The Impact of Behaviour Skills Training on the Knowledge, Skills, and Well-being of Frontline Staff in the Intellectual Disability Sector: A Clustered Randomised Control

Laura Gormley, PhD., BCBA;
Trinity College Dublin & Rehab Group

Olive Healy, PhD., BCBA-D;
Trinity College Dublin

Brona O' Sullivan;
Trinity College Dublin

Darragh O' Regan, BCBA;
Rehab Group

Maeve Bracken, PhD., BCBA-D;
Trinity College Dublin

Corresponding author: Olive Healy, PhD., BCBA-D

School of Psychology, Trinity College Dublin, College Green, Dublin 2, Ireland

olive.healy@tcd.ie

+353 1 8961175

Full trial protocol available from authors on request

Acknowledgements: This research was conducted at Trinity College Dublin and was supported by the Irish Research Council's Employment Based Postgraduate Programme 2015. The authors would like to thank the participants for their role in this study. We would also like to acknowledge the contribution of the Rehab Group behavioural team for their part in delivering the intervention and collecting data.

Conflict of interest: Laura Gormley and Darragh O'Regan were employed by the host organisation at the time of the study.

Abstract

Background

Staff with varying backgrounds and educational qualifications can be effectively trained to implement procedures in line with evidence-based practice (EBP) (Allen and Tynan 2000, Dench 2005, Grey and Hastings 2005). Behaviour Skills Training (BST) is a competency-based training model, used to effectively educate a broad selection of professionals, including frontline staff, in a range of work-related skills (Maffei-Almodovar and Sturmey 2018). However, BST has yet to be evaluated in a large group-based experiment.

Method

This study involved a parallel cluster randomised control trial (RCT). Six service sites, with a total of 54 participants, were assigned to the intervention condition, which used BST to coach intellectual disability staff in reinforcement, systematic prompting, functional communication training (FCT), and task analysis. Fifty participants, across six service sites, were assigned to a control condition, in which generalised training in behavioural interventions was restricted.

Results

Participants in the intervention condition demonstrated statistically significant improvements in their knowledge scores over the study period. Participants in the control condition showed no change or a statistically significant decrease in their knowledge scores. No statistically significant changes to well-being were observed for either group. There was clear evidence of knowledge maintenance, as well as skill acquisition and subsequent generalisation to the workplace environment, among participants in the intervention condition. Participants also evaluated the BST intervention positively.

Conclusions

Results support BST as a method for disseminating EBP to frontline staff working with adults with intellectual and developmental disabilities.

Keywords: Behaviour Skills Training, staff, knowledge, well-being, clustered RCT

Introduction

Over recent years, the importance of embedding EBP within the service provision framework for people with intellectual disabilities has been repeatedly emphasised (National Autism Center 2015, National Institute for Health and Care Excellence 2015, Wong et al. 2014). However, while the evidence base for effective behavioural interventions is well-established (National Autism Center 2015, Sturmey and Didden 2014, Wong et al. 2014), the literature reports a significant disconnect between EBP research and the competencies displayed by frontline staff in applied settings (Campbell 2010, Rapp et al. 2010).

A sizeable portion of frontline staff, supporting people with intellectual disabilities, may not be considered adequately qualified (Campbell 2010, Joint Committee on Social Care Professionals 2002). Even among staff with relevant professional qualifications, the skills required to effectively support this population are typically underdeveloped or absent at the time of qualification (Campbell 2010). However, research shows that frontline staff, with varying backgrounds and educational qualifications, can be upskilled and supported to deliver EBP (Dench 2005, Grey and Hastings 2005, Maffei-Almodovar and Sturmey 2018).

BST is a type of competency-based training that employs instruction, modelling, rehearsal, and feedback (Sarokoff and Sturmey 2004), and has been repeatedly and successfully employed to educate a broad selection of professionals, including frontline staff, in a diverse range of work-related skills to support people with intellectual and developmental disabilities (e.g., Homlitas et al. 2014, Maffei-Almodovar and Sturmey 2018, Parsons et al. 2012). There is also consistent support for the generalisation of the newly learned skills to novel environments, without additional training (e.g., Homlitas et al. 2014, Maggin et al. 2012, Sarokoff and Sturmey 2008).

However, this training model has not yet been evaluated in a large group-based experimental design. Single-subject experimental design has typically been the approach used to test the BST approach. However, the discipline of psychology, more generally, typically relies on group-based designs to analyse the impact of interventions (Kazdin 2011). Therefore, in order to demonstrate the effectiveness of BST to the mainstream psychology community, an RCT evaluation was deemed necessary.

The current study conducted a parallel cluster RCT to evaluate BST, in terms of its effectiveness in disseminating target knowledge and improving psychological well-being among participants. This approach was taken because the intervention was “naturally

applicable to the cluster" (Wears 2002, p. 330); training discrete service sites was more efficient than training individual staff members (Wears 2002), and active engagement was considered more likely (Donner and Klar 2004). Additional analyses provided a closer examination of BST, in terms of knowledge maintenance, skill acquisition and generalisation, and acceptability among participants in the intervention condition. Such analyses can be valuable for informing staff training within clinical practice. However, the logistical constraints controlling this research programme, which included resource allocation and service closures within the host organisation, precluded an evaluation of knowledge maintenance and skill acquisition among participants in the control condition.

Method

Participants, Randomisation, and Setting

Participants were 104 employed frontline staff recruited from 12 service sites within a single intellectual disability service provider organisation. All employed frontline staff in recruited service sites were included. A priori sample size calculation was based on the large effect sizes ($d > 1.0$) obtained for the primary outcome measure (i.e., knowledge acquisition) in a feasibility study. A minimum of 60 participants (i.e., five participants in each service site) were required to demonstrate a large effect size ($d \geq 0.08$) when the p-value was set at 0.05 (Clinical and Transactional Science Institute (UCSF) 2017). Therefore, only those service sites with at least five employed staff were included. Research standards also stipulate that retention rates should reach at least 70-80% in each experimental arm (e.g., Lyles et al. 2007, Whitlock et al. 2004). Therefore, it was considered prudent to recruit at least 100 participants to allow for 20-30% attrition.

Recruited service sites were randomly assigned to the intervention condition (N=6, n=54) or the control condition (N=6, n=50) at one point in time; immediately after recruitment and before baseline testing took place. Allocations were stratified by service type (residential or day) and geographical region. One member of the research team allocated service sites using the 'coin toss' method and another member, blind to the allocations, decided which experimental arm would receive the intervention and which would be designated as control. It was not possible to mask the intervention from participants but they were recruited prior to randomisation.

Within experimental conditions, four service sites provided residential support, while two offered day-based support. Eighty-two participants were female (n=41 intervention;

n=41 control) and 77 participants listed their job title as careworker (n=27 intervention; n=27 control) or programme facilitator (n=10 intervention; n=13 control). Forty-three participants had an undergraduate degree (n=23 intervention; n=20 control) and 18 had earned a postgraduate qualification (n=9 intervention; n=9 control). Overall, 65 participants had been employed with the host organisation for 5 years or less (n=30 intervention; n=35 control). A total of 20 participants (n=14 intervention; n=6 control) reported an employment period of 5-10 years with the host organisation and 19 participants (n=10 intervention; n=9 control) reported an employment period of 10+ years.

There was no significant relationship between experimental condition and participant gender (χ^2 (1, n = 104) = 0.57, p = .449), level of education (i.e., pre-undergraduate degree & other; undergraduate degree; postgraduate qualification) (χ^2 (2, n = 102) = 0.08, p = .963) or length of service within the host organisation (χ^2 (3, n = 104) = 4.91, p = .179). Four behaviour therapists, with postgraduate qualifications in Applied Behaviour Analysis or a related discipline, delivered the training intervention.

Informed consent was obtained from all participants in the study after randomisation and all procedures conducted in the study were in accordance with the Research and Ethics Policy of the host service provider, university and the 1964 Helsinki declaration and its later amendments. Outcome measures were administered in a classroom or office setting.

Outcome Measures

Primary outcomes. ***Knowledge outcomes.*** Knowledge assessments for each training module were developed using the National Professional Development Centre (NPDC) online manuals and associated quizzes (AFIRM Team 2015a, AFIRM Team 2015b, AFIRM Team 2015c, Franzone 2009). Each knowledge assessment contained 10 multiple-choice items, with each correct item receiving a score of '1'; the maximum score on the measure was 10. An analysis of edumetric validity and reliability (Carver 1974) was conducted and found to be sufficient. In addition, the Test of Knowledge, which evaluates general knowledge of behavioural principles and applications, was employed. It contains 20 short answer and multiple-choice questions. Each item answered correctly is scored as a '1'; the maximum score on the measure is 20. The construct validity of this measure was tested and found to be adequate (Denne et al. 2015).

Skill outcomes. Task analyses were developed to evaluate the implementation fidelity of target skills related to each training module. Task analyses were based on the

NPDC guidelines for each target skill (AFIRM Team 2015a, AFIRM Team 2015b, AFIRM Team 2015c, Franzone 2009). Task analyses were divided into 2-3 sections, which listed steps that participants were expected to perform. The percentage of steps correctly performed was calculated for each section. The average percentage of steps performed correctly for each target skill was then determined.

Secondary outcomes. ***Psychological outcomes.*** The battery of measures used to assess psychological well-being were (1) Maslach Burnout Inventory – Human Services Survey (MBI-HSS; Maslach et al. 1996); (2) Minnesota Satisfaction Questionnaire – Short Form (MSQ-SF; Weiss et al. 1967); (3) Short Version of the Occupational Self-Efficacy Scale (Rigotti et al. 2008), and (4) The Shortened Ways of Coping Questionnaire (SWC-R; Hatton and Emerson 1995). The psychometric properties of all measures in this battery have been adequately demonstrated in the studies cited.

Other outcomes. The Training Acceptability Survey (Underwood et al. 2002), a six-item respondent-based measure, was used to examine the training intervention, in terms of relevance, planning, opportunities for participation, resource use, areas of difficulty, providing motivation to learn more, and overall usefulness. The Cronbach alpha coefficient was .88 for the reinforcement module, .87 for the systematic prompting module, .91 for the FCT module, and .90 for the task analysis module.

The Attitudes to Evidence-Based Practice Questionnaire, a 26-item respondent-based rating scale that examines barriers to the use of EBPs in care settings, was also employed. This five-point Likert-scale has demonstrated sufficient internal consistency, in addition to content and face validity (McKenna et al. 2004).

The Perceptions of Supervisory Support Scale (PSS; Fukui et al. 2014), a 19-item respondent-based rating scale, evaluates perceptions of supervisory support that staff receive in the workplace. The psychometric properties of this measure, including reliability and validity, have be adequately demonstrated (Fukui et al. 2014).

Design

A parallel cluster RCT, with a control condition, was used to evaluate the impact of the training intervention on participant knowledge and well-being. A within-subjects, pre-post-test design was used to evaluate knowledge maintenance among participants in the intervention condition. Observations to criterion, a type of event recording (Cooper et al.

2014), was also employed to measure the number of practice opportunities required by participants in the intervention condition, to reach a pre-determined level of target skill proficiency, in the classroom and typical work settings.

Procedure

Figure 1 provides a timeline of study implementation. All participants completed the Attitudes to Evidence-Based Practice Questionnaire, the knowledge outcome measures, and the psychological outcome measures at T1 (baseline) and T3 (4-weeks following the delivery of the training to the intervention condition). The PPS was completed by all participants at T3. Immediately following each training module (T2), participants in the intervention condition completed the associated knowledge assessment and training acceptability survey. After the final module (T2), the intervention group also completed the Attitudes to Evidence-Based Practice Questionnaire and the Test of Knowledge.

The training intervention commenced immediately after baseline testing (T1) and participants in the control condition were precluded from receiving generalised training on behavioural interventions or other approaches used to support people with intellectual disabilities for the duration of the study (4 months). They were not restricted from receiving training as usual according to organisational policy (e.g., instruction on the implementation of specific strategies to support individual service users). Data was not collected on the training as usual provided during the study.

Intervention. The BST intervention was delivered over 3 consecutive days (20 hours in total) via group format in a classroom setting. Each service site was trained separately. The four training modules were: (1) reinforcement; (2) systematic prompting; (3) FCT, and (4) task analysis. Training modules comprised didactic instruction via PowerPoint presentation and an accompanying manual for participants, general case video models with embedded instruction, rehearsal sessions with service user confederates, and individual feedback from the trainer. Mastery criterion for target skills was 90% correct implementation across three consecutive observations. As part of the FCT module, participants were supported by the trainer to develop communication plans for a service user that they were directly working with. Trainers observed the cohort of participants who they had trained, implementing FCT skills, with these service users, in the typical work environment, during the 4-weeks following the classroom-based training intervention.

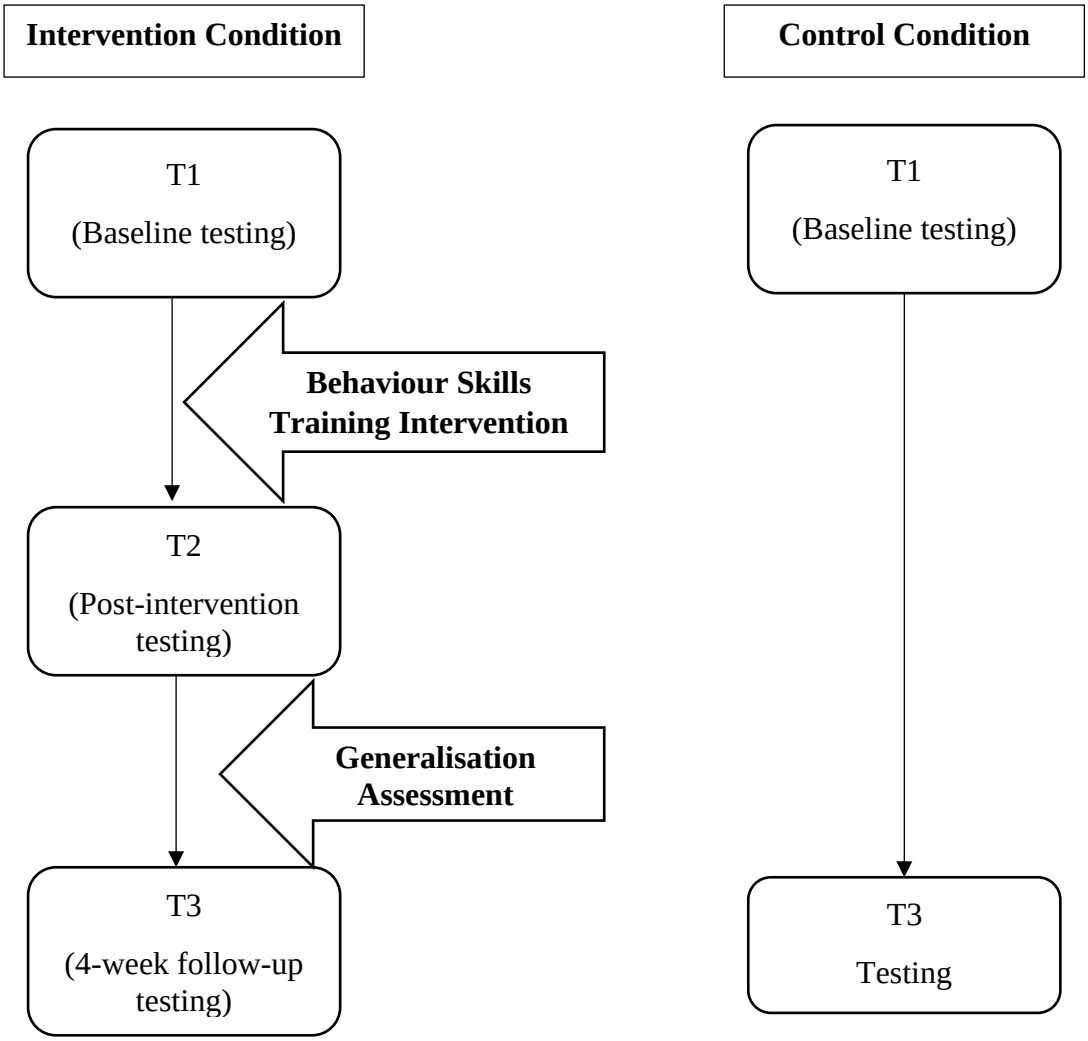

Figure 1. Flow diagram showing the timeline of study implementation.

Procedural Fidelity

Procedural fidelity data were collected for all intervention sessions and measured the trainers' accurate implementation of the training programme, including the rehearsal and feedback sessions. Assessment material is available from the authors on request. The BST intervention was implemented with 100% fidelity by Trainers 1, 2 and 4 and 99.4% by Trainer 3.

Interobserver Agreement

Interobserver agreement (IOA) was collected for 79.2%, 75.5%, 81.5%, and 90.2% of classroom-based rehearsal sessions for the reinforcement, systematic prompting, FCT, and task analysis modules, respectively. The trainer and a second observer independently recorded the accuracy of each participant's implementation of the target skills for all training modules. A point-by-point method was used to calculate IOA and the number of agreements was divided by the number of agreements and disagreements and multiplied by 100. The mean percentage of agreement ranged from 98.7% to 99.3%: reinforcement module (98.7%); systematic prompting module (98.9%); FCT module (access to items: 98.7%; access to breaks: 99.3%), and task analysis module (98.7%). A secondary analysis using Kappa (Cohen 1960) showed "almost perfect" agreement among raters for the four modules.

Data Analysis

Attrition. Eight participants (14.8%) from the intervention condition and four participants (8.9%) from the control condition did not complete the study; a differential attrition rate of 5.9% between the experimental arms. The primary reasons for attrition were absence due to illness or holiday leave. Figure 2 provides a diagram of participant flow during the RCT.

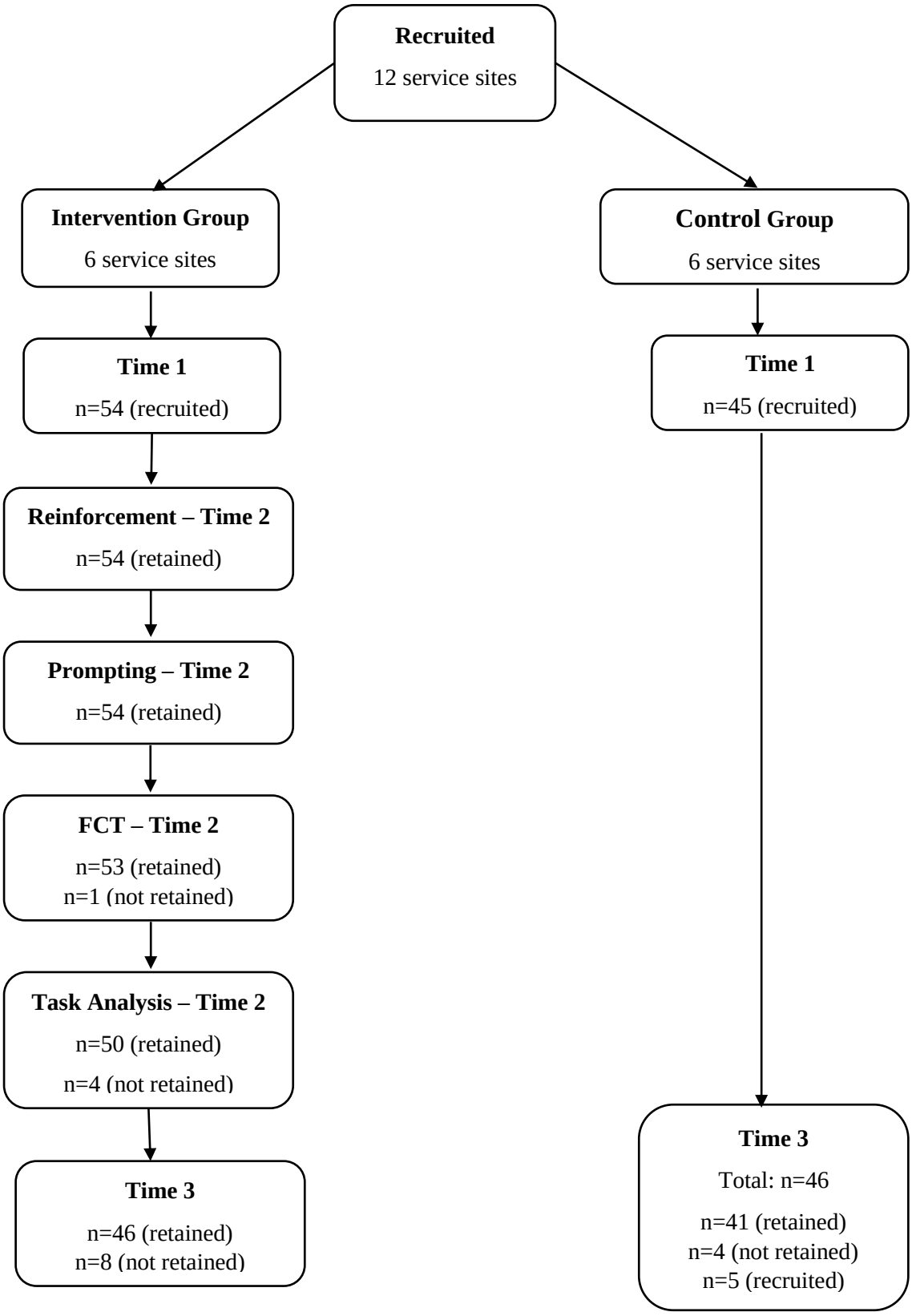

Figure 2. Flow diagram showing participant attrition during the RCT.

Statistical analysis. General linear mixed-effects modelling (GLMM) was used to investigate the impact of the intervention on knowledge and psychological outcomes. GLMM is extremely robust, accounting for correlated data (including that associated with repeated measures), missing data, and the clustering of data within contextual variables (e.g., service sites; Field 2016, Larsen et al. 2010, Seltman 2015). The following random intercept models were tested for goodness of fit: random intercepts across participants; random intercepts across service sites, and random intercepts across participants and across service sites. The analyses are available from the authors on request. The random intercepts across participants model produced the most parsimonious solution in all cases (ICC^2 ranged from 0.39 to 0.68). The fixed parameters tested in all models that compared the experimental conditions were: time of measurement; experimental condition; interaction between time and experimental condition. A fixed parameters GLMM model was used to quantify the relationship between perception of supervisory support and knowledge outcomes for both experimental groups; ICC^2 calculations were precluded because analyses showed that the random intercept covariance parameters were redundant.

GLMM was used to evaluate the impact of the intervention on knowledge outcomes at T1, T2, and T3 for participants in the intervention condition. The random intercepts across participants model produced the most parsimonious solution in all cases (ICC^2 ranged from 0.21 to 0.26) and the fixed parameter tested was time of measurement. Finally, GLMM was employed to examine the relationship between knowledge outcomes for participants in the intervention condition and education level; length of service, and prior exposure to the training content. A random intercepts across participants model was applied (ICC^2 ranged from 0.15 to 0.27), with the fixed parameters as follows: time of measurement; length of service/prior exposure to training content/education level, and the relevant interaction.

Results

Clustered RCT

Training history. At baseline, 75.9%, 87%, 85.2%, and 79.6% of intervention participants reported no prior training in reinforcement, systematic prompting, FCT and task analysis, respectively. At baseline, 48.9%, 71.1%, 68.9%, and 66.7% of control participants reported no prior training in reinforcement, systematic prompting, FCT and task analysis, respectively.

Knowledge outcomes. Compared to participants in the control condition, participants in the intervention condition demonstrated a significantly greater increase in target knowledge from T1 to T3. Time of measurement, experimental condition, and the interaction between time of measurement and experimental condition all significantly predicted participant scores on all knowledge outcome measures (Table 1).

Table 1

A summary of the GLMMs knowledge outcome measures (N=12, n=54), the pairwise comparisons of the experimental conditions at T1 and at T3 across the knowledge outcomes, and the pairwise comparisons of changes in knowledge outcomes from T1 to T3 for both experimental conditions

Assessment		**F**	**Numerator Df**	**Denomenator Df**	***P***		**t**	**M**	**SE**	***P***		**t**	**M**	**SE**	***P***
Reinforcement	***Time***	20.93	1	96.08	.00**	***T1***	0.16	0.63	3.85	.87	***Intervention***	11.29	28.01	2.48	.00**
	Condition	30.42	1	104.15	.00**	***T3***	8.84	35.39	4.00	.00**	***Control***	-2.13	-7.87	3.69	.04*
	Time X Condition	66.52	1	96.08	.00**										
Systematic prompting	***Time***	17.42	1	90.80	.00**	***T1***	2.22	7.89	3.55	.03*	***Intervention***	7.04	20.62	2.93	.00**
	Condition	47.15	1	100.23	.00**	***T3***	9.20	33.47	3.64	.00**	***Control***	-1.71	-4.22	2.47	.10
	Time X Condition	41.49	1	90.80	.00**										
FCT	***Time***	19.90	1	90.49	.00**	***T1***	0.25	1.37	5.54	.81	***Intervention***	9.19	35.93	3.91	.00**
	Condition	37.44	1	99.54	.00**	***T3***	11.07	49.23	4.45	.00**	***Control***	-2.61	-11.49	4.39	.01*
	Time X Condition	74.50	1	90.49	.00**										
Task analysis	***Time***	7.86	1	87.12	.006*	***T1***	1.02	6.00	5.90	.31	***Intervention***	6.15	24.46	3.98	.00**
	Condition	24.32	1	97.84	.00**	***T3***	8.28	40.93	4.94	.00**	***Control***	-2.04	-8.11	3.97	.047*
	Time X Condition	33.91	1	87.12	.00**										

Test of Knowledge	***Time***	22.35	1	96.02	.00**	***T1***	1.72	3.33	1.94	.09	***Intervention***	7.42	12.99	1.75	.00**
	Condition	37.18	1	104.79	.00**	***T3***	7.49	18.33	2.45	.00**	***Control***	-1.02	-1.61	1.58	.31
	Time X Condition	38.31	1	96.02	.00**										

*p<.05; ** p < .001 N = number of clusters; n = number of participants

Pairwise comparisons, which relied on estimated marginal means and a Bonferroni adjustment[1] based on two comparisons ($p<.025$), showed no significant difference between the two experimental conditions across all knowledge outcome measures, except the systematic prompting assessment, at T1 (Table 1). A second set of pairwise comparisons, showed that participant scores on all knowledge outcome measures increased significantly from T1 to T3 for the intervention condition but did not change or decreased significantly for the control condition (Table 1).

Psychological outcomes. Time of measurement, experimental condition, and the interaction between time of measurement and experimental condition did not predict scores on any psychological outcome measures.

Other outcomes. For both experimental conditions, participants' scores on the PSS did not demonstrate a significant relationship with knowledge scores. Attitudes towards EBP were generally positive for both experimental conditions. The majority of participants agreed or strongly agreed that there were advantages to implementing EBP and changing their approach according to recommendations from high quality research. However, at T3, a larger proportion of the control condition agreed or strongly agreed that research findings do not easily transfer into their own work. Furthermore, the control condition consistently reported time and cost restraints as barriers to implementing EBP in their workplace. The intervention group, on the other hand, were not so decisive on these issues at either T1 or T3.

Within-Subjects Analysis for Intervention Group

Knowledge outcomes. Time of measurement significantly predicted knowledge scores for participants in the intervention condition (Table 2). Pairwise comparisons, which relied on estimated marginal means and used a Bonferroni adjustment[1] based on three comparisons ($p<.017$), showed a significant increase in scores from T1 to T2 across all knowledge measures, and no significant change from T2 to T3 for all knowedge outcomes, except the systematic prompting assessment. Participant scores on the systematic prompting assessment decreased significantly from T2 to T3 (Table 2).

Length of service and prior exposure to training content did not signifcantly predict participant scores on the knowledge outcome measures. Level of education predicted scores on the task analysis assessment only ($F(2, 48.20) = 5.79$, $p = .006$); participants without an undergraduate degree consistently scored lower than those with an undergraduate (M = -

12.09, $p = .010$) or postgraduate degree ($M = -13.00$, $p = .049$). There was no interaction between time of measurement and education level across any of the knowledge outcomes.

Table 2

A summary of the GLMMs knowledge outcome measures (N=6, n=54) and the pairwise comparisons of changes in knowledge outcomes from T1-T2 and T2-T3 for the intervention condition (N=6, n=54)

Assessment		**F**	**Numerator Df**	**Denomenator Df**	***P***		**M**	**SE**	***P***
Reinforcement	***Time***	94.27	2	100.60	.00**	***T1-T2***	31.30	2.47	.00**
						T2-T3	-3.23	2.64	.67
Systematic prompting	***Time***	67.35	2	99.91	.00**	***T1-T2***	28.52	2.53	.00**
						T2-T3	-7.69	2.69	.02*
FCT	***Time***	97.12	2	97.12	.00**	***T1-T2***	41.22	3.18	.00**
						T2-T3	-5.27	3.38	.37
Task analysis	***Time***	40.97	2	91.91	.006*	***T1-T2***	29.54	3.50	.00**
						T2-T3	-4.70	3.71	.63
Test of Knowledge	***Time***	51.72	2	97.94	.00**	***T1-T2***	13.63	1.50	.00**
						T2-T3	-0.68	1.59	1.00

*p<.05; ** p < .001 N = number of clusters; n = number of participants

Skill outcomes. The average number of observations taken by participants in the intervention group to master the target skills during classroom-based rehearsal sessions was 3.2 (range = 3-6) for the positive reinforcement skill, 3.4 (range = 3-9) for least-to-most prompting, 3.9 (range = 3-6) for FCT to request access to items, 3.3 (range = 3-8) FCT to request access to a break, and 3.6 (range = 3-7) for the task analysis and backward chaining skill. On average, participants were implementing at least 87% of the steps correctly on the first rehearsal of each target skill and at least 97% of the steps correctly for the final rehearsal.

Data show that 30 (55.6%) participants had the opportunity to implement FCT with a service user in their typical work setting. Of these 30 participants, 25 (83.3%) displayed mastery on the first observation, three required two observations, one participant required three observations, and one participant did not reach mastery within the allocated three observations.

Social validity. Participants, in the intervention condition, evaluated the BST intervention positively. Scores on the training acceptability survey could range from '8' to '32'. The task analysis module (*M* = 28.8; range = 22-32) recorded the highest mean score, followed by the FCT module *(M* = 28.7; range=19-32), the systematic prompting module *(M* = 28.1; range=21-32), and the reinforcement module *(M* = 27.6; range=18-32).

Discussion

Results from the RCT confirmed that the BST intervention was effective in disseminating target knowledge. Knowledge scores for participants assigned to the intervention condition improved significantly between T1 and T3. The approximate average increase was 28% (T1:41%-T3:69%), 21% (T1:34%-T3:55%), 36% (T1:36%-T3:72%), and 25% (T1:51%-T3:76%) for the reinforcement, systematic prompting, FCT, and task analysis modules, respectively. In contrast, participants in the control condition showed either no change or a statistically significant decrease in knowledge scores over the study period.

A closer examination of the intervention group showed evidence of knowledge maintenance. There was a significant improvement in knowledge scores from T1 to T2, but no significant change from T2 to T3 on the Test of Knowledge and the reinforcement, FCT, and task analysis assessments. Participant scores on the systematic prompting assessment decreased significantly from T2 to T3. However, the scores at T2 and T3 still remained significantly higher than T1 scores, indicating that participant knowledge on this topic

improved as a result of the BST intervention. Subsequent analyses also revealed that regardless of educational level, length of service or prior training, participants benefited equally from the BST intervention. The only exception to this related to the task analysis knowledge scores; participants who had not yet achieved an undergraduate degree consistently scored lower than those with an undergraduate or postgraduate degree. However, there was no interaction between time of measurement and education level, indicating that that all educational groups were improving and maintaining knowledge at a comparable rate.

Results also indicated rapid levels of skill acquisition among participants in the intervention group. Prior to the first feedback session, average procedural integrity levels reached at least 87% across all target skills, despite 75% or more of participants stating that they had not received previous training in the target practices. Therefore, it would appear that the instruction and modelling components of the BST intervention played an important role in target skill development. Previously, feedback has been shown to be the most effective component of BST (Ward-Horner and Sturmey 2012). However, findings from the current study support the position that training outcomes are optimised when all four components are used in combination (Barnes et al. 2011, Homlitas et al. 2014).

Although the general training literature reports that successful post-intervention generalisation of trained skills to the workplace is difficult (Ager and O'May 2001, Campbell 2010) and documents "a staff culture frequently hostile to the structure and language of behavioural interventions" (Ager and O'May 2001, p. 252), findings from the present study produced evidence to the contrary. In total, 29 of the 30 participants who had the opportunity to implement FCT with a service user, achieved mastery within three observations and 25 of these participants displayed mastery on the first observation. Alongside these findings, participants generally reported positive attitudes towards EBP and consistently evaluated the BST intervention positively, in terms of relevancy and usefulness in the workplace. These findings would be strengthened by a concurrent analysis of service user behaviour and skill development and this approach is recommended for future research.

A key limitation of the current study was the inability to definitively conclude that the BST intervention had no impact on participants' psychological well-being. Service user skill acquisition has been positively correlated with staff well-being in previous research (Zaharia and Baumeister 1978). Therefore, it was hypothesised that effective staff training in

behavioural interventions, designed to promote skill acquisition among service users, would positively impact staff psychological well-being. However, the study did not objectively assess the intervention effects on service users, making it impossible to know if staff training resulted in skill acquisition among this group. Furthermore, skill development takes time and considering the complexity of the psychological constructs being measured (Fukui et al. 2014, Lazarus 1999, Locke 1969, Maslach 1993, Maslach 1999) and the sensitivity of the measures used, the follow-up period of 4 weeks may have been inadequate to facilitate meaningful change. Therefore, future studies must facilitate follow-up over a substantially longer period than 4 weeks, to adequately evaluate the impact of BST on relevant psychological constructs. An extended follow-up period would also increase confidence in the clinical value of the training package, both in terms of knowledge maintenance and accurate skill implementation in the workplace.

The sample size employed and variation within the data may also have precluded the detection of changes to participants' well-being. A priori estimations of sample size were based on the large effect sizes observed during a feasibility study and these effect sizes were exclusively related to participant performance on the knowledge measures. Furthermore, the feasibility study had failed to yield adequate data to calculate reliable effect sizes pertaining to the well-being variables. Nevertheless, given that the primary focus of the RCT was the dissemination of EBP and there is an ethical obligation to avoid resource wastage in applied research (Devane et al. 2004), the approach taken for sample size calculation was considered prudent.

Another key limitation relates to the evaluation of target skill acquisition over the course of the study. Observations to criterion, a type of event recording, was employed during training to measure the number of practice opportunities required by participants to reach a pre-determined level of proficiency in the implementation of skills related to each target EBP (Cooper et al. 2014). However, no baseline measures of procedural integrity were recorded. Therefore, although it was hypothesised that the instruction and modelling components played an important role in the high levels of procedural integrity observed during participants' first rehearsal of the target skills, the impact of a pre-existing skill-set cannot be ignored.

Notwithstanding these limitations, results clearly support BST as an effective method for disseminating knowledge relating to EBP, to frontline staff working with adults with

intellectual disabilities. Although it is the first time this has been shown through a large group-based experimental design, improved knowledge is not equivalent to improved procedural integrity (Grey et al. 2007). Therefore, while there is persuasive evidence of skill development among staff in the intervention group, the absence of a control comparison precludes more definite conclusions. This limitation should be addressed in future research. Likewise, the extent to which staff training can affect outcomes for people with intellectual disabilities is, as yet, unclear (van Oorsouw et al. 2009) and a recent meta-analysis (Maffei-Almodovar and Sturmey 2018) indicates that the number of BST research studies, which examine the impact of staff training on service user outcomes is relatively sparse. Therefore, future studies should examine the effects of BST on relevant service user outcomes.

Endnote

1. A Bonferroni adjustment minimises the chances of committing a Type 1 error with multiple comparisons. This involves setting a more stringent alpha level for each comparison by dividing the original alpha level ($p<.05$) by the number of comparison you intend to make.

2. Intracluster correlations (ICC) were calculated by comparing within and between cluster variance based on random intercept across participants mixed effect models.

References

AFIRM Team (2015a) Prompting. National Professional Development Center on Autism Spectrum Disorders, FPG Child Development Center, University of North Carolina, Chapel Hill, NC.

AFIRM Team (2015b) Reinforcement. National Professional Development Center on Autism Spectrum Disorders, FPG Child Development Center, University of North Carolina, Chapel Hill, NC.

AFIRM Team (2015c) Task Analysis. National Professional Development Center on Autism Spectrum Disorders, FPG Child Development Center, University of North Carolina, Chapel Hill, NC.

Ager, A. & O'May, F. (2001) Issues in the definition and implementation of 'best practice' for staff delivery of interventions for challenging behaviour. *Journal of Intellectual and Developmental Disability,* **26,** 243-256. doi: 10.1080/13668250120063412

Allen, D. & Tynan, H. (2000) Responding to aggressive behavior: Impact of training on staff members' knowledge and confidence. *Mental Retardation,* **38,** 97-104.

Barnes, C. S., Dunning, J. L. & Rehfeldt, R. A. (2011) An evaluation of strategies for training staff to implement the picture exchange communication system. *Research in Autism Spectrum Disorders,* **5,** 1574-1583.

Campbell, M. (2010) Workforce development and challenging behaviour: Training staff to treat, to manage or to cope? *Journal of Intellectual Disabilities,* **14,** 185-196.

Carver, R. P. (1974) Two dimensions of tests: Psychometric and edumetric. *American Psychologist,* **29,** 512-518. doi: 10.1037/h0036782

Clinical and Transactional Science Institute (UCSF) (2017) Sample size calculators: Sample size calculators for designing clinical research. In: *http://www.sample-size.net/means-sample-sizeclustered/*. http://www.sample-size.net/means-sample-sizeclustered/, http://www.sample-size.net/means-sample-sizeclustered/.

Cohen, J. (1960) A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement,* **20,** 37-46. doi: 10.1177/001316446002000104

Cooper, J. O., Heron, T. E. & Heward, W. L. (2014) Applied Behavior Analysis. Merrill Publishing Company, Columbus, OH.

Dench, C. (2005) A model for training staff in positive behaviour support. *Tizard Learning Disability Review,* **10,** 24-30.

Denne, L. D., Thomas, E., Hastings, R. P. & Hughes, J. C. (2015) Assessing competencies in applied behavior analysis for tutors working with children with autism in a school-based setting. *Research in Autism Spectrum Disorders,* **20,** 67-77.

Devane, D., Begley, C. M. & Clarke, M. (2004) How many do I need? Basic principles of sample size estimation. *Journal of Advanced Nursing,* **47,** 297-302.

Donner, A. & Klar, N. (2004) Pitfalls of and controversies in cluster randomization trials. *American Journal of Public Health,* **94,** 416-422.

Field, A. (2016) Discovering statistics using IBM SPSS statistics SAGE Publications Ltd., London, England.

Franzone, E. (2009) Steps for implementation: Functional communication training. National Professional Development Center on Autism Spectrum Disorders, Waisman Center, University of Wisconsin, Madison, WI

Fukui, S., Rapp, C. A., Goscha, R., Marty, D. & Ezell, M. (2014) The Perceptions of Supervisory Support Scale. *Administration and Policy in Mental Health and Mental Health Services Research,* **41,** 353-359.

Grey, I. M. & Hastings, R. P. (2005) Evidence-based practices in intellectual disability and behaviour disorders. *Current Opinion in Psychiatry,* **18,** 469-75. doi: 10.1097/01.yco.0000179482.54767.cf

Grey, I. M., Hastings, R. P. & McClean, B. (2007) Staff training and challenging behaviour. *Journal of Applied Research in Intellectual Disabilities,* **20,** 1-5.

Hatton, C. & Emerson, E. (1995) The development of a shortened 'Ways of Coping' questionnaire for use with direct care staff in learning disability services. *Journal of Applied Research in Intellectual Disabilities,* **8,** 237-251.

Homlitas, C., Rosales, R. & Candel, L. (2014) A further evaluation of behavioral skills training for implementation of the picture exchange communication system. *Journal of Applied Behavior Analysis,* **47,** 198-203.

Joint Committee on Social Care Professionals (2002) Final report of Joint Committee on social care professionals. Department of Social and Family Affairs, http://hdl.handle.net/10147/46110.

Kazdin, A. E. (2011) Single case research designs. Oxford University Press, New York, NY.

Kurtz, P. F., Boelter, E. W., Jarmolowicz, D. P., Chin, M. D. & Hagopian, L. P. (2011) An analysis of functional communication training as an empirically supported treatment for problem behavior displayed by individuals with intellectual disabilities. *Research in Developmental Disabilities,* **32,** 2935-2942.

Larsen, T. K., Melle, I., Auestad, B., Haahr, U., Joa, I., Johannessen, J. O., et al. (2010) Early detection of psychosis: Positive effects on 5-year outcome. *Psychological Medicine,* **41,** 1461-1469. doi: 10.1017/S0033291710002023

Lazarus, R. S. (1999) Stress and coping: A new synthesis. Springer, New York.

Locke, E. A. (1969) What is job satisfaction? *Organizational Behavior and Human Performance,* **4,** 309-336. doi: https://doi.org/10.1016/0030-5073(69)90013-0

Lyles, C. M., Kay, L. S., Crepaz, N., Herbst, J. H., Passin, W. F., Kim, A. S., et al. (2007) Best-evidence interventions: Findings from a systematic review of HIV behavioral interventions for US populations at high risk, 2000-2004. *American Journal of Public Health*, **97,** 133-43. doi: 10.2105/ajph.2005.076182

Maffei-Almodovar, L. & Sturmey, P. (2018) Change agent training in behavior analytic procedures for people with developmental and intellectual disabilities: A meta-analysis. *Review Journal of Autism and Developmental Disorders*, **5,** 129-141. doi: 10.1007/s40489-018-0128-6

Maggin, D. M., Fallon, L. M., Sanetti, L. M. H. & Ruberto, L. M. (2012) Training paraeducators to implement a group contingency protocol: Direct and collateral effects. *Behavioral Disorders,* **38,** 18-37.

Maslach, C. (1993) Burnout: A multidimensional perspective. In: *Professional burnout: Recent developments in theory and research.* (eds W. B. Schaufeli, C. Maslach, T. Marek, W. B. Schaufeli, C. Maslach & T. Marek). pp. 19-32. Taylor & Francis, Philadelphia, PA, US.

Maslach, C. (1999) A Multidimensional Theory of Burnout. In: *Theories of Organisational Stress* (ed C. L. Cooper). pp. 66-85. Oxford University Press, Oxford.

Maslach, C., Jackson, S. E. & Leiter, M. P. (1996) Maslach Burnout Inventory. Consulting Psychologists Press, Palo Alto, CA.

McKenna, H. P., Ashton, S. & Keeney, S. (2004) Barriers to evidence-based practice in primary care. *Journal of Advanced Nursing,* **45,** 178-189.

National Autism Center (2015) National Standards Project - Phase 2. http://www.nationalautismcenter.org/national-standards-project/history/.

National Institute for Health and Care Excellence (2015) Challenging behaviour and learning disabilities: prevention and interventions for people with learning disabilities whose behaviour challenges. NICE Guidelines. http://nice.org.uk/guidance/ng11.

Parsons, M. B., Rollyson, J. H. & Reid, D. H. (2012) Evidence-based staff training: A guide for practitioners. *Behavior Analysis in Practice,* **5,** 2-11.

Rapp, C. A., Etzel-Wise, D., Marty, D., Coffman, M., Carlson, L., Asher, D., et al. (2010) Barriers to evidence-based practice implementation: Results of a qualitative study. *Community Mental Health Journal,* **46,** 112-118.

Rigotti, T., Schyns, B. & Mohr, G. (2008) A short version of the occupational self-efficacy scale: Structural and construct validity across five countries. *Journal of Career Assessment,* **16,** 238-255.

Sarokoff, R. A. & Sturmey, P. (2004) The effects of behavioral skills training on staff implementation of discrete-trial teaching. *Journal of Applied Behavior Analysis,* **37,** 535-538.

Sarokoff, R. A. & Sturmey, P. (2008) The effects of instructions, rehearsal, modeling, and feedback on acquisition and generalization of staff use of discrete trial teaching and student correct responses. *Research in Autism Spectrum Disorders,* **2,** 125-136.

Seltman, H. J. (2015) Experimental design and analysis. http://www.stat.cmu.edu/~hseltman/AboutMe.html.

Sturmey, P. & Didden, R. (2014) Evidence-based practice and intellectual disabilities. Wiley-Blackwell.

Underwood, M., O'Meara, S. & Harvey, E. (2002) The acceptability to primary care staff of a multidisciplinary training package on acute back pain guidelines. *Family Practice,* **19,** 511-515. doi: 10.1093/fampra/19.5.511

van Oorsouw, W. M., Embregts, P. J., Bosman, A. M. & Jahoda, A. (2009) Training staff serving clients with intellectual disabilities: A meta-analysis of aspects determining effectiveness. *Research in Developmental Disabilities,* **30,** 503-11.

Ward-Horner, J. & Sturmey, P. (2012) Component analysis of behavior skills training in functional analysis. *Behavioral Interventions,* **27,** 75-92.

Wears, R. L. (2002) Advanced statistics: statistical methods for analyzing cluster and cluster-randomized data. *Academic Emergency Medicine,* **9,** 330-41.

Weiss, D. J., Dawis, R. V., England, G. W. & Lofquist, L. H. (1967) Manual for the Minnesota Satisfaction Questionnaire. University of Minnesota, Industrial Relations Center, Work Adjustment Project, Minneopolis, USA.

Whitlock, E. P., Polen, M. R., Green, C. A., Orleans, T. & Klein, J. (2004) Behavioral counseling interventions in primary care to reduce risky/harmful alcohol use by adults: A summary of the evidence for the U.S. Preventive Services Task Force. *Annals of Internal Medicine,* **140,** 557-68.

Wong, C., Odom, S. L., Hume, K., Cox, A., Fettig, A., Kucharczyk, S., et al. (2014) Evidence-based practices for children, youth and young adults with autism spectrum disorder. The University of North Carolina, Frank Porter Graham Child Development Institute, Autism Evidence-Based Practice Review Group, Chapel Hill.

Zaharia, E. S. & Baumeister, A. A. (1978) Technician turnover and absenteeism in public residential facilities. *American Journal of Mental Deficiency*, **82,** 580-593.