaluated. The number medications at admission were counted and additional trials were counted if: a) an admission medication was discontinued, b) a medication was added, but was then discontinued prior to discharge, d) a new medication was added and continued until their date of discharge. All medications were counted at most, only once regardless of changes in dosage or the number of times it was discontinued and subsequently re-added.

Question 2: This question was designed to quantify the number of medications the patient was receiving at discharge. Medications that targeted medical conditions, illness, or sleep, were not included in this count. A medication was scored as being present at discharge if the date of the patient's discharge matched the date that the medication was discontinued.

Questions 3: This question referred the total number of behavioral components included in the patient's final treatment package at discharge. A behavior component was defined as any proactive or reactive procedure put in place to either support the emergence of appropriate behavior, or reduce the likelihood that inappropriate behavior will continue to occur. A behavioral component was also defined as any protocol put in place to keep the patient safe (i.e., protective equipment or mechanical restraint). Some examples of behavior components include functional communication training, non-contingent reinforcement, any differential reinforcement procedure, timeout, or a protective helmet.

Question 4: This question is designed to quantify the number of proactive procedures included in a patient's treatment package at their time of discharge. A proactive procedure was defined as any procedure that is implemented throughout the patient's day, either across all settings, or in a specific context (e.g., academics) that supports the emergence and maintenance of appropriate behavior. In this way, the implementation of a proactive procedure is not contingent on the occurrence of inappropriate behavior, rather, it is implemented throughout an activity regardless of ongoing behavior. Often, proactive procedures arrange for, or signal the delivery of a possible reinforcer contingent on appropriate behavior.

Questions 5: This question was designed to quantify the complexity of the reactive procedures included in a patient's treatment package. A reactive procedure was defined as any procedure whose implementation depends wholly on the occurrence of inappropriate

behavior. That is, if inappropriate behavior were never to occur, reactive procedures would never be implemented. Some examples of reactive procedures include, extinction, response cost, response blocking or timeout. Protective equipment and mechanical restraint were not included in this question, even if they were applied contingent on the occurrence of problem behavior as they are captured separately in subsequent questions.

Question 6: This question referred to what type of crisis management procedures were required to maintain staff and patient safety during large bouts of high intensity and high frequency problem behavior.

Question 7: This question quantified the staffing levels required to both maintain the patient's safety.

Question 8: This question reflected the amount of protective equipment worn by the patient and the frequency of which it was worn. Examples of protective equipment include padded helmets, wrist guards, and knee pads.

Question 9: This question quantified the amount of protective equipment worn by staff members while working with a particular to maintain staff safety.

Question 10: This question quantified the amount of mechanical restraint worn by patients and the frequency of which the person wore it. Mechanical restraint can restrict the patient's ability to engage in dangerous behaviors such as self-injury. An example of mechanical restraints are arm splints.

Figure 1.
The percentage of medications (by category) prescribed for participants at discharge for each study group.

Demographic information for participants delineated by function, gender, age, and percentage reduction in challenging behavior at the end of treatment.

Table 1

Demographics										
Variable	Gender				Age				Percentage Reduction	
	Male		Female		min		max		n	mean
	n	%	n	%						
Social	37	21	56.76	16	43.24	5	22	13.28	37	83.41
Automatic	38	30	78.95	8	21.05	8	22	14.50	37	81.31
Total	75	51	68.00	24	32.00	5	22	13.90	74	82.37

Table 2

Mean TIRF item score and standard deviation across study groups.

Variable	TIRF Score Variables Across Groups					
	Social (N = 37)			Automatic (N = 38)		
	n	mean	SD	n	mean	SD
I. Pharmacological Intervention						
1: Number of Medication Trials	37	2.27	0.96	38	3.39	1.39
2: Number of Medications at D/C	37	2.84	1.17	38	3.08	1.28
II. Behavior Supports						
3: Behavioral Intervention Plan	37	4.49	0.87	38	4.76	0.59
4: Proactive Procedures	37	3.00	1.08	38	3.11	1.20
5: Reactive Procedures	37	3.00	1.15	38	2.89	0.89
6: Crisis Management Procedures	33	1.58	1.00	32	1.72	0.77
7: Staffing Levels Required	35	0.43	1.04	35	0.43	0.92
III. Protective Equipment						
8: Protective Equipment-Child	37	0.43	0.96	37	1.70	1.66
9: Protective Equipment-Staff	24	1.92	1.44	25	2.04	1.57
10: Mechanical Restraint	37	0.11	0.52	38	1.34	1.28

* significant

SD = standard deviation