



Kent Academic Repository

MacDonald, Anne (2016) *An Evaluation of Staff Training in Positive Behaviour Support*. Doctor of Philosophy (PhD) thesis, University of Kent, University of Kent.

Downloaded from

<https://kar.kent.ac.uk/54604/> The University of Kent's Academic Repository KAR

The version of record is available from

This document version

UNSPECIFIED

DOI for this version

Licence for this version

UNSPECIFIED

Additional information

Versions of research works

Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in *Title of Journal*, Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).

An Evaluation of Staff Training in Positive Behaviour Support

Abstract

AIMS: Challenging behaviour is common for many people with learning disabilities and has a negative impact on the lives of these individuals. It is linked to decreased levels of support from staff, reduced opportunities for inclusion in the community, use of restrictive interventions, and placement breakdown. Equipping staff with the necessary knowledge, skills and experience to support people with challenging behaviour in a positive, respectful and effective way has proved a challenge for care agencies. Positive Behaviour Support (PBS) has been shown to be effective in minimising challenging behaviour. The aim of this study was to evaluate the impact of training managers of social care services in PBS. **METHOD:** A longitudinal training programme in PBS was delivered to 50 managers of community-based services for people with learning disabilities and challenging behaviour. The training programme lasted a year; data were collected pre and post training, and at 6 month follow-up. A non-randomised control group design was used. **RESULTS:** Data demonstrated significant reduction in challenging behaviour which was sustained over time. However, there was no change in quality of life for service users, and very limited changes in staff support to service users. **CONCLUSION:** This study has demonstrated that training managers in PBS can have a positive impact on challenging behaviour in people with learning disabilities. There are a number of aspects to the results which are unexpected and these are discussed with reference to the relevant literature.

Tizard Centre, University of Kent

2016

Word Count: 92,074

Acknowledgements

Thanks to all the staff and service users who participated in this research.

Thanks to my team for all their hard work in collecting data and coding videos; I couldn't have done it without your help.

Thanks to my manager for his support and flexibility throughout the study.

Thanks to my supervisor, Peter McGill for his guidance, encouragement, patience and good humour over the past 5 years.

And finally, thank you John, for all you've had to put up with, and for everything you've done to help make this possible.

Table of Contents

Acknowledgements	ii
Table of Contents	iii
List of Figures.....	vi
List of Tables.....	viii
1 Thesis Overview, Learning Disability and Challenging Behaviour	1
1.1 Thesis Overview	1
1.2 Learning Disability	2
1.3 Challenging Behaviour	3
1.4 Chapter Summary	9
2 Positive Behaviour Support	10
2.1 Chapter Outline	10
2.2 Development and Definition.....	10
2.3 Evidence for Effectiveness.....	21
2.4 Chapter Summary	40
3 Review of the Literature in Relation to Staff Training in Positive Behaviour Support	41
3.1 Chapter Outline.....	41
3.2 Development of Positive Behaviour Support Training	41
3.3 Definition of Positive Behaviour Support Training	44
3.4 Method	46
3.5 Results.....	56
3.6 Discussion	74
3.7 Conclusions.....	79
3.8 Chapter Summary	80
4 Introduction to Study	81
4.1 Purpose of Study	81
4.2 Theory of Process of Change	82
5 Method.....	86
5.1 Chapter Outline.....	86
5.2 Design	86
5.3 Participants.....	86
5.4 Group Size	90
5.5 Ethics	91
5.6 Measures	91
5.7 Procedure	99

5.8	Data Collection and Reliability.....	103
5.9	Missing Data.....	107
5.10	Data Analysis.....	109
5.11	Results from Reliability Checks.....	114
5.12	Results from Missing Data.....	118
5.13	Theory of Process of Change.....	118
5.14	Chapter Summary.....	120
6	Managers' Results.....	122
6.1	Chapter Outline.....	122
6.2	Managers' Results at T1.....	122
6.3	Managers' Results at T2: Impact of Training.....	124
6.4	Managers' Results at T3: Maintenance of Training Effects.....	132
6.5	Chapter Summary.....	139
7	Staff Results.....	141
7.1	Chapter Outline.....	141
7.2	Staff Results at T1.....	141
7.3	Staff Results at T2: Impact of Training.....	144
7.4	Staff Results at T3: Maintenance of Training Effects.....	157
7.5	Chapter Summary.....	168
8	Service User Results.....	170
8.1	Chapter Outline.....	170
8.2	Service User Results at T1.....	170
8.3	Service User Results at T2: Impact of Training.....	174
8.4	Service User Results at T3: Maintenance of Training Effects.....	192
8.5	Chapter Summary.....	209
9	Discussion of Managers' and Staff Results.....	211
9.1	Chapter Outline.....	211
9.2	Study Limitations.....	211
9.3	Maintenance of Training Effects (T1-T3).....	235
9.4	Chapter Summary.....	241
10	Discussion of Service User Results.....	244
10.1	Chapter Outline.....	244
10.2	Study Limitations.....	244
10.3	Impact of Training (T1-T2).....	246
10.4	Maintenance of Training Effects (T1-T3).....	262
10.5	Chapter Summary.....	266
11	Overall Discussion.....	268

11.1	Chapter Outline.....	268
11.2	Summary of Results.....	268
11.3	Theory of Process of Change.....	269
11.4	Actual Process of Change for Challenging Behaviour	270
11.5	Actual Process of Change for Quality of Life.....	273
11.7	Concluding Comments.....	277
12	References.....	285
13	Appendices	311

List of Figures

Figure 1.	Process for Assessing Abstracts	50
Figure 2.	Theory of Process of Change for Challenging Behaviour	84
Figure 3.	Theory of Process of Change for Quality of Life	85
Figure 4.	Theory of Process of Change in Relation to Measures Used in this Study.....	121
Figure 5.	Mean Percentage Practice Leadership T1 and T2	125
Figure 6.	Managers' Mean Percentage PBS Knowledge T1 and T2	127
Figure 7.	Mean Percentage Practice Leadership T1-T3	133
Figure 8.	Managers' Mean Percentage PBS Knowledge T1-T3.....	134
Figure 9.	Mean Percentage Assistance at T1 and T2.....	145
Figure 10.	Mean Percentage Positive Contact at T1 and T2	146
Figure 11.	Mean Percentage No Contact at T1 and T2	146
Figure 12.	Mean Percentage ASM at T1 and T2	148
Figure 13.	Staff Mean Percentage PBS Knowledge T1 and T2	153
Figure 14.	Mean Percentage Assistance T1-T3	159
Figure 15.	Mean Percentage Positive Contact T1-T3.....	160
Figure 16.	Mean Percentage No Contact T1-T3.....	160
Figure 17.	Mean Percentage ASM T1-T3	163
Figure 18.	Staff Mean Percentage PBS Knowledge T1-T3.....	164
Figure 19.	Mean ABC Total at T1 and T2.....	175
Figure 20.	Mean ABC Severe Behaviours at T1 and T2.....	175
Figure 21.	Individual Change in ABC Total Scores T1-T2 Experimental.....	177
Figure 22.	Individual Change in ABC Total Scores T1-T2 Control.....	177
Figure 23.	Reliable Change in ABC Total Scores T1-T2 Experimental.....	179
Figure 24.	Reliable Change in ABC Total Scores T1-T2 Control	180
Figure 25.	Mean Frequency at T1 and T2	181
Figure 26.	Mean Severity at T1 and T2.....	182
Figure 27.	Mean Percentage Total Engagement T1 and T2.....	185
Figure 28.	Mean Percentage Disengagement T1 and T2.....	185
Figure 29.	Mean Percentage Challenging Behaviour T1 and T2	186
Figure 30.	Mean GCPLA Range T1 and T2	189
Figure 31.	Mean ABC Total T1-T3	195
Figure 32.	Mean ABC Severity T1-T3	195
Figure 33.	Individual Change in ABC Total Scores T1-T3	197
Figure 34.	Reliable Change in ABC Total Scores T1-T3	198
Figure 35.	Mean Frequency T1-T3.....	199

Figure 36.	Mean Severity T1-T3	200
Figure 37.	Mean Percentage Total Engagement T1-T3	203
Figure 38.	Mean Percentage Disengagement T1-T3	203
Figure 39.	Mean Percentage Challenging Behaviour T1-T3	204
Figure 40.	Mean GCPLA Range T1-T3	207
Figure 41.	Mean GCPLA Busy T1-T3	207
Figure 42.	Actual Process of Change for Challenging Behaviour.....	271
Figure 43.	Actual Process of Change for Quality of Life.....	275

List of Tables

Table 2.1: Studies used to define PBS	19
Table 2.2: Key Features Defining PBS	20
Table 2.3: Reviews of PBS Research	24
Table 3.1: Studies Defining PBS Training Format.....	46
Table 3.2: Studies Included and Excluded in Systematic Review (from the 28 studies read in full).....	51
Table 3.3: Details of Studies Included in the Systematic Review.....	57
Table 3.4: Length, Format and Content of Training in Studies Included in Systematic Review	67
Table 3.5: Staff Outcome Measures from Studies in Systematic Review.....	69
Table 3.6: Service User Outcome Measures from Studies in Systematic Review	74
Table 5.1: Managers’ Characteristics at T1, Experimental and Control Groups.....	89
Table 5.2: Service Users’ Characteristics at T1, Experimental and Control Groups	90
Table 5.3: Measures Used in the Study for Each Participant Group.....	91
Table 5.4: Summary of Areas in PBS Knowledge Test.....	97
Table 5.5: Content of Training Workshops	101
Table 5.6: Data Collection Points for Current Study	103
Table 5.7: Type of Test Used with Each Measure in the Comparative Literature.....	111
Table 5.8: Methods of Calculating Effect Size in Current Study	112
Table 5.9: Effect Sizes for Tests Used in Current Study	112
Table 5.10: Reliable Change Formula Definitions.....	114
Table 5.11: Inter-observer Reliability Data for MTS Behaviours	115
Table 5.12: Percentage Engagement and Assessment in Re-coding of Videos.....	116
Table 5.13: Service Users’ Number of Observations at T1 and T2 and Reasons for Short Observations	117
Table 5.14: Data Entry Agreement in Current Study	117
Table 5.15: Summary of Dropouts, Data Imputed and Type of Imputation.....	119
Table 6.1: Difference between Cohort One and Cohort Two at T1 for Managers’ Measures for Experimental and Control Groups.....	123
Table 6.2: Difference between Experimental and Control at T1 for Managers’ Measures ...	124
Table 6.3: Practice Leadership Descriptive Statistics T1 and T2.....	125
Table 6.4: Practice Leadership Sensitivity Analysis T1-T2	126
Table 6.5: Practice Leadership Questionnaire Individual Areas Experimental Group T1-T2	126
Table 6.6: Managers’ PBS Knowledge Descriptive Statistics T1 and T2	127

Table 6.7: Managers' PBS Knowledge Sensitivity Analysis T1-T2	128
Table 6.8: Managers' CHABA Descriptive Statistics T1 and T2	128
Table 6.9: Managers' CHABA Significance Testing T1-T2	129
Table 6.10: Managers' CHABA Sensitivity Analysis T1-T2	130
Table 6.11: Correlations for Managers' Measures at T2	131
Table 6.12: Practice Leadership Descriptive Statistics T1-T3	132
Table 6.13: Practice Leadership Sensitivity Analysis T1-T3	134
Table 6.14: Managers' PBS Knowledge Descriptive Statistics T1-T3.....	134
Table 6.15: Managers' PBS Knowledge Sensitivity Analysis T1-T3 and T2-T3.....	135
Table 6.16: Managers' CHABA Descriptive Statistics T1-T3	136
Table 6.17: Managers' CHABA Significance Testing T1-T3	137
Table 6.18: Managers' CHABA Sensitivity Analysis T1-T3	137
Table 6.19: PSR Descriptive Statistics T2 and T3	138
Table 6.20: PSR Sensitivity Analysis T2-T3.....	138
Table 6.21: Correlations for Managers' Measures at T3	139
Table 7.1: Difference between Cohort One and Cohort Two at T1 for Staff Measures for Experimental and Control Groups	142
Table 7.2: Difference between Experimental and Control at T1 for Staff Measures	144
Table 7.3: Staff MTS Descriptive Statistics T1 and T2	145
Table 7.4: Staff MTS Sensitivity Analysis T1-T2.....	147
Table 7.5: ASM Descriptive Statistics T1 and T2.....	148
Table 7.6: Percentages of Group in Each Category of ASM Score at T1 and T2	148
Table 7.7: ASM Sensitivity Analysis T1-T2	149
Table 7.8: ASM Means per Question at T1 and T2, Percentage Change from T1-T2, and Comparison of Change Scores from T1-T2.....	150
Table 7.9: Percentage Scoring 'Good' per Question at T1 and T2 for Experimental and Control	152
Table 7.10: Staff PBS Knowledge Descriptive Statistics T1 and T2	153
Table 7.11: Staff PBS Knowledge Sensitivity Analysis T1-T2.....	153
Table 7.12: Staff CHABA Descriptive Statistics T1 and T2.....	154
Table 7.13: Staff CHABA Significance Testing T1-T2.....	154
Table 7.14: Staff CHABA Sensitivity Analysis T1-T2	155
Table 7.15: Correlations for Staff and Managers' Measures T1 and T2	156
Table 7.16: Practice Leadership Scores for Each Category of ASM Score at T1 and T2	156
Table 7.17: Staff MTS Descriptive Statistics T1-T3.....	158
Table 7.18: Staff MTS Sensitivity Analysis T1-T3.....	161
Table 7.19: ASM Descriptive Statistics T1-T3	162

Table 7.20: Percentage of Group in Each Category of ASM Score T1-T3	162
Table 7.21: Practice Leadership Scores for Each Category of ASM Score T1-T3	162
Table 7.22: ASM Sensitivity Analysis T1-T3	163
Table 7.23: Staff PBS Knowledge Descriptive Statistics T1-T3.....	164
Table 7.24: Staff PBS Knowledge Sensitivity Analysis T1-T3.....	165
Table 7.25: Staff CHABA Descriptive Statistics T1-T3	166
Table 7.26: Staff CHABA Significance Testing T1-T3.....	167
Table 7.28: Correlations for Staff and Managers' Measures at T3.....	168
Table 8.2: Difference between Cohort One and Cohort Two at T1 for Service User Measures for Experimental and Control Groups.....	171
Table 8.3: Difference between Experimental and Control for Service User Measures and Characteristics T1	172
Table 8.4: Correlations at T1 for Managers', Staff, and Service User Measures.....	173
Table 8.5: Descriptive Statistics for ABC T1 and T2.....	174
Table 8.6: Sensitivity Analysis ABC Total and Severe Behaviours T1-T2	176
Table 8.7: Reduction to the Baseline in ABC Total for Experimental and Control T1-T2 ..	178
Table 8.8: Reliable Change Data.....	179
Table 8.9: Descriptive Statistics for ABC Difference for Each Trainer	180
Table 8.10: Descriptive Statistics for BRF T1 and T2	181
Table 8.11: Sensitivity Analysis BRF Frequency and Severity T1-T2	183
Table 8.12: Reduction to the Baseline in BRF for Experimental and Control T1-T2.....	184
Table 8.13: Service Users' MTS Descriptive Statistics T1 and T2	184
Table 8.14: Services Users' MTS Sensitivity Analysis T1-T2.....	187
Table 8.15: Descriptive Statistics for Observed and Videoed Groups	188
Table 8.16: GCPLA Descriptive Statistics T1 and T2	188
Table 8.17: GCPLA Range and Busy Sensitivity Analysis T1-T2.....	189
Table 8.18: ABC Correlations T2.....	190
Table 8.19: Correlations T2 for all Measures.....	191
Table 8.23: ABC Descriptive Statistics T1-T3.....	194
Table 8.24: ABC Total and Severe Behaviours Sensitivity Analysis T1-T3 and T2-T3	196
Table 8.25: Reduction to the Baseline in ABC Total and Severe Behaviours T1-T3.....	198
Table 8.26: BRF Descriptive Statistics T1-T3	199
Table 8.27: Frequency Sensitivity Analysis T1-T3 and T2-T3 and Severity Sensitivity Analysis T1-T3	201
Table 8.28: Reduction to the Baseline in BRF T1-T3.....	201
Table 8.30: MTS Sensitivity Analysis T1-T3	205
Table 8.31: GCPLA Descriptive Statistics T1-T3.....	206

Table 8.32: GCPLA Range and Busy Sensitivity Analysis T1-T3	208
Table 8.33: Correlations T3 for Managers', Staff, and Service User Measures.....	209
Table 9.1 Comparisons with Other Studies for Practice Leadership Index and for Team Meetings, Supervision and Observation	216
Table 9.2: Item Comparison on the ASM for Mean Scores at T1 and Percentage Scoring 'Good' at T1	221
Table 9.3 ASM and Assistance Comparisons.....	222
Table 9.4: Setting Comparisons	224
Table 9.5: Type of Service in Current Study	225
Table 9.6: Observation Comparisons	226
Table 9.7: PBS Knowledge Mean Percentage Comparisons.....	229
Table 9.8: CHABA Comparisons.....	233
Table 9.9: PSR Comparisons	236
Table 9.10: Assistance Comparisons T1-T3.....	240
Table 9.11: PBS Knowledge Comparisons at T3	241
Table 9.12: CHABA Comparisons at T3.....	243
Table 10.1: Comparison of ABC Scores at T1 and T2 for Total and Severe Behaviours, and Mean Reduction to Percentage of Baseline	249
Table 10.2: Comparative Reductions to Percentage of Baseline from T1-T2 Using BRF...	253
Table 10.3: MTS Challenging Behaviour Comparisons T1-T2	255
Table 10.4: Comparisons for Engagement T1-T2 and ABS.....	258
Table 10.5: GCPLA Comparisons at T1 and T2 for Range and Busy.....	261
Table 10.6: ABC Comparisons T1- T3.....	263
Table 10.7: Comparative Reductions to Percentage of Baseline from T1-T3 Using BRF...	264
Table 10.8: MTS Challenging Behaviour Comparisons T1-T3	265
Table 10.9: Comparisons for Engagement T1-T3	266
Table 13.1: Practice Leadership ANOVA Results for Time and Group	323
Table 13.2: Managers' PBS Knowledge ANOVA Results for Time and Group	323
Table 13.3: Managers' CHABA Learned Positive ANOVA Results for Time and Group ...	323
Table 13.4 Managers' CHABA Learned Negative ANOVA Results for Time and Group...	323
Table 13.5: Managers' CHABA Biomedical ANOVA Results for Time and Group.....	324
Table 13.6: Managers' CHABA Stimulation ANOVA Results for Time and Group.....	324
Table 13.7: Managers' CHABA Emotional ANOVA Results for Time and Group.....	324
Table 13.8: Managers' CHABA Physical Environment ANOVA Results for Time and Group	324
Table 13.9: MTS Assistance ANOVA Results for Time and Group.....	325
Table 13.10: MTS Positive Contact ANOVA Results for Time and Group	325

Table 13.11: MTS No Contact ANOVA Results for Time and Group	325
Table 13.12: Staff PBS Knowledge ANOVA Results for Time and Group.....	325
Table 13.13: Staff CHABA Learned Positive ANOVA Results for Time and Group	326
Table 13.14: Staff CHABA Learned Negative ANOVA Results for Time and Group.....	326
Table 13.15: Staff CHABA Biomedical ANOVA Results for Time and Group.....	326
Table 13.16: Staff CHABA Stimulation ANOVA Results for Time and Group.....	326
Table 13.17: Staff CHABA Emotional ANOVA Results for Time and Group.....	327
Table 13.18: Staff CHABA Physical Environment ANOVA Results for Time and Group..	327
Table 13.19: ABC Total ANOVA Results for Time and Group.....	328
Table 13.20: GCPLA Range ANOVA Results for Time and Group.....	328
Table 13.21: GCPLA Busy ANOVA Results for Time and Group.....	328

1 Thesis Overview, Learning Disability and Challenging Behaviour

1.1 Thesis Overview

The current chapter will outline the current definition of learning disability, and will then go on to consider its prevalence in the United Kingdom (UK); as this study is based in Scotland there will be specific attention paid to relevant Scottish data. The chapter will then go on to consider the definition, prevalence and impact of challenging behaviour, along with some reflection on explanatory models. A range of interventions for challenging behaviour will be outlined, with links made to Positive Behaviour Support (PBS).

Chapter 2 traces the development of PBS from its roots in Applied Behaviour Analysis (ABA) to becoming a discipline in its own right. This provides a number of key defining features of PBS based on a range of studies that have defined it over the past 25 years. Following this, there is a review of the effectiveness of PBS based on evidence in the literature; a number of individual studies are considered as well as findings from literature reviews and meta-analyses.

Chapter 3 provides a definition of PBS training, which combines an emphasis on PBS content as well as a particular approach to providing training, and then the findings from a systematic review of PBS training are presented. Results from the studies are detailed, with reference to outcomes for both staff and service users.

Chapter 4 provides an overview of the study which is the focus of this thesis. An outline is given of the theory of process of change which was expected in the study, both in relation to service users' challenging behaviour and also quality of life for service users.

Chapter 5 describes the method used to carry out this study. This includes study design, participants, procedure, measures, data collection, reliability checks, approach taken to missing data, data analysis, and results from reliability checks. A version of the theory of process of change is presented in relation to the measures used in this study

Chapters 6-8 contain the results from the study, a chapter each considering results for managers, staff and service users. Each of these chapters first considers results pre training, in terms of any differences between the groups at baseline. The main part of each chapter considers the effectiveness of the training, presenting results from pre to post training. The chapters conclude with consideration of whether any training effects are sustained and present results from pre training to follow-up.

Chapters 9 and 10 discuss the results of the training firstly for both the managers and staff, which are considered together as there is overlap of measures, and then for service users. A range of relevant literature is considered in order to put the results into context.

Chapter 11 presents an overall discussion, based on results for all participants: managers, staff and service users. An updated model of the process of change is presented based on the model in chapter 4, updated to reflect the actual changes. The thesis ends with some concluding thoughts and implications, both for PBS training and for future research.

1.2 Learning Disability

1.2.1 Terms and Definition

Historically, there are many different terms that have been used to describe learning disability and today there are a number of terms in use. Until recently the term ‘mental retardation’ was common in the United States, although now ‘developmental disability’ is more commonly used. In the UK the term ‘intellectual disability’ tends to be used in research; however, since commissioning bodies and service providers in the UK mainly refer to ‘learning disability’, this study uses that term throughout.

Learning disability is a condition characterised by significant impairments of both intellectual and adaptive functioning with an onset before the age of 18 (American Psychiatric Association, 1994). The degree of learning disability can be mild, moderate, severe or profound, with over 90% of those affected falling within the mild range (Department of Health (DOH), 2001).

1.2.2 Prevalence

It is difficult to be clear about numbers of people with learning disabilities as these vary depending on the different ways of estimating this; the three main approaches involve the use of general statistical principles of distribution of intelligence within the whole population (usually known as ‘true prevalence’); using specifically collected data, which can be resource intensive and time-consuming to collect (e.g. Lowe et al, 2007a); or by combining information collected by government departments on the presence of people with learning disabilities using particular services (usually known as ‘administrative prevalence’). Estimates based on general population and distribution of intelligence may be impacted by the fact that people with learning disabilities die earlier than the rest of the population, and therefore in the older age group this estimate will be less accurate. In addition, estimates can be influenced by the fact many people whose IQ may fall within the learning disability range (less than 70), may be living without any support and functioning adaptively without difficulty, and therefore may not be known to services, making government figures low compared to true prevalence.

In England, the most recent estimate of the number of people with learning disabilities appears to be from Public Health England which states that in England in 2012 there were 1.14 million people with learning disabilities, of which 908, 000 were over 18 (Emerson et al, 2013).

This is based on combining information collected by government departments about numbers of people with learning disabilities using services, overall population predictions, and epidemiological research.

In Scotland, learning disability data is collected in three ways: via an annual census of Scottish schools, annual returns from local authorities about the number of adults with learning disabilities using services, and information from General Practitioners' databases of all patients with learning disabilities over the age of 18 (Scottish Government (SG), 2013). In 2014, which is the most recent information available from the 32 local authorities in Scotland, the total figure of adults with learning disabilities was 28,786 (Scottish Consortium for Learning Disability (SCLD), 2015), that is 6.0 per 1000 adults in the general population; this figure is only based on people known to local authorities and therefore does not include anyone with a learning disability who was not using services.

1.3 Challenging Behaviour

1.3.1 Definition

The most recent definition of challenging behaviour is that from the Royal College of Psychiatrists: 'behaviour of such an intensity, frequency or duration as to threaten the quality of life and/or the physical safety of the individual or others and is likely to lead to responses that are restrictive, aversive or result in exclusion' (Royal College of Psychiatrists et al, 2007, p.10). This definition clearly links the behaviour with its social consequences, that is, exclusion, restrictions or aversive responses.

Challenging behaviour is a socially constructed concept that covers a diverse group of behaviours and may present differently in people with profound and severe learning disabilities from the way it does in people with mild learning disabilities and mental health issues. The term was originally adopted in the United States by The Association for Severe Handicaps, and was used to emphasise that the behaviours represent a challenge to services, rather than a set of problems located in the person, which may be the implication with terms such as 'problem behaviour' or 'difficult behaviour'. Whether behaviour is regarded as challenging depends on many factors including the setting and the impact of the behaviour upon the setting; the expectations and social rules within the setting; and the perceived causes for the behaviour (Emerson & Einfeld, 2011). Challenging behaviour includes behaviours such as physical aggression, self-injury, property destruction, disruptive behaviours, sexualised behaviours, rituals and stereotypical behaviours.

1.3.2 Prevalence

In terms of prevalence of challenging behaviour, estimates of this have been around 5% to 15% of people with learning disabilities (Ball, Bush & Emerson, 2004; Emerson & Einfeld, 2011; Lowe et al, 2007a); variety exists due to different ways of measuring this and also due to differences in definition. Emerson et al (2001) suggested a prevalence rate of 10-15% and Lowe et al (2007a) reported on a study screening all services for people with learning disabilities in seven local authorities; 4.5 per 10,000 of the general population were rated as severely challenging, equalling 10% of the learning disability population. Both of these studies also found that two thirds of those with severe challenging behaviour were male.

1.3.3 Impact

Challenging behaviour is associated with a range of negative outcomes which may significantly impact on the person and those around them. This includes abuse (Rusch et al, 1986), placement breakdown (Broadhurst & Mansell, 2007), staff stress and negative staff reactions (Bromley & Emerson, 1995; Hastings, 1995), neglect (Mansell, 1995), and increased costs (Knapp et al, 2005). Challenging behaviour is also likely to persist over time (Taylor et al, 2011; Totsika et al, 2008) and is the most common reason given for placement breakdown and the use of out of area placements in specialist units (Emerson & Einfeld, 2011).

1.3.4 Explanatory Models

Studies regarding epidemiology of challenging behaviour have suggested a number of risk markers. In 2003 McClintock et al carried out a meta-analysis of risk indicators for challenging behaviour looking at 86 studies. They described four types of challenging behaviours: self-injury, aggression, stereotypical behaviours and destruction of property and found that males were more likely to have challenging behaviour than females, as were people with autism or with more severe learning disabilities. Challenging behaviour can also be a manifestation of an underlying mental or physical health issue and is associated with a range of specific syndromes, sensory impairment, physical disabilities, and a number of additional disabilities. Emerson & Einfeld (2011) summarised this research, and the recent guidelines from the National Institute for Health and Care Excellence (British Psychological Society and Royal College of Psychiatrists, 2015) also referred to a range of research in relation to characteristics associated with challenging behaviour. However, little is known about the exact process by which these factors influence susceptibility to the development of challenging behaviour (Allen, 2013a).

Hastings et al (2013) outlined a conceptual framework for understanding why challenging behaviours occur in people with learning disabilities; this included both biological and psychosocial vulnerabilities including physical health problems, genetic factors, negative life events,

poor communication skills, limited social networks, lack of meaningful activity, and underlying mental health issues. However the authors noted that factors which contribute to the maintenance of challenging behaviour also need to be considered in order to ensure a full understanding; they noted that ‘challenging behaviours must be useful in some way if they continue to occur’ (p.8). An understanding of functions of behaviour and the purpose of challenging behaviour from the person’s point of view is central to offering helpful intervention and support; this is not a straightforward process as behaviours can serve multiple functions, and different behaviours can have the same function. The final important factor in this overall framework for making sense of challenging behaviour is the behaviour of carers; understanding their values system and underlying beliefs, for instance about people with learning disabilities, is vital in order to ensure helpful responses. Staff will call upon their own cultural background and experiences in order to understand the challenging behaviour they are presented with; if this is unhelpful, for example, coming from a culture in which behaviour should be punished, then it is important to identify and address this as part of any intervention. A variety of studies have argued that conflict between service aims and personal beliefs is detrimental to effective intervention, for example, Hastings & Remington (1994b).

1.3.5 Intervention

1.3.5.1 Specialist Units

Many people with challenging behaviour find it difficult to get services that can meet their needs; this has resulted in the treatment of people with challenging behaviour within specialist health-based assessment and treatment units (Beadle-Brown et al, 2006). These specialist units provide support for the most challenging individuals and are often regarded as a last resort for those whose community placements, have broken down (Broadhurst & Mansell, 2007).

However, these units are often out of area placements, which are associated with a number of negative factors: they are often geographically distant from people’s homes and therefore have a negative impact on family life and on the ability of families to maintain connections with their loved one. They are based on an institutional model, which has a number of associated negative factors: a closed culture where it is difficult to challenge practice, resulting in a higher likelihood of poor practice and abuse (DOH, 2012), and the potential for increasing problem behaviours due to the congregating of so many people with behavioural challenges in one location (Newman & Emerson, 1991). They are expensive (Allen et al, 2007; Hassiotis et al, 2008; McGill & Poynter, 2012) and have struggled to evidence good outcomes for residents (Beadle-Brown et al, 2006; Beadle-Brown et al, 2009a; DOH, 2012). In addition, the creation of these units may also undermine the development of local services with the necessary skills to support people with complex and challenging behaviours.

Out of area specialist placements for people with challenging behaviour appear to generally be a symptom of a wider systems failure which frequently includes lack of effective local services for people with challenging behaviours (DOH, 2007). It is recognised that for some people with mild learning disabilities who may be under criminal justice legislation, then some specialist units are more successful and appropriate; however that is not the client group included in this study and therefore this discussion relates more to people with more severe learning disabilities and complex needs who are placed in out of area specialist units. There have been a number of recommendations about the need to reduce the use of these units. For example, in 2007, the DOH provided guidance on the development of responsive local services for people with challenging behaviour, including recommendations around service improvements, the need to produce better outcomes, demonstrate value for money, and support the families of people with complex needs. There have also been a number of policy statements by both UK and Scottish governments about the need to reduce their use (DOH, 2014; SG, 2013).

Over the last 30 years there has been a reduction in adult inpatient beds in Scotland from 7,000 in 1980 to 318 in the most recent annual Scottish survey carried out in 2012 (SG, 2013). Clearly this trend in reduction is not specific to Scotland (e.g. see Mansell & Ericsson, 1996 for a discussion of deinstitutionalisation in the wider UK), but as this study was conducted in Scotland, that is the focus of this discussion. In 2000, *'The Same as You?'*, the Scottish Government's review of services for people with learning disabilities, recommended that health boards should aim to reduce their assessment and treatment places for people with learning disabilities to four for every 100,000 people across the country, and that they should plan for appropriate community services to avoid inpatient assessment and treatment. More recently, the Scottish Government's current strategy document for people with learning disabilities, *The Keys to Life* (2013) gave an update on this recommendation and stated 'some NHS boards have still not made community-based service re-provision for individuals in 'longer-stay' NHS residential services, as opposed to inpatient assessment and treatment beds, and some of these longer-stay beds remain on the geographical site of the former long-stay hospitals' (p.124). Of the 318 individuals identified in the 2012 survey, 52 individuals had been placed out of area. Additionally, 78 of these 318 had been admitted to inpatient services directly from long-stay hospitals and 97 individuals had been in hospital for more than five years.

In 2014 a census of in-patient beds in Scotland for learning disability and mental health was carried out; this indicated that out of the 3,909 people in inpatient beds, 230 people were diagnosed with a learning disability although the report noted that there was no diagnostic information for 272 people so this figure could not be taken as the entire learning disability inpatient population; 181 were in a learning disability unit, and 226 had a learning disability consultant. It also found that people with learning disabilities were more likely to have longer stays

in these hospitals than patients with no identified learning disability; the average time since admission was 22 months as opposed to five months for mental health or addictions (SCLD, 2015).

In England, since the abuse scandal that took place at Winterbourne View in 2011, there has been a considerable impetus to move people from assessment and treatment units, to smaller and more local social care provision. A number of high level reports and papers have been written addressing the issue (*Winterbourne View: Time is Running Out*, Association of Chief Executives of Voluntary Organisations (ACEVO), 2015; *Positive and Proactive Care: Reducing the Need for Restrictive Interventions*, DOH, 2014; *Winterbourne View Review. Transforming Care One Year On*, DOH, 2013; *Transforming Care: A National Response to Winterbourne View Hospital*, DOH, 2012; *Supporting people with a learning disability and/or autism who have a mental health condition or display behaviour that challenges*, 2015, Local Government Association, Association of Directors of Adult Social Services, and NHS England; *Ensuring Quality Services*, Local Government Association, 2014; *Winterbourne View: Time for Change*, NHS England, 2014), and in 2015, the Chief Executive of NHS England announced a national closure plan for assessment and treatment units (House of Commons, Committee of Public Accounts, 2015). The most recent figure for people with learning disabilities in inpatient facilities comes from the Committee of Public Accounts which reported 3230 people with learning disabilities in inpatient facilities in September 2014 (House of Commons, 2015). This is despite a UK Government commitment in 2012, that if anyone with a learning disability and challenging behaviour in England would be better off supported in the community, then they should be moved out of hospital by June 2014. This report also noted that a fifth of people with learning disabilities in inpatient settings had been there for over five years, and over a third lived more than 50 kilometres from their home; it also stated that ‘having people with learning disabilities living in hospitals is incompatible with the Department’s model of care for people with learning disabilities and challenging behaviour’ (p.5).

1.3.5.2 Specialist Behavioural Teams

These are peripatetic teams that are usually health-based and multi-disciplinary, providing a range of training and direct support to people with learning disabilities with challenging behaviours and their staff/services (Emerson et al, 1987; Toogood et al, 1994). A survey in 1996 (Emerson) found 65 such teams in England and Wales; a more recent survey covering both Scotland and England contacted 46 services and received 20 returns (Davison et al, 2015). This survey reported that 80% of the teams were provided by the NHS and that these were mainly managed by a nurse; only 16% were managed by a behaviour analyst. Reasons for referral were mainly linked with challenging behaviour (increase in behaviour, risk of placement breakdown

or behaviour causing limitations to access), and 83% of service users were displaying challenging behaviour at the time of referral. As regards interventions, 47% reported using PBS.

These teams have been found to be effective both in reducing challenging behaviour and in impacting positively on quality of life (Allen et al, 2011; Hassiotis et al, 2009), although research is mixed and some studies reported less successful outcomes. Lowe et al (1996) found only one of the two teams evaluated over three years had positive outcomes; they also reported the team using interventions based on ABA was more successful. In the Hassiotis study (2009), treatment as usual from community learning disability teams was compared with a group receiving enhanced treatment, from a specialist behavioural team; the study reported that both groups demonstrated reduction in challenging behaviour, although the specialist team group reduced more. Allen et al (2011) reported outcomes for two specialist behavioural teams using a PBS approach; they found reductions in challenging behaviour and reduced use of restrictive practices, as well as improvements in use of community and leisure.

1.3.5.3 Restrictive Interventions

Challenging behaviour is also managed via a range of restrictive interventions, such as physical restraint and the use of medication (Emerson et al, 2000a; Sturmey, 2009). A survey of residential services in the south east of England in 2008 found that physical intervention was used by 47%, and that of these, only 65% had a policy with regard to its use (Deveau & McGill, 2009). Physical intervention has been associated with abusive practice (MacDonald et al, 2011), with pain, and with risk of injury to both staff and service users (Allen & Tynan, 2000; Leadbetter, 2002).

Anti-psychotic medication is also widely used (McGillvray & McCabe, 2005; Public Health England, 2015; Sheehan et al. 2015), despite lack of evidence for its effectiveness (Tyrer et al, 2008). Emerson et al (2000a) compared treatment approaches and outcomes for three different service models for people with challenging behaviour; village communities, NHS residential settings, and community-based housing. They found that different models of support were more likely to use specific approaches, for example, health settings used more restraint and community services used more sedation. Overall, they found that service users were more than three times more likely to receive medication as treatment than behavioural support, and that nearly half were subject to physical restraint. More recently, Public Health England (2015) reported on the use of a range of psychotropic medication for people with learning disabilities and/or autism. Based on GP records from the Clinical Practice Research Datalink (which collects clinical data from a substantial number of GP practices in England and is considered to provide a good representation of practice across England), they found that 29.5% of people with learning disabilities were on psychotropic drugs of some kind, including antipsychotics, antidepressants, hypnotics and anxiolytics.

1.3.5.4 Behavioural Interventions

Since the 1960s, behavioural approaches to treat challenging behaviour have come to the fore with a growing understanding that environmental factors can influence behaviour, rather than challenging behaviour being a manifestation of internal pathology (Emerson & Einfeld, 2011, chapter 4). Around this period ABA became seen as an increasingly successful approach to treating challenging behaviour (Baer et al, 1968). The growing influence of this approach was seen with the launch of the Journal of Applied Behavior Analysis in 1968 which since then has published many studies demonstrating the success of ABA in reducing challenging behaviour.

Over the past 25 years PBS has been increasingly regarded as an appropriate and effective intervention for people with learning disabilities and challenging behaviour (Carr et al, 1999). Particularly since the abuse scandal that took place at Winterbourne View in 2011, there has also been a considerable focus on PBS in England at a national and governmental level. It has been recommended in a series of high-level reports and good practice guidance (ACEVO, 2015; DOH, 2014).

Although there is less of a focus on PBS in Scotland, it is referred to in the Scottish Government's new strategy for people with learning disabilities in Scotland, launched in 2013. This states that 'PBS is the recommended approach to supporting people with learning disabilities and behavioural difficulties' (p.121).

1.4 Chapter Summary

This chapter firstly presented an overview of this thesis. It then went on to briefly outline the definition, prevalence, and impact of challenging behaviour. Some consideration was given to explanatory models and risk factors for challenging behaviour, before an outline of treatments was presented, including health-based assessment and treatment units, specialist behavioural teams, restrictive interventions, and behavioural interventions including PBS.

2 Positive Behaviour Support

2.1 Chapter Outline

This chapter will outline the development of Positive Behaviour Support (PBS) from its roots in Applied Behaviour Analysis (ABA) to becoming a fully-formed model in its own right. The development will be traced through a range of studies which attempted to define and describe PBS over a 25 year period. From these studies a number of potential defining features of PBS will be identified; these will then be assessed against a range of papers considering the definition of PBS in order to check their accuracy in forming a composite and generally agreed definition of PBS. From this process 11 key features are specified which combine to make up PBS.

Following on from this definition of PBS, there is a review of the evidence for its effectiveness, summarising findings from a number of literature reviews and meta-analyses, as well as individual studies with a range of participant group sizes. Based on the evidence in the published literature, some observations are made about the effectiveness of PBS and also about the enablers of effective PBS, that is, the factors that are more likely to facilitate a positive outcome from PBS intervention. Finally, some implications for the future direction of PBS research are discussed, specifically in relation to quality of life outcomes.

2.2 Development and Definition

In this section the development of PBS is traced via a number of studies which defined and described it over a 25 year period. The key studies discussed are summarised in Table 2.1 on p.19.

PBS came to the fore in the 1980s and 1990s, and is widely regarded as emerging from the debate in this period about the use of aversive techniques (e.g. Allen et al, 2005; Dunlap & Carr, 2007; Singer & Wang, 2009). In this context aversive techniques came to mean more than the technical behavioural meaning of aversive, but became synonymous with approaches that involved painful, degrading or extremely unpleasant events. The formation of PBS was a reflection of the development of the rights movement and values-based approaches to supporting people with learning disabilities such as Person Centred Planning (PCP) and the inclusion movement (Kincaid, 1996; O'Brien & O'Brien, 2002). The principles and values of these movements increased the momentum to use alternative behaviour management approaches which did not use aversive techniques. The move towards de-institutionalisation also demonstrated some of the limitations of traditional ABA, as it became clear that aversive behavioural practices were at odds with the new value-base. In this sense, PBS was created out of moral concerns and has at its heart a commitment to improving the quality of life and social roles of vulnerable people (Singer & Wang, 2009).

In 1990, Horner et al published what is the first clear definition of PBS, and this is generally regarded as the first use of the term 'Positive Behaviour Support' in the literature (Allen et al,

2005; Dunlap et al, 2009; Dunlap et al, 2014; Johnston et al, 2006). However at this point they also called it ‘non-aversive behavioural support’, although they noted that the non-aversive nature of PBS is not in fact the most important aspect; rather it is the positive intervention strategies which are more innovative and important. They identified nine main characteristics: an emphasis on lifestyle change; being based on functional analysis; utilising multi-component interventions; focusing on ecological manipulation; emphasising antecedent control; teaching new skills/behaviours; building environments with effective reinforcement; minimising use of aversive approaches; and utilising proactive and reactive approaches. Although this was the first time that an attempt had been made at a definition of PBS, earlier work had also referred to many of these nine characteristics (Carr et al, 1988; Donnellan et al, 1985; Emerson & McGill, 1989; La Vigna & Donnellan, 1986; La Vigna et al, 1989), and in fact the authors noted that ‘the basic concepts are being promoted from several different perspectives’ (p.125), mentioning functional communication training, positive programming, and gentle teaching as some examples of this.

However, although this was the first attempt to define PBS, these nine characteristics could also be said to be true of ABA, certainly in terms of its theory (e.g. Baer et al, 1968), although there are indications that common ABA practice at the time did not reflect all of these features. For example, in a meta-analysis of ABA (Didden et al, 1997) which considered 482 studies reporting on treatment for challenging behaviour from 1968-1994, lifestyle changes and teaching adaptive behaviours were not reported in the analysis, and only 8% of interventions were antecedent-based compared to 57% being consequence-based. In addition there was a high percentage (24%) of studies which reported use of aversive techniques (see Table 2.3 for comparison figures of studies reporting on use of aversives).

Didden et al (1997) also referred to the aversives debate and voiced concern that decisions may be made about interventions for challenging behaviour on the basis of how acceptable they are, rather than how effective; this could be seen as a comment on PBS and its clear statement of avoiding extreme aversives and instead using only respectful, positive strategies (Bambara et al, 1994; Horner et al, 1990; La Vigna et al, 1989).

To some extent Horner et al’s definition focused more on the behavioural technology side of PBS where it overlapped more obviously with ABA, rather than the values side. However, throughout the 1990s, the emphasis would change with more of a focus on defining PBS through values than technology. For example, Anderson and colleagues’ (1993) description of PBS training indicated that PBS was evolving. Their description of PBS included a move away from more specialist-led approaches, and a movement towards approaches such as PCP (Kincaid, 1996) which involved a range of stakeholders in the process of developing behaviour support plans for individuals. This included family, friends, direct care staff and other professionals, as well as potentially the person themselves, and was a clear move away from the more

'expert-based' approaches of ABA. One of the results of this inclusion of stakeholders is that functional assessment was widened out to be something undertaken by staff teams and those closest to the individual, rather than by external specialists; assessment also moved from controlled settings such as the laboratory to more ordinary settings in the community. Logically there was also therefore less of an emphasis on carrying out actual functional analysis, but rather using staff report, direct observation and other more flexible methods.

Bambara et al (1994) continued the idea of PBS emerging from a range of previous work in their description of it as 'an emerging multi-component approach' (taken from the article title). Their article attempted to introduce PBS to school psychologists and outlined the approach both conceptually and in terms of the framework of interventions; the emphasis was on how psychologists can support teachers and schools to introduce the approach. At this stage, as with Horner et al (1990), there were still some elements missing of what would come to be the full PBS definition, for example, there was no reference to organisational change, a future key component of PBS.

In 1996, Koegel et al presented further evidence of the evolution of PBS. In the preface to the book, *'Positive Behavioral Support: Including People with Difficult Behavior in the Community'* the editors described their vision of PBS and noted that as well as being founded on functional assessment, and utilising multiple intervention strategies, PBS was fundamentally based on person-centred values, and was committed to outcomes which were meaningful to the person. They specifically noted that interventions must not be demeaning or humiliating or painful, but should support development of a full, participative life. These characteristics were repeated by various contributors throughout the book, with an emphasis on introducing the approach to families, in schools and in developing inclusive communities. Commenting on the first section of the book, Haring & De Vault (1996, p.115-120) described PBS as interventions that consider the context of the behaviour, that address the function of the behaviour, that can be justified by the outcomes, and that are acceptable to the individual and those around them. This notion of the social validity of interventions became a key feature of the PBS definition, and within this volume the chapter on 'goodness of fit' outlined this concept in detail (Albin et al, 1996, p.81-98). Goodness of fit, or contextual fit, is the compatibility between the components of the plan and characteristics of both the people who will implement it and the environment in which it will be implemented; when contextual fit is high, the support plan is consistent with the values, skills and resources of those implementing it.

These developments in PBS and its attempt to increasingly distinguish itself from ABA were reflected in a variety of publications over the next few years (e.g. Anderson & Freeman, 2000; Carr et al, 1999; Dunlap & Horner, 2006; La Vigna & Willis, 1996). For example, in 1999, Carr et al published their research synthesis on PBS. They emphasised the need to recognise that in order to change problem behaviours, change to poor environments first needed to occur; they

also referred to increases in positive behaviour and lifestyle change as the primary focus of PBS, with a side effect of decrease in problem behaviours. The authors noted the importance of systems change as an essential element of PBS, both in terms of broad organisational changes and also individual support staff changing their behaviour. They also emphasised social validity, and noted that unless the primary stakeholders judge the intervention to be worthwhile, then it is unlikely to succeed. They suggested questions such as: Would you be able to use this intervention? Would you be willing to use this intervention? Does the intervention reduce the problem behaviour to a level that you find acceptable? And, does the intervention make a difference in the life of the individual? (p.19) However, arguably, social validity had long been an element of ABA; back in 1978 Wolf had called for ABA to begin to pay attention to social validity more systematically. He suggested that behavioural interventions be evaluated as to the social significance of the goals, that is are the goals really what consumers want; the social appropriateness of the procedures, that is do others around the person consider the interventions appropriate; and the social importance of the effects, that is whether consumers are satisfied with the outcomes of the interventions (p.207). These are very similar to the questions suggested by Carr et al (1999) 20 years later.

Anderson & Freeman (2000) also compared PBS with ABA and refer to PBS as ‘expanding the application of ABA’ (taken from the article title). Their summary of the development taking place was that ‘ABA provides the technology with which to teach skills, whereas PBS suggests the areas in which the technology should be applied’ (p 91), echoing the theme of PBS being a development or evolution from ABA which is referred to by many authors (Allen et al, 2005; Buschbacher & Fox, 2003; Carr et al, 2002; La Vigna & Willis, 2005a; La Vigna & Willis, 1996).

In 2000, Horner chose not to continue the PBS/ABA comparison stating that there was no merit in comparing PBS with conventional behaviour management, since there was no difference in either values or technology. However, he went on to say that if there was a difference it is that PBS has a focus on much larger outcomes for the lives of the individuals involved – which in turn caused an expansion of the behavioural technology in order to meet these wider lifestyle outcomes. Some examples of this were the necessity for a multi-intervention approach across a range of environments and contexts relevant for the individual; the commitment to contextual fit, that is, that the strategies must work for all the individuals involved, including other service users, staff and family, and must also have social validity for the mediators implementing them; and the requirement to apply PBS within larger settings and systems. Horner et al (2000) also gave guidance about PBS plans offering a detailed description of their content, format and purpose.

In 2002 Carr et al produced what is now widely regarded as the seminal definition article where the additional characteristics of PBS were clearly specified and defined. Their nine characteristics added to the more technology-based definition of Horner et al (1990), and although they referred to functional analysis and ABA technology, the focus within the definition was on values: lifespan perspective or a long-term focus; ecological validity and the use of PBS in real-life situations; stakeholder participation, with stakeholders providing valuable perspectives; social validity, emphasising good fit of interventions and acceptability of interventions to stakeholders; systems change in order to allow behavioural change to occur; emphasis on prevention and proactive approaches; flexibility with respect to scientific practices and willingness to utilise other theoretical perspectives. Although many of these elements, such as social validity and stakeholder views, long-term focus, and the need for system-wide change, were present in ABA literature (Baer et al, 1968; Baer & Wolf 1987; Wolf, 1978), their emphasis, and the priority given to them by Carr et al (2002), moved PBS into a more obviously separate category from ABA. As the authors put it, PBS ‘evolved beyond the parent discipline to assume its own identity’ (p.5).

However, some continued to argue that ABA and PBS were one and the same, and that ABA met many of these characteristics also. For example, J Carr & Sidener (2002) outlined eight features of PBS taken from a review of the literature at the time, although they also noted that there were a variety of definitions of PBS, and in particular that there were different ways of conceptualising its relationship with ABA. The features that they found commonly described were: PCP, functional assessment, positive non-aversive interventions, multi-component interventions, environmental alterations, meaningful outcomes linked to quality of life, ecological validity, and systems-level interventions. They argued that each of these could also be found in ABA, and noted that those who supported the idea of PBS as a separate discipline may have been influenced by the perceived failure of ABA to address the issues that they most valued; a reflection perhaps that the theoretical values of ABA were not reflected in the literature reporting on outcomes of ABA intervention.

Johnston et al (2006) also disagreed with the description of PBS as a new approach, separate from ABA, saying ‘it is clear that PBS emphasises certain values...these values...have long been accepted by diverse professions, including ABA’ (p.54). However they went on to note that whereas PBS puts values first rather than efficacy, ABA regards the effectiveness of interventions as more important than values about whether these interventions should be used; which could be regarded as a clear indication of a fundamental difference between the two models. Although they acknowledged that ‘PBS evolved from and remains similar to ABA in certain ways’ (p. 59), they characterised PBS as a watered-down or less technical version of ABA, more marketable to service providers without ABA training as it requires less technical expertise, which is less robustly evaluated and involves less vigorous training. They also criticised

what they felt to be the dishonest representation of PBS by its proponents, who refused to acknowledge the debt to ABA and in fact portrayed PBS as more forward-thinking, having better values, being more flexible and more proactive than ABA.

However others have argued that with its emphasis on values, systems change, contextual fit and the role of family, PBS makes a unique contribution to ABA. Filter (2007) suggested that PBS had made behaviour analysis much more accessible to the general population by embedding it in a service delivery model, a view also proposed by Wacker & Berg (2002) who argued that PBS could not be a science as it is not based on scientific methods. They stated that the blending of the science of ABA with the values of PCP and inclusion, did not produce a new science but a new service delivery system. Dunlap et al (2008) tried to address some of the misunderstanding which they felt had led to ‘unnecessary counterproductive polarisation within the field’ (p.684). Their take on the controversy between supporters of ABA and supporters of PBS was that although the individual elements of PBS were not new, the way that they had been taken from a variety of models and put together into a different model was new, in particular in the way that PBS had evolved from the parent approach of ABA to become a new discipline. They noted that PBS and ABA clearly have many areas of overlap, but although PBS was exploring new and additional ideas, its theoretical core came from ABA. Similar to Filter (2007), they suggested that PBS was an expansion of ABA, through the use of other methods and models, for example school-wide interventions; ‘PBS has taken the indispensable foundations of ABA and extended them to applications in community-based settings’ (p. 693).

Moving on from the controversy about whether PBS is part of ABA or a new discipline, other definitions have followed since Carr et al (2002): Allen et al (2005) provided a PBS definition, based on a summary of the current relevant literature, including Horner et al (1990) and Carr et al (1999). They traced the history of PBS from its link with ABA, noting that although ABA offered a menu of non-aversive options, it is aversive techniques which dominated the literature until the 1990s. They also highlighted the importance of not losing the behavioural science contained in ABA, as without this PBS is only a values-based approach without a technology to put the values into practice. Also in 2005, La Vigna & Willis described their PBS model which outlined a range of proactive and reactive strategies as well as being based on functional assessment and taking into account both the views and the skills of the mediators of the plans (2005a).

In Dunlap & Carr (2007), the authors outlined the characteristics of PBS, referring to its scientific approach, rigorous validation and accountability to evaluation; and at the same time its openness to innovation and the use of other approaches, and its grounding in values, particularly a person-centred approach and family-based perspective. They noted the foundations of PBS both empirically and conceptually have much in common with those of ABA. Other elements

which featured in their definition include functional analysis; contextual fit; quality of life focus; non-aversive nature; enriching deficient environments; teaching new skills/behaviours; use of reinforcement and antecedent manipulation; and the need for PBS to be a systems-wide approach, although they noted that at that point there were few evaluations of large-scale PBS.

Around this time, debates about the origins of PBS became less common in the literature and there was more of a focus on how to implement PBS, rather than describing where it had come from. For example Allen (2009) described how PBS could be implemented as a service system, a point returned to in Allen et al (2013b), which focused on the contexts where PBS takes place and identified organisational factors more likely to support the implementation of PBS. This focus on widespread and systematic implementation is noted by some to be a defining element of PBS (Carr, 2007; Dunlap et al, 2008). La Vigna & Willis (2012) also moved on from the origins debate and addressed the issue of the effectiveness of PBS, in addition to costs and accessibility.

Dunlap et al (2009) in *The Handbook of Positive Behaviour Support* (Sailor, et al, 2008, p. 3-16) described how PBS could be used in families, schools, wider communities and with a variety of client groups such as mental health services, foster care, and youth justice. In their introduction they defined the core features of PBS as: a behavioural science which uses functional assessment, environmental change to prevent problem behaviours, teaching of adaptive behaviours and the use of consequences to promote behaviour; the integration of multiple intervention elements, effective in a range of contexts; commitment to durable lifestyle outcomes, the quality of life outcomes for the individual are the measure of success; and organisational systems that facilitate sustained effect. In a later chapter of the same book, while tracing the intellectual roots of PBS, Singer & Wang (2009, p. 17-46) noted that in PBS functional assessment goes beyond the functional relationship that behaviour has with antecedents and consequences, and recognises the importance of understanding the cultural and social contexts of the behaviour. They noted that with this understanding ‘the possibility of empathy and recognition of fellow humanity of people who otherwise seem threateningly strange, is more likely’ (p. 27), and that therefore an increased imperative exists to find a way to make the benefits of full inclusion in community life accessible to those individuals who have previously been denied them.

The most recent definition of PBS came from Gore et al (2013) who defined PBS in terms of values, theory and evidence base, and process. They listed ten elements which they noted were both overlapping and were not a menu of options, but that all of the elements needed to be in place for the approach to be truly PBS. Elements described in the values category were:

- reduction of challenging behaviour occurs within the context of increased quality of life, including inclusion, participation and valued social roles
- approaches to intervention should help the individual build their skills and adaptive behaviours, and reject aversive or restrictive approaches

- stakeholder input is essential both to determine priorities for intervention and also to inform the whole PBS process

Elements described in the theory and evidence base category were:

- an understanding that challenging behaviour is functional
- the use of ABA as the primary approach to address behaviour change
- the secondary use of other complementary approaches, for example, mental health interventions

In the process category the elements described were:

- a data-driven approach at every stage of PBS, including assessment, planning and implementation
- functional assessment as the basis of interventions, in order to ensure that all interventions are based on an understanding of the behaviour within its context and social environment
- the use of multiple interventions, both proactive to change behaviour over time, and reactive to manage behaviour when it occurs
- the need for monitoring and evaluation of PBS implementation in order to ensure interventions remain effective over time

Although the Gore et al (2013) article can be regarded as the most recent attempt to define PBS, the following year Dunlap et al (2014) published a short essay on the term ‘Positive Behaviour Support’, and their preference for this term over others such as ‘Positive Behavioural Support’ or ‘Positive Behavioural Interventions and Support’. They restate their understanding of PBS, which is very similar to that in other works by Dunlap (e.g. Dunlap & Carr, 2007; Dunlap et al, 2009); however they do also call for ‘a more extended consideration of the PBS definition’ (p. 136) as there is potentially some confusion about the variation of terminology and the different categories of PBS.

Based on this discussion of the literature, a number of key features of PBS can be seen as occurring regularly in the range of studies which have attempted to define and describe PBS, from Horner et al (1990) to Gore et al (2013), albeit at times with some slight differences in the definition and explanation of each. While it is acknowledged that there is considerable overlap between ABA and PBS, it is felt that PBS has sufficiently evolved into a different model to be defined separately. These features and their occurrence in the relevant literature are outlined in Table 2.1, listed from the most commonly used to the least used. Out of the 13 features referred to by 17 studies and commentaries which have attempted to define PBS, there are 11 key features which occur in more than half the studies; for the purposes of the current study, PBS was therefore defined as being made up of these 11 features. These 11 key defining features are outlined in more detail in Table 2.2, with brief explanations of each, which are a synthesis of their descriptions in a range of studies.

After Carr et al (2002), the majority of PBS writings reflect these 11 features, with some small exceptions. For example, Horner (2000), Horner et al (2000), and La Vigna & Willis (2012) do not specifically include flexibility in scientific practice as part of their definition, and some writers do not specify the need for systems or organisational change (Anderson & Freeman, 2000; Bambara et al, 1994; Horner et al, 1990). Another exception is Carr, whose writings on PBS make little direct reference to reactive strategies (Carr et al, 1999; Carr et al, 2002; Dunlap & Carr, 2007); the emphasis in his descriptions of PBS is very much on proactive approaches, so much so that he says ‘the proactive nature of PBS stands in sharp contrast to traditional approaches’ (2002, p9). La Vigna and colleagues on the other hand, emphasise the need for reactive as well as proactive approaches in their multi-element model (La Vigna & Willis, 2012; La Vigna & Willis, 2005a), as do Allen et al (2005) and Allen (2009), and when the UK-based International Journal of Positive Behavioural Support was launched in 2011, the editors specified that the journal would welcome contributions on reactive strategies due to ‘the clear need to comprehensively link together what has until now been two separate literatures’ (Baker & Allen, 2011, p. 5). However, despite these small differences, there appears to be broad agreement in the defining features of PBS.

Table 2.1: Studies used to define PBS

Features of PBS	Horner et al 1990	Bambara et al 1994	Koegel et al 1996	Carr et al 1999	Horner 2000	Anderson & Freeman 2000	Horner et al 2000	Lucyshyn et al 2002	Carr et al 2002	Allen et al 2005	La Vigna & Willis 2005a	Dunlap & Carr 2007	Allen 2009	Dunlap et al 2009	La Vigna & Willis 2012	Gore et al 2013	Dunlap et al 2014
Emphasis on lifestyle change/quality of life	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Altering deficient environments	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Minimal use of aversive interventions	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Long-term focus	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Multi-element interventions	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Based on functional analysis	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Social validity/contextual fit	X	X	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Flexibility in scientific practice	✓	✓	✓	✓	X	✓	X	✓	✓	✓	✓	✓	✓	✓	X	✓	✓
Systems/organisational change	X	X	✓	✓	✓	X	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Proactive and reactive interventions	✓	✓	✓	X	✓	✓	✓	X	X	✓	✓	X	✓	X	✓	✓	X
Ecological validity	X	X	✓	✓	X	X	✓	✓	✓	✓	✓	✓	✓	✓	X	✓	X
Data driven	X	X	✓	✓	X	X	✓	✓	X	✓	✓	✓	X	X	✓	✓	X
Person Centred Planning	X	X	✓	X	X	✓	✓	✓	X	X	✓	✓	✓	X	X	X	X

Table 2.2: Key Features Defining PBS

<p>1. An emphasis on lifestyle change/quality of life</p> <ul style="list-style-type: none"> Improving quality of life is a starting point and not dependent on behaviour – it is the primary goal of PBS; this includes improved relationships, increased community integration, greater self-confidence, increased choice and a fuller range of opportunities A commitment to inclusion and normalisation values, including PCP
<p>2. Altering deficient environments/ecological changes</p> <ul style="list-style-type: none"> A belief that challenging behaviours are a product of the person’s environment and that many people with learning disabilities live within barren, un-stimulating and difficult environments; PBS therefore incorporates broad and lasting changes to the environment in order to address challenging behaviour A wide range of variables is addressed – e.g. relationships, employment, activities, leisure, skills, staffing, diet, routines, physical environment, instructional methods
<p>3. Use of interventions which are justifiable and minimally aversive</p> <ul style="list-style-type: none"> The level of intrusiveness of intervention must be justifiable and comparable to interventions/treatment offered to non-disabled individuals Minimal or no use of aversives, and rejection of those which cause actual pain, result in physical harm, or do not support the dignity of the individual
<p>4. Long-term focus</p> <ul style="list-style-type: none"> A belief that as challenging behaviours are often long-term behaviours, interventions need to address long-term changes; intervention is seen as a never-ending, evolving process Aiming for generalised and lasting changes in behaviour
<p>5. Multi-element interventions</p> <ul style="list-style-type: none"> Simultaneous manipulation of many variables – as challenging behaviour is often maintained by multiple variables Interventions include antecedent manipulation, teaching adaptive behaviours, use of reinforcement and reactive strategies
<p>6. Based on functional assessment</p> <ul style="list-style-type: none"> Functional assessment is used in order to understand what purpose behaviours serve, what events are likely to maintain the behaviour Ensuring a direct link between functional assessment and interventions
<p>7. Social validity/contextual fit</p> <ul style="list-style-type: none"> Family, friends, co-workers, other professionals, and the individual must inform the intervention process, agree strategies and evaluate effectiveness Interventions must make sense for the values, skills, resources and social context of these stakeholders
<p>8. Flexibility in scientific practice/ use of other approaches secondary to ABA</p> <ul style="list-style-type: none"> More flexible practices, e.g. descriptive functional assessment and use of observation, interviews and self-report, rather than formal functional analysis Acceptance of other theoretical perspectives and use of approaches complementary to ABA which are used as secondary to ABA technology
<p>9. Systems/organisational change</p> <ul style="list-style-type: none"> A belief that meaningful and lasting change is only possible if organisations and systems around people are changed A good PBS plan is not enough – organisational systems must be introduced to support this and ensure its implementation
<p>10. Necessity of proactive and reactive interventions</p> <ul style="list-style-type: none"> Use of reactive interventions with a short-term emphasis, to manage behaviours in crisis or emergency situations Use of proactive interventions with a long-term emphasis, to change behaviour over time
<p>11. Ecological validity</p> <ul style="list-style-type: none"> Involvement of ‘typical agents’ in implementing the intervention, i.e. family and ordinary support staff, rather than researchers or behaviour specialists Use of natural settings to implement the intervention, e.g. home or community, rather than wards, specialist units or institutions

2.3 Evidence for Effectiveness

PBS has gained widespread acceptance over the past 25 years as an effective and appropriate response to challenging behaviour in individuals with learning disabilities. It has been used with a wide range of client groups including people with neurological conditions (Gardner et al, 2003; Rothwell et al, 1999), young people with emotional and behavioural difficulties (La Vigna et al, 2005) and in mainstream schools (Sugai et al, 2000). However as the focus of this study is people with learning disabilities, then that is the focus of the following review of the effectiveness of PBS. PBS training will be considered in a later chapter, so outcomes for service users in relation to training inputs (or even PBS interventions introduced via training), will not be considered in this chapter. There is extensive literature in relation to school-wide PBS, particularly from America, but as most of this is in relation to children without a learning disability, this literature is also not considered here.

From reviewing the literature, it is clear that different studies have adopted different approaches to evaluation of the effectiveness of PBS. This includes differences in criteria for including studies in literature reviews; a variety of methods of evaluating outcomes; and differences in the way outcomes are described. With this lack of consistent approach it is difficult to directly compare the findings; instead below there is a summary of the literature with the relevant evaluation findings reported for each. Literature reviewed in this chapter has primarily focused on literature reviews which have brought together a range of studies. In addition, a range of individual studies which highlight particular aspects of PBS are selected for further consideration. This part of the chapter concludes with a summary of PBS effectiveness and some reflections with regard to potential enablers in the successful implementation of PBS, based on the evidence in the literature.

2.3.1 Meta-Analyses and Literature Reviews

A search of the literature was made for studies reviewing effectiveness of PBS for people with learning disabilities; although this was not a systematic review, similar principles were followed. Searches took place using the terms “Positive Behav* Support” and “effectiveness”, and “Positive Behav* Support” and “review”. From the reviews that were found, reference lists were also studied in order to search for other similar studies. Additional searches were then made of the content lists of the Journal of Positive Behavior Interventions (JPBI) and the International Journal of Positive Behavioural Support (IJPBS), as these are the two PBS-specific journals.

Many of the reviews of behavioural interventions with people with learning disabilities do not focus specifically on PBS (e.g. Didden et al, 2006; Matson & Gorman-Smith, 1986; Scotti et al, 1991) and therefore studies were not included in this review unless they specifically referred

to PBS in the method. The exceptions to this were any studies published in the JPBI, or in the IJPBS, where the reference to PBS was assumed. Seven studies were found which reviewed PBS research, and data from these are summarised in Table 2.3. Not all of these studies reported on outcomes in terms of reduction of challenging behaviour (Clarke & Dunlap, 2008; Dunlap et al, 1999); however they contain other useful information about the implementation of PBS and therefore are still included in this review of evidence for its effectiveness.

The most comprehensive study of PBS effectiveness remains the research synthesis carried out by Carr et al in 1999. Standards for inclusion in their analysis were that studies were published between 1985 and 1996, in English, peer-reviewed, related to learning disability or autism, addressed self-injury aggression, property destruction or ‘tantrums’, and included either stimulus-based or reinforcement-based interventions. Significantly they included studies that did not report use of functional analysis or functional assessment, a fundamental part of the PBS approach. They also included studies only using one intervention, so the multi-component approach of PBS was another key element not considered essential. The analysis included 109 articles, relating to 230 individuals and 366 different outcomes for these individuals; studies only using an AB design were excluded, as were those with fewer than three baseline points or fewer than three intervention points. Effect on challenging behaviour was measured as percentage reduction from baseline, with success being considered 90% reduction from baseline. This found that 51.6% of the outcomes achieved a 90% reduction from baseline; using 80% reduction from baseline, 68% of outcomes achieved this. Using 90% reduction as success, studies reporting use of a functional assessment had a success rate of 59%, whereas those not using any functional assessment had a success rate of 32%. In this study, the authors class reactive strategies as non-PBS; they report that for 71.1% of the outcomes, PBS procedures were used alone, with no reactive strategies. Despite lifestyle changes being an essential element of PBS and perhaps the most significant feature that distinguishes it from ABA, they found that it was a stated intervention goal for only 10.4% of the sample, was included as an intervention strategy in 3.5%, and included data on impact on quality of life in only 2.6%. In addition, although measures of challenging behaviour were noted in all studies, changes in positive behaviour were only measured in 45%, thus somewhat undermining the claim that PBS has a significant focus on developing new behaviours and skills. They also noted that although there was an increase over time in short-term maintenance, no outcomes were tracked for more than two years, thus leaving unanswered the question of whether PBS can produce long-term change, which is also one the key aims of PBS. Other findings included: stimulus-based interventions were becoming more common than reinforcement-based; multi-component interventions were not becoming more common; use of functional assessment prior to intervention was becoming more common; punishment was used in 13.8% of outcomes. Greater effectiveness of PBS was associated with the use of functional assessment and with the use of typical agents, that is, people who normally

support the individual such as family or their ordinary support staff, rather than researchers or behaviour specialists (untypical agents).

Marquis et al (2000) used the same database as Carr et al (1999) to carry out their meta-analysis of PBS which provided an in-depth statistical analysis reporting on effect sizes to complement Carr et al's descriptive synthesis. They confirmed that effect sizes were larger when a functional assessment was carried out and the results used to design the intervention, and when typical agents and typical settings were used (as this study used the same data as Carr et al (1999), its data are not included in Table 2.3).

Also in 1999, Dunlap et al carried out a descriptive analysis of articles published in 10 journals between 1980 and 1997 for young people (up to age 21) with learning disabilities and behavioural difficulties. Although the studies were not described as PBS-specific, this was published in the JPBI and was therefore included. In addition, the interventions used and the variables studied would meet the PBS criteria as well as those included in the Carr et al (1999) review. There are a number of interesting findings which are relevant to this review. They split their findings into six-year categories and thus were able to track change in the research over this time, thus noting an increase in assessment-based interventions (although still only an average of 11% of studies reported using assessment), and these were not specified as functional behavioural assessments. The majority of intervention agents were untypical and this remained relatively constant, with an average of 63% of studies using untypical agents; in particular parents were very unlikely to be intervention agents (under 5%). However research in typical settings was increasing and ecological validity had a mean of around 25% (this figure is an average of the three types of ecological validity measured: physical, social and activity context). There was also an increase in social validity with an average of 14% of studies reporting on this; reporting of maintenance and generalisation data were at 40% and 21% respectively.

Table 2.3: Reviews of PBS Research

Review Details	Carr et al, 1999	Dunlap et al, 1999 *	Snell et al, 2005	Clarke & Dunlap, 2008 **	O'Dell et al, 2011*	La Vigna & Willis, 2012	Goh & Bambara, 2012
Dates of review	1985-1996	1980-1997	1997-2002	1999-2005	1999-2008	Not specified	1997-2008
Type of review	Research synthesis	Descriptive analysis	Descriptive analysis	Descriptive analysis	Descriptive analysis	Literature review	Meta-analysis
Number of studies	109	1197	111	59	118	12	83
Number of participants	230	-	-	-	-	423	145
Number of outcomes	366	-	-	-	-	-	
Type of participant	Mixed age group, all learning disability (LD)	21 years and under	School-age	21 years and under	Mixed age group, 48% LD or autism	Mixed age group	School-age, 48% LD
Any use of FA %	73	11	100 ***	54 (45)	-	100 ***	100***
Experimental	29	-	53	-			21
Descriptive	8	-	18	-			42
Both	36	-	29	-			37
Intervention agent %:							
Typical	44	22	83	25 (6)	52	-	83

Review Details	Carr et al, 1999	Dunlap et al, 1999 *	Snell et al, 2005	Clarke & Dunlap, 2008 **	O'Dell et al, 2011*	La Vigna & Willis, 2012	Goh & Bambara, 2012
Non-typical	56	63	95	47 (90)	48	-	17
Intervention setting %:							
Typical	35	25	82	67 (22)	86	-	82
Non-typical	65	-	45	-	14	-	18
Intervention type %:							
Stimulus-based	42	-	-	-	-	-	-
Reinforcement-based	46	-	-	-	-	-	-
Antecedent	-	31	65	44 (46)	69 (includes skills)	-	23
Consequence	-	32	78	20 (70)	-	-	15
Skills training	-	44	49	41(25)	16	-	17
Self-management	-	7	7	17 (1)	-	-	-
Multi-component (2 or more)	11	29	78	29 (41)	48	100***	46
Use of punishment %	14	-	10	-	-	0 ***	-
Social validity reported %	7	14	22	31 (3)	33	-	39

Review Details	Carr et al, 1999	Dunlap et al, 1999 *	Snell et al, 2005	Clarke & Dunlap, 2008 **	O'Dell et al, 2011*	La Vigna & Willis, 2012	Goh & Bambara, 2012
Quality of life reported %	3	-	2	-	52	-	1
Adaptive behaviour %:							
Skills	-	57	55	58 (38)	-	-	-
Engagement	-	17		37 (24)	-	-	-
Appropriate behaviour	-	-		-	-	-	52
Social interaction	-	16		15 (6)	-	-	-
Non-specified positive	45	-		-	-	-	-
Effectiveness for challenging behaviour	90% reduction from baseline – for 52% of outcomes; 80% reduction – for 68% of outcomes	-	97% of studies reported a reduction (level not specified)	-	-	70% reduction reported in one study; to 22% of baseline in another. Severity reduced	PND 85% - moderate effect
Reported generalisation %	7	21	26	34 (6)		-	7
Reported maintenance %	41	40		44 (13)	31	One study 20 years and another of 2 years data.	20
Reported data post 6 months %	14	12	13	15 (4)	5		-

Review Details	Carr et al, 1999	Dunlap et al, 1999 *	Snell et al, 2005	Clarke & Dunlap, 2008 **	O'Dell et al, 2011*	La Vigna & Willis, 2012	Goh & Bambara, 2012
Other Info	Typical agents – 66% success; non-typical – 44% success. Using FA – 59% success; not using FA – 42 % success		Generalisation more common in natural settings		Intervention in schools becoming more common; 52% of studies included systems change.	PBS is not more expensive than traditional approaches. PBS does not require highly trained specialists.	Antecedent and skills training interventions had more impact than consequence-based and multi-component interventions

* Dunlap et al, 1999 & O'Dell et al, 2011 are not PBS-specific reviews; ** Clarke & Dunlap, 2008 compared data from Journal of Positive Behavior Interventions with data from Journal of Applied Behavior Analysis (latter is in brackets); *** Indicates that this feature was part of the criteria for inclusion in the review.

Snell et al (2005) carried out a descriptive analysis of 111 studies published between 1997 and 2002 on school-age individuals with disabilities in order to determine how often certain key aspects of PBS were being used. Their criteria were similar to Carr et al (1999), but in addition stipulated that the study must report on interventions based on either functional assessment or functional analysis¹, so this was perhaps a more pure-PBS study. They found that consequence-based interventions were most widely used; 10% of the studies reported using punishment, a similar level to Carr et al's findings and 22% reported on social validity, in comparison with an average of 7% (depending on the different social validity questions asked) in Carr et al (1999). In terms of functional assessment and analysis, functional assessment alone was used in 18%; most frequently this included interview and recording of antecedents, behaviour and consequences, but other methods included videoed observations and reviewing records. Functional analysis alone was used by 53%, and both functional assessment and functional analysis were used by 29%; there were a number of clear differences between studies that used functional assessment from those using functional analysis. Studies using functional analysis were more likely to use atypical settings for assessment and interventions, less likely to involve family members or teachers and more likely to use fewer interventions. In terms of outcomes, 97% reported reduction in challenging behaviours, 55% reported increase in positive behaviours, 26% on generalisation, and 13% on maintenance. Studies carried out in natural settings were more likely to report on generalisation; they were also more likely to involve the team in assessment and implementation. The authors noted that involving stakeholders, which is an important part of PBS was missing in many of the studies, and therefore the element of contextual fit was also missing.

In 2008 Clarke & Dunlap produced a descriptive analysis of intervention research published in the JPBI from 1999 to 2005, and compared this with research published in two other journals, the Journal of Applied Behavior Analysis (JABA) and Education and Training in Mental Retardation and Developmental Disabilities (ETMRDD) over the same period. Participants had to have a disability and be aged 21 or under, and the focus of the analysis was to compare the JPBI publication record with the other two journals, in relation to research reflecting the major features of PBS, specifically ecological and social validity, and the use of assessment-based interventions. In relation to ecological validity, this was scored in relation to typical physical, activity and social contexts; JPBI had higher rates of ecological validity than the other two journals. JPBI also had higher rates of articles with measures for social validity, 31% as compared to

¹ Functional assessment is based on descriptive methods such as observation, records review, informant-based questionnaires and rating scales; its purpose is to gain an in-depth understanding of the context of the challenging behaviour and to generate a hypothesis about the function. Functional analysis is the use of experimental methods to test hypothesis about the function served by the challenging behaviour (Emerson & Einfeld, 2011, chapter 6).

3% for JABA and 20% for ETMRDD, thus indicating that ecological and social validity, which are key elements of PBS, are more evident in the journal devoted to PBS research than they are in other comparable journals. In relation to assessment-based intervention, the authors found that 54% of JPBI's articles were assessment-based, while this was only true of 45% of JABA and 22% of ETMRDD articles. Other findings included were in relation to maintenance and generalisation data; 44% of studies included maintenance data and 34% generalisation data, these percentages were considerably higher than in JABA, but comparative to ETMRDD. However, articles containing data after six months of intervention accounted for 15% of the JPBI research and only between 2-4% of the other two journals. The authors concluded from these data that JPBI was meeting its purpose of publishing research reflecting many of the key features of PBS.

In 2011 O'Dell et al provided an analysis of the literature published in JPBI over the past 10 years with a view to exploring how PBS was being practiced and whether the research reflected good principles of PBS. They reported on 118 studies and found that the majority were conducted in typical settings, and by typical agents, 31% reported maintenance data, and multi-component interventions were present in nearly half of the studies. There was a substantial emphasis on prevention, via the use of antecedent strategies or teaching new skills to prevent challenging behaviour occurring. Interventions including quality of life changes were included in the majority; however they did not report on effectiveness of PBS.

La Vigna & Willis (2012) carried out a literature review in relation to the effectiveness of PBS. They addressed a number of specific questions: is PBS effective with severe challenging behaviour; is PBS effective with high-rate behaviour; does PBS require specialist expertise; is PBS expensive to implement; and is PBS effective in institutions. The authors noted that they were not addressing whether PBS is effective in terms of quality of life changes, an interesting omission, since quality of life is a key feature of PBS. Their criteria for inclusion were that studies had to report on multi-element interventions, rather than single variables; that any studies using aversives specifically as a punishment were excluded; and studies reporting on non-severe challenging behaviour were excluded. In that sense their criteria were arguably more PBS-specific than in any of the previous literature reviews. Based on these criteria they selected 12 studies which were then reviewed in relation to the specified questions; these included some single-subject multiple baseline designs, as well as control group studies and also case studies were included if they met Type 3 criteria (Kazdin, 1981). The 12 studies include 423 individuals, of which data for 295 came from Type 3 case studies. If the studies involving PBS training rather than direct intervention are excluded, then there are 10 studies involving 225 individuals, from five countries (England, Scotland, Australia, US, Ireland); 97 of these individuals are from Type 3 case studies. In terms of outcomes, and excluding the training studies, the authors reference five studies that addressed very severe behaviours, for a total of 24 individuals (Donnellan

et al, 1985; La Vigna & Willis, 1992; La Vigna et al, 1989; MacDonald et al, 2010; McClean et al, 2007) and two studies that addressed high-rate behaviours (La Vigna & Willis, 1992; MacDonald et al, 2010). This appears to be a limited amount of evidence for these two critical factors in relation to effectiveness; however part of the reason may be the stricter criteria of only including studies using multi-element interventions and excluding those that used punishment. This does however, appear to indicate that when using these criteria (both which are arguably essential elements of PBS), then there is limited evidence of effectiveness for PBS in the published literature.

Goh & Bambara (2012) carried out a meta-analysis of individualised PBS in school settings. Although their criteria included participants without a learning disability, 48% of the participants had a learning disability, and therefore the study is included here; however this fact may mean that some of their findings are less directly applicable to the learning disability field. The criteria that they used were peer-reviewed articles with school-aged participants, using functional behaviour assessment (FBA)² and evaluating at least one function-based intervention, using single-participant research design with experimental control, that is, no AB designs, and with at least two data points in both baseline and intervention phases. They found 83 studies with 145 participants. They coded a number of features of the studies: participant characteristics; FBA characteristics, including whether it was experimental, descriptive or both, who conducted the assessment, where it was done, and whether the team were involved (team was defined as at least two individuals, involving at least one school-based employee, p. 274); and intervention characteristics, including type of intervention either skills training, antecedent-based, consequence-based, or multi-component, who implemented the intervention, where it was implemented, whether the team were involved, and if it had social validity. The method they used for analysing effectiveness was percentage of non-overlapping data points (PND), that is, calculating the percentage of intervention data that do not overlap with the lowest baseline data point (if the purpose of the intervention is to increase the measure, then the highest baseline data point would be used). PND scores were calculated individually by participant for reduction of challenging behaviour, increase in appropriate behaviour, and overall behaviour change. A large effect was 90% or greater, moderate 70%-90%; 50%-70% was a small effect; less than 50% was not effective.

The overall effectiveness of PBS in reducing challenging behaviour was a PND of 80%, a moderate effect and in increasing adaptive behaviour was 90%, on the border of moderate/large effect. They found that there was a moderate effect both for people with learning disabilities (85%), and those without (90%). In terms of FBA, descriptive measures were most common, in

² In this study, FBA is used as an over-arching term which includes functional assessment (the descriptive element of FBA) and functional analysis (the experimental element of FBA).

42% of the cases, whereas experimental methods alone were only used with 21% of participants, substantially lower than in Snell et al (2005) who reported functional analysis used alone by 53%. The majority of FBA were carried out by typical participants, involving school staff, and were carried out in a typical setting, for example, a classroom, so this can be seen as a positive in terms of utilising the PBS feature of flexibility in relation to scientific practice. However, most did not involve the team, although those that did, did not achieve any better effect sizes. In terms of interventions, multi-component interventions were more common than antecedent-based, skills training, or consequence-based, but there was no significant difference in effect size among the types of intervention. Most of the interventions used typical agents and were implemented in typical settings, although there was no significant difference from those that did not. Team involvement in decisions during intervention planning only took place in 39% of cases; an involvement of the team resulted in a significantly higher PND score compared to those that did not involve the team. However, social validity data were only collected in a minority of cases, and there was no significant difference in effect sizes between those cases that measured social validity and those that did not.

These findings by Goh & Bambara (2012) showed that PBS applied in school settings can reduce problem behaviour and increase adaptive behaviour. It was also clear that PBS can be used by ordinary staff in ordinary environments, without the need for specialist settings or intervention agents. However, the authors noted that few studies reported on generalisation, and even fewer assessed maintenance, with only four reporting maintenance beyond two months. Poor measures of generalisation were also mentioned in Carr et al (1999) and Snell et al (2005). Another issue raised is that only half of the studies which focused on reducing challenging behaviour also measured increase in adaptive behaviour; given that this is a key feature of PBS, this is somewhat surprising. In addition, only one study included lifestyle interventions, and this is also notable, given that it is regarded as such an important PBS outcome; this reflected findings in Carr et al (1999) and Snell et al (2005). However, a positive finding was that the majority of interventions were multi-component, as with Snell et al (2005), and the use of consequence-based interventions was reduced compared to Carr et al (1999).

In terms of functional assessment, Goh & Bambara (2012) found that there was no evidence of greater effect sizes with experimental functional assessment compared with descriptive. It is of note that the only area where statistically significant differences were found between effect sizes was in the team involvement in intervention planning, suggesting that team involvement is a factor in better results; as social validity is a key feature of PBS, this was an encouraging result. Snell et al (2005) reported a quarter of the team were involved in decision making, so this is an improvement; however the authors note that it was difficult to know exactly what this team involvement consisted of, and suggested that future research should document the process to allow for replication and further investigation

2.3.2 Individual Studies

There is now a substantial amount of published literature reporting outcomes from studies in PBS; there are both UK and US based PBS journals publishing a range of evidence from the past 20 years. As it is beyond the scope of this study to conduct a systematic review, a small number of studies have been selected for consideration which are illustrative of the range and variety of PBS literature. The studies were selected as they addressed a number of factors:

- All commentators agree that impact on quality of life is a major aim of PBS, therefore it was felt appropriate to highlight studies that specifically addressed this
- A concern sometimes raised regarding PBS is in relation to its efficacy (e.g. Johnstone et al, 2006), therefore studies that demonstrated impact on severe behaviours were included
- The need for PBS to be effective across settings is important, and therefore specialist health settings were specifically considered
- The need for PBS to be effective for families is important, and therefore family settings were specifically considered
- Specialist behavioural teams are a useful way of addressing challenging behaviour and achieving impact on larger amounts of people than individual interventions, therefore studies reporting on effectiveness of these teams are included
- Maintenance is a key factor in any behavioural intervention and particularly since research demonstrates that challenging behaviour is often a lasting issue (Taylor et al, 2011; Totsika et al, 2008), therefore studies with long-term data were included.

2.3.2.1 Quality of Life Outcomes

PBS interventions should have a substantial focus on quality of life since this is acknowledged as a key element of PBS. In 2002, Clarke et al noted that the substantial focus which PBS has on lifestyle outcomes rather than just on problem behaviour, was not well reflected in the literature, partly because of the difficulty in finding a single measure that could effectively evaluate quality of life. The authors report on a single case study with a young girl with autism and a learning disability, where they used a number of different measures in order to evaluate quality of life. Functional assessment was carried out and involved a range of stakeholders, following this a range of interventions were introduced. Measures used to evaluate included direct observation based on video recordings which considered engagement, affect, happiness and problem behaviour; and also the Quality of Life Survey (Knoster, 1999a) to assess relationships and participation; and the Positive Behavioural Satisfaction Survey (Knoster, 1999b) to assess stakeholders' satisfaction with the process. Impact was evaluated at a two-year follow-up and data show durable improvements in all of the variables which were evaluated; as well as reduced

challenging behaviour there was evidence of better relationships, positive affect and happy behaviours. In addition the authors reported that team members were happy with the process and the PBS interventions, thus being one of the few studies to comment on social validity and contextual fit.

Kincaid et al (2002) reported on behavioural and quality of life outcomes for 78 teams working with children and young people. Rather than focusing on actual behavioural reductions, they developed a behaviour outcomes survey which assessed the team member's perception of how and why the individual's behaviour had changed and measured whether staff felt comfortable with implementing the interventions, as well as evaluating the acquisition of new behaviours and skills. This showed that 86% of respondents were comfortable to implement the interventions and 82% believed behaviour was occurring less frequently; 71% also indicated that alternative skills were being used more frequently. In addition a 22-item scale to measure quality of life was used via interview with stakeholders and included areas such as social inclusion and interpersonal relationships; this demonstrated that respondents perceived modest improvements in quality of life.

Another study that addresses quality of life is West & Patton (2010) who reported on a supported employment intervention for four adults with severe disabilities and challenging behaviours. Functional behavioural assessment took place and the findings from this were used to successfully implement supported employment procedures; job coaches were used to provide prompts and to support the process. The authors reported that all participants met job performance criteria, and that there was a complete and immediate absence of challenging behaviours within their job settings. Quality of life improvements have also been noted in a range of other studies (e.g. Allen et al, 2011; McClean et al, 2007; McClean & Grey, 2012).

2.3.2.2 Use with Severe Behaviours

In 2007, McClean et al reported on an evaluation of PBS for five people with what they described as very severe challenging behaviours. Functional behavioural assessment was carried out and a multi-element behaviour support plan was implemented for each individual. This included environmental adaptations, skills teaching, reinforcement strategies and reactive strategies; a Periodic Service Review (PSR) (La Vigna et al, 1994) was introduced to monitor implementation. Behaviours reduced to near zero and were sustained over a two year period. Quality of life was also measured and was seen to improve significantly for three of the five individuals.

McClean & Grey (2012) reported on the use of PBS for four individuals with autism and very severe challenging behaviour, as indicated by the Checklist of Challenging Behaviours (Harris et al, 1994). Following functional behavioural assessment, each individual had interventions in a five step sequence; low arousal, rapport building, visual scheduling, skills teaching

and reinforcement strategies. Data were collected over a three year period and the PSR was used to monitor implementation including procedural reliability. Target behaviours reduced in frequency to 14% of the baseline and improvements in quality of life were also reported, via the Quality of Life Scale (Kincaid et al, 2002).

2.3.2.3 Use in Institutional Health Settings

Allen et al (2012) reported on the implementation of PBS within a specialist health service for acute admissions and long-stay inpatient units. PBS had previously been introduced as the main model of support and monthly use of physical interventions was tracked. Results showed reductions in use of physical intervention over the period of the study; a 74% reduction in use of physical interventions was noted from 2004-5 to 2010-11. Although no quality of life data were reported, and it was not possible to establish that the reduction in physical interventions was directly attributable to the implementation of PBS, these results are nonetheless a good indication of the effectiveness of PBS interventions when used systemically.

Gray et al (2013) reported on the use of PBS within an acute assessment and treatment health service, where service users were admitted at times of crises. An adapted and shortened form of PBS was implemented, including a shorter assessment process via the Brief Behavioural Assessment Tool (Smith & Nethell, 2010) and a brief PBS plan with limited focus on long term strategies. Outcomes were reported for 75 service users who had an average stay of 140 days within the unit; both Aberrant Behaviour Checklist (Aman & Singh, 1986) and Health of the Nation Outcome Scale-Learning Disability (Roy et al, 2002) average scores reduced significantly, demonstrating that PBS can be implemented within a crisis in-patient health service. The authors noted however, that no quality of life measures were used and therefore this important aspect of PBS was not able to be evaluated.

2.3.2.4 Use in Family Settings

Durand et al (2013) described a multi-site randomised clinical trial assessing the impact of PBS training sessions for parents. The two groups were PBS or PBS plus optimism training for parents; both groups of parents received eight weekly sessions of 90 minutes focusing on analysing their child's behaviour and then developing interventions based on this and which would be a good fit for the family circumstances. In addition to this, the experimental group also received optimism training added in as part of the sessions, for example, they were helped to identify patterns in their own thoughts and feelings and taught strategies to address these. Results showed that both groups improved in measures of the child's challenging behaviour, measured both directly through time sampling observation and also through standardised measures reported by parents. However, the PBS plus optimism training resulted in a significantly improved score

compared with PBS alone; this group also had a greater number of children who achieved reliable change (Jacobson & Truax, 1991) in both behavioural observations (56% as compared to 29% in PBS alone group) and in the use of the standardised measure (72% as compared to 35% in the PBS alone group). The authors commented that there were a number of important aspects to this study, in particular the fact that adding optimism training to PBS training appeared to make this more successful in terms of reducing behavioural challenges.

2.3.2.5 Specialist Behavioural Teams

Hassiotis et al (2009) described a randomised control trial in the UK for adults with learning disabilities. The 63 participants were split into either standard treatment from a community learning disability team, or into standard treatment plus PBS from a specialist behaviour therapy team. The approach included functional analysis, multi-element interventions and the use of an implementation monitoring tool, the PSR. After six months challenging behaviour as measured by the total score on the Aberrant Behaviour Checklist (ABC) reduced by 43% in the PBS group compared to treatment as usual. A follow-up to this study in 2011 (Hassiotis et al, 2011) demonstrated that the PBS group continued to have significantly lower ABC scores two years on from initial randomisation. In 2012, Hassiotis et al further updated this study with quality of life information; they reported that at 24 months participants in both groups were engaged in more activity but that there was no significant difference between the groups.

Allen et al (2011) reported findings from the Positive Behavioural Support: Clinical Practice and Outcomes Project (P-CPO), which was formed from three agencies providing clinical behavioural support in England and Ireland. Data were collected as part of the ordinary clinical work of the teams at point of referral to the services, at discharge, and at follow-up around 12 months after discharge. The authors noted that their intention was not to carry out a time-limited trial under specific conditions, but to provide ongoing data about the use and effectiveness of PBS in real-life conditions. The study reported on T2 data for 26 people, and results showed significant decrease on the ABC, a significant increase on the Guernsey Community Participation and Leisure Assessment (GCPLA) (Baker, 2000), and reductions in use of restrictive interventions and injuries caused by challenging behaviour.

2.3.2.6 Long-term Use

In 2010 Dunlap et al conducted a descriptive analysis of PBS for 21 individuals over two years with a view to measuring the long-term impact of PBS when implemented in community settings by natural implementation agents. Teams of relevant stakeholders were established for each individual and a PCP process was undertaken, followed by a functional behavioural assessment. This led on to a multi-element PBS plan with a range of interventions and an

implementation integrity measure; implementation was supported by on-site coaching by the researchers who acted as consultants and facilitators. A range of different tools were used to measure changes in behaviour and quality of life, which was defined as including five domains: material well-being, social well-being, personal well-being, leisure and recreation, and health and safety. Due to the variety of measures used, data were summarised and rated for changes over time on a nine-point scale ranging from 'situation (or behaviour) has deteriorated very substantially relative to baseline' to 'situation has improved substantially relative to baseline'. The authors reported that despite intervention integrity being variable, there was a general reduction in behaviour which was maintained over the two years, despite some relapses, and there was also evidence of improved quality of life, which tended to improve slightly further in the second year. They noted that this study evidenced the ability of natural agents to implement PBS in natural settings over time with positive outcomes for both behaviour and quality of life. In addition they noted that although behaviour changes were modest in comparison to other literature, this may be due in part to the fact that this was a longer study and implemented in natural settings; they concluded that long-term reductions in challenging behaviour in natural settings 'are not likely to be dramatic' (p. 273). Although this study presented encouraging changes in quality of life, its use of a rating scale for this rather than a reliable and valid measure means that there is some limitation to the application of their findings.

In 2014 Potter reported on a 21 year follow-up of a man with severe self-injurious behaviour; this is one of the longest studies involving PBS within the literature. Following a functional behavioural assessment, multi-element PBS was implemented with a range of strategies such as environmental manipulations, teaching new skills, stimulus change and non-aversive reactive strategies. Results show a significant reduction of self-injury, reducing to zero and this was maintained at three follow-ups over the 21 year period; community activities also increased and were maintained over this time, thus addressing the quality of life element of PBS.

2.3.3 Effectiveness of PBS

From this consideration of the evidence contained both in individual studies and in literature reviews, there are a number of observations that can be made in relation to effectiveness of PBS, and of potential enablers for the successful implementation of PBS, as well as some comments in relation to characteristics of PBS for which there is little evidence of impact within the literature and which may therefore have implications for future research.

2.3.3.1 Is PBS Effective

Despite quality of life changes being an important and arguably a defining feature of PBS, there has been limited success in the literature in evidencing this, particularly over a long period of time. Of the seven PBS review studies only three found that quality of life was reported on, and

even then it was a very small percentage of studies that reported, with the exception of O'Dell et al (2011). In addition, in terms of maintenance, the maximum that any review found was 15% of PBS studies reporting data post six months (Clarke & Dunlap, 2008). In the individual studies focused on here, there is more evidence of quality of life changes in some; however many studies note that it was either not possible to assess quality of life, or that this was not the focus of their study. It appears that it is widely acknowledged that PBS needs to be more successful in evidencing effectiveness in terms of quality of life, but that research has not yet successfully demonstrated this systematically.

In terms of impact on behaviour the results are clearer. In Carr et al (1999), success is taken as reduction of 90% from the baseline and over half the outcomes achieve this. If 80% reduction was taken as success, then the results would be even more impressive with 68% of outcomes achieving it. Snell et al (2005) reports that 97% of outcomes reported reductions in challenging behaviour, but no specific reductions are given. Goh & Bambara (2012) found a PND of 80% which is a moderate effect. The individual studies also reported reductions in behaviour, for example, McClean et al (2007) reported challenges reduced to near zero for five people with severe challenging behaviours, McClean & Grey (2012) reported reduction to 14% of the baseline for four individuals, and Hassiotis et al (2009) reported the ABC reduced by 43% for the PBS group compared to the control group. It is clear therefore that PBS interventions can demonstrate significant reductions in challenging behaviour; however since it is likely that successful interventions studies are more likely to be published than interventions having no impact, it is difficult to establish how effective PBS truly is when used in daily practice in the community. It is clear that it can be very effective in reducing challenging behaviour, and that there is a good range of evidence demonstrating its effectiveness in this respect.

It is also worth noting that in the studies considered here, many of the research designs are weak, using only a quasi-experimental design (an A-B format), rather than an experimental design with A-B-A-B reversal. There are also a number of single case or small *n* studies, for example Clarke et al (2002) and Potter (2014) are both single case studies, and West & Patton (2010) and McClean & Grey (2012) both report data for four service users; however it is noted that this small size of group is common in learning disability research (British Psychological Society and Royal College of Psychiatrists, 2015) and that there are guidelines available clarifying the defining features of single-subject research methodology (Horner et al, 2005), although these are not followed in many of the studies under consideration here. There are very few studies reporting on the impact of PBS interventions using a control group; both Hassiotis et al (2009) and Durand et al (2013) carried out single-blind randomised control trials, but these are the exception.

Earlier in this chapter the development and definition of PBS was traced and consideration was given to how different PBS is from ABA, or whether it is just a more modern, values-based

approach to ABA. The question can perhaps be returned to here, as it is clear in terms of the meta-analyses and research reviews that it is difficult to find reviews which focus only on PBS; many of the reviews just consider behavioural interventions more broadly. In fact the study regarded as providing the most in-depth analysis of PBS interventions (Carr et al, 1999), includes many studies published in the JABA and studies using interventions which would be described as ABA. With the individual studies it is easier to find PBS-specific studies; however many of them when describing the interventions carried out are also describing a range of ABA interventions. It appears therefore that when the implementation of PBS is described, it is very difficult to differentiate it from ABA.

2.3.3.2 Enablers to PBS

In their consideration of factors supporting the successful implementation of PBS Carr et al (1999) noted that carrying out a functional assessment could increase success rate by almost half. However as this is an essential element of PBS input, it is perhaps not legitimate to regard it as an enabler to effective PBS, since arguably without functional assessment the intervention cannot be truly described as PBS.

Ecological validity was also noted as an important factor by Carr et al (1999); they reported greater success with typical agents (families, direct care staff, teachers) than with non-typical (researchers, behaviour specialist and psychologists), a success rate of 66% for the former and 44.3% for the latter. Collaborative partnerships with direct care staff were also noted as an enabler by Goh & Bambara (2012).

Systems or organisational change were noted as helping in achieving more successful implementation of PBS by Carr et al (1999), although they noted that the information is limited. The use of PSR is one method of addressing the need for systems change, and its use has been reported in a small number of studies.

Hieneman and Dunlap (2000) conducted a descriptive study reporting on interviews carried out with family members, service providers and trainers in PBS, to discover their views on some of the factors that can increase the likelihood of PBS being implemented successfully. They identified 12 categories including: degree of personal investment by support staff, system responsiveness, alignment of the plan with the values of the support providers, and collaboration between providers. The authors noted that the results re-emphasised the importance of ensuring that PBS is implemented in a person-centred way, with reference to the resources and preferences of the key stakeholders. In phase two of the study the same authors (2001) identified the priority factors; personal investment or 'buy-in' of support staff was identified as the most important. They reflected that at this stage, this factor had not been the subject of empirical research, and noted that there is little information about the variables which impact on the effectiveness of PBS. Despite these limitations, these findings are associated with some of the key

features of PBS, that is, the importance of contextual fit and social validity, and the need for systems change as part of the process.

2.3.3.3 Future Research Implications

There are a number of factors for which there is little supporting evidence within the literature, despite being regarded as key elements of PBS. These include: use of typical settings (Carr et al, 1999; Goh & Bambara, 2012) and involving the team in assessment (Goh & Bambara, 2012). In addition, though not key factors, there is also limited evidence for the effect of different types of intervention, stimulus-based interventions versus reinforcement-based interventions (Carr et al, 1999; Goh & Bambara, 2012) and experimental functional analysis compared to descriptive (Goh & Bambara, 2012).

Factors of which there has been little consideration in the literature, therefore perhaps indicating the need for further research include: measures of generalisation (Carr et al, 1999; Goh & Bambara, 2012; Snell et al, 2005); measures of maintenance (Carr et al, 1999; Goh & Bambara, 2012); evidence re impact of lifestyle interventions (Carr et al, 1999; Dunlap & Carr, 2007; Goh & Bambara, 2012; Snell et al, 2005; La Vigna & Willis, 2012); effect of good social validity (Carr et al, 1999; Dunlap & Carr, 2007); impact of PBS interventions on the family (Dunlap & Carr, 2007); implementing and sustaining PBS on a systems-wide basis (Allen et al, 2013b; Dunlap & Carr, 2007).

One of the aspects noted in this review of the published literature purporting to describe outcomes of PBS, is that many of the studies included could be challenged in relation to whether what they are implementing is actually PBS. For example, in the Carr et al (1999) research synthesis, which is widely regarded as the seminal work on PBS evidence in research, it is clear that many studies were included which did not carry out a functional assessment, as were studies that did not use multi-element interventions, and also studies that used punishment were included. It could be argued that any one of these factors should disqualify the study from being defined as PBS and therefore these studies should not be included in the analysis or in the evidence base for PBS. Further research may be appropriate in order to assess the effectiveness of PBS studies that are truly implementing all the key elements of PBS and not just some aspects of it.

In contrast, La Vigna & Willis (2012) excluded studies which did not use functional assessment and multi-element interventions and which did use punishment. However, they did not report any quality of life outcomes in any of the 12 studies that they identified, and therefore another vital element of PBS, was missing. A similar picture emerges from the meta-analysis in Goh & Bambara (2012) which identified 83 studies, only one of which reported on lifestyle outcomes. It is of note that there are so few studies including lifestyle outcomes, when this is generally regarded as one of the key defining elements of PBS. This presents questions as to

why this has not been better demonstrated: for example is PBS, despite the considerable discussion on the subject, in fact not able to improve quality of life, and that is why there is such a lack of evidence in the literature? It may be that evidencing change in quality of life is more of a challenge than reducing behavioural problems, perhaps because quality of life is more difficult to define and to measure. Given the current lack of evidence it is not clear if quality of life changes are more difficult to achieve, that is, that quality of life is not actually being improved by PBS, despite all the rhetoric associated with it, or if it is just harder to evidence and report. Further research is indicated in order to address some of these areas specifically around quality of life.

2.4 Chapter Summary

This chapter has attempted to trace the development and definition of PBS from its first use in the literature as a specific term (Horner et al, 1990) to the most recent definition article (Gore et al, 2013). Through this process a number of key features of PBS have been identified and evidence for the selection of these features has been presented based on a range of studies describing and defining PBS over the past 25 years.

The chapter then went on to consider the evidence for the effectiveness of PBS within published literature, using literature reviews and research synthesis, as well as individual studies. The main enablers of effective PBS were identified as were possible areas for further research.

3 Review of the Literature in Relation to Staff Training in Positive Behaviour Support ³

3.1 Chapter Outline

This chapter provides a systematic review of the literature in relation to Positive Behaviour Support (PBS) training. The development of PBS training is discussed with reference to a number of leading studies which were the first to attempt to describe PBS training in detail, and from this a definition of the format of PBS training is presented. Following this a systematic review of PBS training is presented. The findings are summarised and discussed with reference to the type of study, measures used, and results. Lastly some implications are discussed in relation to further research in PBS training.

3.2 Development of Positive Behaviour Support Training

Chapter 2 described the gradual evolution of PBS over time which has allowed a clear definition to emerge with reference to some core essential elements; a similar approach can also be taken in terms of defining what constitutes PBS training. Through examining a range of early studies which describe PBS training, it is possible to come to a definition of PBS training. It is suggested that in addition to having a content which contains the 11 elements of PBS as outlined in chapter 2, PBS training must also involve a particular approach to training, with specific objectives and a particular format. There are a number of studies that have described PBS training in some detail, either from a theoretical standpoint or by describing the training actually delivered; generally these have described PBS training in relation to the training objectives and the format of the training. Considering how each of the studies addresses these will help towards a greater understanding of what constitutes PBS training.

The training model presented by Anderson et al (1993) and further described in Anderson et al (1996) proposed that PBS training, because it required the transmission of a new value base, as well as learning new skills and processes in practice, had to reach beyond traditional brief training and must be delivered in a specific way. They emphasised a case study approach, which would allow implementation of the learning through practical application in the workplace; this longitudinal training had a focus on long-term change in the service user's life and behaviour. Rather than just reducing problem behaviours, it aimed to teach new, more adaptive behaviours in addition to introducing a range of quality of life changes, designed to support them to live a fuller and more ordinary life in the community.

³ A version of this chapter was accepted as a chapter in the book *Clinical Handbook of Evidence-based Practices for Individuals with Intellectual and Developmental Disabilities*, N.S. Singh (Ed), Springer, and was also the basis for an article in the *Journal of Physical and Developmental Disabilities* (2013) 25: 17-33.

They noted the foundation of PBS is that ‘behaviour changes occur in the context of a rich lifestyle, not as a prerequisite to a rich lifestyle’ (p.364), thus specifying at the outset a key feature of PBS, that lifestyle enhancement or quality of life approaches are not just an outcome but are a technique, they are not an optional extra, but an essential element of PBS. They noted that training must be delivered in a way that promotes generalisation and maintenance, so that PBS skills learned can be used in a variety of contexts over a period of time. Specific objectives for the training were to achieve positive impact on: lives of service users (improving lifestyles as well as reducing behaviour); skills and knowledge of staff (focusing on proactive strategies, shifting from single to multi-component approaches, seeing behaviour change in the context of a good life, recognising the need for long-term support, not quick fixes); agencies and systems (PBS to be adopted into organisational policies, training to be prioritised proactively rather than crisis-led, resulting in flexible support options for individuals involving inter-agency and family collaboration).

Their description of how the training should be delivered is very detailed and was strategic in its aim to develop local expertise and to change systems. This included:

- Ensuring sustainability and developing local expertise via a training for trainers model, with good support to new trainers
- Targeting multiple audiences – training should be useful to a wide range of participants and should seek to create systems of comprehensive support across all of the individual’s environments
- Interspersing teaching with supported application via longitudinal training which is combined with periods of practice in the service setting, along with coaching and mentoring
- Providing a comprehensive PBS curriculum (including values, functional analysis, PBS planning, skill building, reinforcement, emergency management, evaluation and systems issues)
- Facilitating the development of PBS communities via ongoing implementation of PBS at a local level, and multiple levels of training, dependent on roles.

Dunlap et al (2000) refined and extended the definition of PBS training and their description of training objectives and format have considerable overlap with that of Anderson and colleagues. They also emphasised the need to work collaboratively in order to build multi-disciplinary partnerships between those people that supported an individual in a range of settings, with an aim to ensure enduring capacity to provide effective PBS. Some of the areas they proposed that training must involve were:

- Creating lifestyle changes for individuals with challenging behaviours

- Addressing systems change and enhancing systems to promote PBS
- Using a case study format – participants being brought together to apply their learning to an actual person
- Undertaking a dynamic training process – practical application interspersed with formal learning
- Providing comprehensive training addressing a broad range of topics (collective goals; building teams; functional assessment; designing PBS plans; implementing strategies; evaluation; infusing PBS into systems).

La Vigna and colleagues produced a number of papers describing their PBS multi-element model (La Vigna & Willis, 1995; La Vigna et al, 1989) but it is the paper in 2002 (La Vigna et al, 2002) where they first published a description of their intensive PBS training programme. The format of this training was intensive field-based supervised practice which involved hands-on implementation of PBS, with practice exercises, writing assignments and feedback sessions. There was an emphasis on carrying out functional analysis and developing and implementing multi-element plans with both proactive and reactive strategies. The training objectives of the programme were:

- To train participants to provide competency-based assessment including the development of multi-element support plans, designed to achieve valued outcomes in cost effective ways
- To provide participants with the skills, materials and procedures to continue delivering PBS in their workplaces
- To teach participants to use PBS interventions with consistency and accuracy, and to utilise ongoing quality improvement systems
- For participants to be able to support individuals with learning disabilities and severe challenging behaviours to integrate into ordinary work and living; and to be able to achieve durability and generalisation of effects for individuals.

The paper by Carr et al (2002) is regarded as seminal in the definition and development of PBS; in this they articulated a vision for how PBS should continue to develop and it is here that they outline their views of PBS training. This includes the need for more on-site education and less use of lecture format and formal training; real-life problem solving, in context, with sufficient time duration; and also the creative use of technology. Their objectives from PBS training were:

- Training a range of stakeholders, not just professionals
- Movement away from experts – more collaboration, with stakeholders and also between agencies
- Not simply a transfer of information from expert to provider, but involving capacity building and leading to system change
- Not based on a list of intervention techniques, but learning how to use interventions within systems and how to integrate these with broader infrastructures.

In 2006, Kincaid et al reviewed the ‘Positive Behavior Support Training Curriculum’, a training model implemented by the Rehabilitation Research and Training Centre on Positive Behavior Support in the US. In order to carry out this review, they first described what they regarded as the essential elements of PBS training and their summary therefore provides another definition of PBS training. It included: collaboration and team building between families and professionals; interventions that include lifestyle enhancements; addressing broader systems issues. They also noted the importance of a case study format, with a dynamic training process to allow practical implementation.

3.3 Definition of Positive Behaviour Support Training

In defining PBS training it is important not just to consider the content of the training, but also to identify any specific description for how PBS training is delivered. From reviewing these five studies above describing PBS training it became clear that there was a specific format to PBS training as described by different authors. Format refers to how the training is delivered, for example, how it is structured, timescales for sessions, and who participates in the training. This is separate from the actual content or curriculum of the training, which would be based on the 11 key features of PBS described in the previous chapter. Each of the five studies described above identified a number of features in relation to the format of PBS training; these are that training will be person-focused, have a longitudinal structure, involve stakeholders, and address systems change. These are described in more detail below.

Person-focused – PBS is a practical approach which deliberately focuses on application rather than theory, therefore through PBS training participants are supported to apply PBS in practice and are encouraged to carry out the training with a specific individual in mind. This leads on to teaching about functional assessment which is person focused and supports participants to an understanding of why behaviour occurs in particular environments and contexts for a specific individual. This, in turn leads to a range of person-specific multi-element plans, both proactive and reactive, being developed via the training process. PBS training is not intended to be a scientific-focused demonstration of techniques, rather it is intended to support trainees to understand and participate in a collaborative process. It is not about learning skills per se, but

rather about learning how to implement PBS in real settings, with individuals known to the trainees, with a focus of improving those individuals' lives.

Longitudinal – the training has a long-term focus and is comprised of more than one session; it is modular, with breaks in between sessions to apply the learning in practice. This element fits closely with being person-focused, as it is through a longitudinal process with opportunity for real-life practice that the training can best be truly person-focused and multi-element, and it is through this process that trainees can most effectively develop PBS skills in practice.

Involving stakeholders – within the definition of PBS, involvement of stakeholders is important, and so it also is within the training model. PBS training should not be focused solely on professionals or specialists; training is the primary means to make PBS more widespread in natural settings, therefore family carers and direct care staff are the essential groups to receive the training. A further aspect of stakeholder involvement is that of 'contextual fit', that is the congruence between the PBS interventions and the context in which they will be implemented (Albin et al, 1990); involving stakeholders in developing PBS plans during training helps ensure that the strategies implemented will fit with their values and make sense to their life experience. If those who implement the PBS interventions find the interventions a good fit with their own values, a good match for their goals, suited to their skills and resources, and congruent with the needs and demands of the environments in which they live and work, it is anticipated that they will be more likely to implement them. Involving stakeholders in the formation of these plans through a person-focused training model is therefore likely to increase contextual fit and therefore may also increase implementation. There is some evidence of this: Carr et al's (1999) meta-analysis of PBS interventions found that natural settings and carers demonstrated better outcomes than external intervention agents in controlled settings, a success rate in 61% of cases as opposed to 44%. Carr et al (2002) in fact go even further and state 'the primary goal of PBS is to help an individual change his or her lifestyle in a direction that gives all relevant stakeholders.....the opportunity....to enjoy an improved quality of life' (p.5), thus describing the needs of the individual, family, friends, staff, employers as central to the outcomes of PBS.

Systems change – PBS is a whole-systems approach and is most effective when not seen in isolation from the rest of the organisation; therefore when introducing PBS to a service or organisation, it must reach further than just behavioural changes. PBS training must address this issue, particularly if participants are to leave the training and implement what they have learnt. This issue is so central to PBS that Carr (2007) suggested that 'the central independent variable in PBS is systems change' (p4), and a range of other studies have shown associations between implementing a whole-system approach to PBS and improved outcomes for service users both in terms of challenging behaviour and quality of life (Allen et al, 2012; Allen et al, 2011; Perry et al, 2011). Allen et al (2013b) propose that for PBS to be successful, behavioural intervention

skills at the service user level need to be supplemented by organisational changes in order to ensure whole system commitment and widespread implementation. La Vigna et al (1994) suggested a model for this, their Periodic Service Review (PSR), which is a quality assurance system that can be used to support the implementation of PBS.

In order to check each of the five studies against these features, a reliability check was carried out by a colleague of the writer's⁴; there was 100% agreement between the writer and the colleague in relation to the key features of the definition of PBS training format. This information is summarised in Table 3.1

Table 3.1: Studies Defining PBS Training Format

	Anderson et al 1993	Dunlap et al 2000	La Vigna et al 2002	Carr et al 2002	Kincaid et al 2006
Person-focused	✓	✓	✓	✓	✓
Longitudinal	✓	✓	✓	✓	✓
Stakeholder involvement	✓	✓	✓	✓	✓
Systems change	✓	✓	✓	✓	✓

3.4 Method

The objective of this review is to summarise the results from published studies in the last 25 years regarding outcomes of PBS staff training in relation to adults with learning disabilities. Studies which include either outcomes for staff (e.g. changes in knowledge or skills) and/or outcomes for service users (e.g. changes in rate or severity of challenging behaviour, impact on quality of life, reductions in use of restraint) will be considered.

3.4.1 Identification of Literature

Searches of Google Scholar and of electronic databases (Web of Science, Pub Med, and Psych-INFO) were conducted in February 2015 to identify relevant studies, published in English from 1990 onwards. In each database searches were made firstly for “positive behav* support”, and then with additional search terms such as “training” and “learning disab* OR developmental disab* OR intellectual disab*”. Searches were also made under “challenging behav*” with the

⁴ Behaviour Support Advisor with the organisation's Positive Behaviour Support Team

same search terms added in. Further details about the searches conducted are available in appendix one.

In addition to these initial searches, reference lists of studies found by the initial search were then accessed and their relevance assessed. Web of Science and PsychINFO were used for citation searches in order to identify further studies citing those already identified. Finally, the contents list of the International Journal of Positive Behavioural Support was checked. All studies were assessed for relevance by reading abstracts. If abstracts indicated that the training may be relevant to the review, the full text of the article was assessed. In reviewing studies, particular attention was paid to the method section and to the description of training given, and this description was matched for *content* against the 11 defining features for PBS outlined in chapter 2, and also against the four key features in relation to *format* of training as outlined above. For studies where the training was not described in detail within the study, but where reference was made to other published work which described the approach used in the training, then these original sources were also sourced and read. The description of the approach in the source work could then be taken as the basis for the training described in the study, and the study was scored accordingly. Reliability checks were done by a colleague of the writer and there was 100% agreement in articles included and excluded.

3.4.2 Inclusion Criteria

Studies were included that described outcomes from training with PBS content, as defined in chapter 2 where a review of the major literature on PBS concluded that there are 11 key features. A judgement was taken that if any eight of these were included in the description of the training, then the study would be regarded as having PBS content. The reason for eight being taken as the cut-off is because Horner et al (1990) (widely regarded as the original definition of PBS) included eight of these features; the rationale being that if this is generally regarded as PBS, then other studies which also include eight or more PBS features could also be regarded as PBS.

In addition to meeting the criteria for PBS content, studies also had to meet all four key features of format for PBS training: to be person-focused; to be longitudinal and to allow opportunity for trainees to apply the learning in practice; to involve stakeholders, that is, direct care staff not just specialists; and to have a focus on systems or organisational change, not just change for the service user or for individual staff.

Participants for the studies in this review were individuals with learning disabilities and challenging behaviour, and/or the staff that provide their support.

Due to the limited number of studies found, criteria for inclusion were kept wide and included training of different lengths, with any type of research design, including with or without control groups, and with or without reliability measures. No limitations were placed on type of

outcome measures, whether based on outcomes for staff or for service users. Studies which included outcomes for both staff and families were included, but only the staff data are reported. Studies focusing purely on family carers were not included; the current study was focused only on staff training in PBS, therefore this was felt to be a more relevant criterion.

3.4.3 Included and Excluded Studies

Following the search process outlined above, 38 studies were identified from their titles as being potentially part of the review. The process for assessment is outlined in Figure 1. Firstly, abstracts were read and this resulted in 10 being excluded for the following reasons:

- Studies which did not report on PBS training, but more generic challenging behaviour training or physical interventions (e.g. McDonnell et al, 2008b; van Oorsouw et al, 2010)
- Studies relating to training carried out with family carers only, rather than staff (Durand et al, 2013)
- Studies which described PBS training for working with service users without a learning disability (e.g. La Vigna et al, 2005)
- Studies relating to school-wide PBS, as these were in relation to service users without a learning disability (e.g. Sugai et al, 2000)
- Studies which were focused on PBS interventions, not training (e.g. West & Patton, 2010).

This left 28 studies which needed to be read in detail in order to decide on their inclusion or exclusion. For some studies, the training was not sufficiently described, and therefore reference was made to a previous ‘source article’ with more detail about the training; where this occurred the source article was accessed and training content or format was assessed via that. Each of the studies was assessed in detail against the inclusion criteria and the results are shown in Table 3.2. Detailed assessment was necessary in relation to content as some studies referred to PBS and yet did not meet the criteria (Allen et al, 1997; Allen & Tynan, 2000; Tierney et al, 2007) while one study described itself as Applied Behaviour Analysis, and yet came very close to being included as it met seven features of PBS content (Grey et al, 2005). Allen et al (1997) discussed the use of PBS within their unit and made reference to the La Vigna multi-element model (La Vigna et al, 1989); however the study described behaviour management training with a focus on reactive strategies and therefore it was excluded. Likewise the later study from Allen & Tynan (2000) was excluded for similar reasons. Tierney et al (2007) referred to PBS within their training (they discussed the importance of PBS plans), yet with insufficient detail to be able to demonstrate that they met the necessary inclusion criteria; this study was therefore

excluded. The study by La Vigna et al (2005) was in relation to people without a learning disability, so despite meeting all other criteria, it was excluded. Studies were also excluded based on their format, despite having a PBS content. For example, Rose et al (2014) and Kraemer et al (2008) were both excluded as they only provided one-day training and therefore could not meet the format criteria, despite meeting the content criteria. If studies did not meet the PBS content criteria, then format was not assessed.

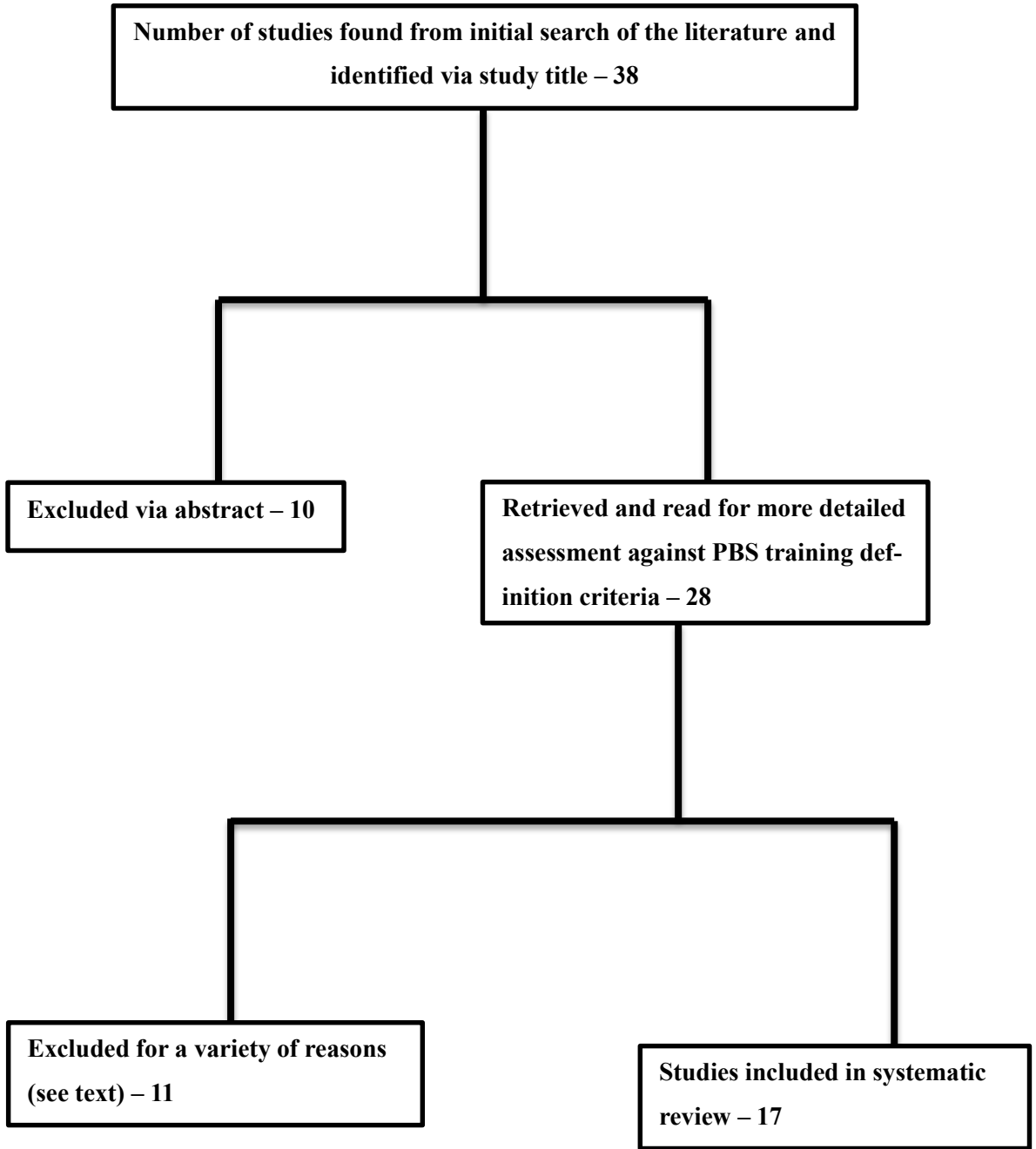


Figure 1. Process for Assessing Abstracts

Table 3.2: Studies Included and Excluded in Systematic Review (from the 28 studies read in full)

Study	Source Articles	Content Standards (11) (see Table 2.2)	Content Met (at least 8)	Format Standards				Format Met (all)	Included or not
				Person-focused	Longitudinal	Stakeholder involvement	Systems change		
Berryman et al, 1994	-	11	✓	✓	X	✓	X	X	No
Allen et al, 1997	-	5 (2, 3, 5, 10, 11)	X	-	-	-	-	NA	No
Baker, 1998	-	10 (1, 2, 4, 5, 6, 7, 8, 9, 10, 11)	✓	✓	✓	✓	✓	✓	Yes
Allen & Tynan, 2000	-	5 (2, 3, 5, 10, 11)	X	-	-	-	-	NA	No
Grey et al, 2002	La Vigna & Donnellan (1989)	11	✓	✓	✓	✓	✓	✓	Yes
La Vigna et al, 2002	-	11	✓	✓	✓	✓	✓	✓	Yes

Study	Source Articles	Content Standards (11) (see Table 2.2)	Content Met (at least 8)	Format Standards				Format Met (all)	Included or not
				Person-focused	Longitudinal	Stakeholder involvement	Systems change		
Reid et al, 2003	-	9 (1, 2, 3, 4, 5, 6, 8, 10, 11)	✓	✓	✓	✓	✓	✓	Yes
Freeman et al, 2005	Anderson et al (1993 & 1996) & Dunlap et al (2000)	11	✓	✓	✓	✓	✓	✓	Yes
La Vigna et al, 2005	-	11	✓	✓	✓	✓	✓	✓	No
McClellan et al, 2005	La Vigna & Willis (1995)	11	✓	✓	✓	✓	✓	✓	Yes
Dench, 2005	Donnellan et al (1985)	11	✓	✓	✓	✓	✓	✓	Yes
Grey et al, 2005	-	7 (2, 5, 6, 7, 8, 10, 11)	X	-	-	-	-	NA	No

Study	Source Articles	Content Standards (11) (see Table 2.2)	Content Met (at least 8)	Format Standards				Format Met (all)	Included or not
				Person-focused	Longitudinal	Stakeholder involvement	Systems change		
Grey & McClean, 2007	McClellan et al (2005)	11	✓	✓	✓	✓	✓	✓	Yes
Tierney et al, 2007	-	5 (2, 5, 6, 10, 11)	X	-	-	-	-	NA	No
McGill et al, 2007	La Vigna et al (1989)	8 (1, 2, 3, 4, 5, 6, 10, 11)	✓	✓	✓	✓	✓	✓	Yes
Lowe et al, 2007b	-	11	✓	✓	✓	✓	✓	✓	Yes
Browning-Wright et al, 2007	Browning-Wright et al (2003)	11	✓	X	X	X	X	X	No
Dowey et al, 2007	-	4 (1, 2, 6, 11)	X	-	-	-	-	NA	No

Study	Source Articles	Content Standards (11) (see Table 2.2)	Content Met (at least 8)	Format Standards				Format Met (all)	Included or not
				Person-focused	Longitudinal	Stakeholder involvement	Systems change		
Kraemer et al, 2008	Browning-Wright et al (2003)	11	✓	X	X	X	X	X	No
Macurik et al, 2008	-	5 (3, 5, 6, 10, 11)	X	-	-	-	-	NA	No
Gore & Umizawa, 2011	Carr et al (1999)	11	✓	✓	✓	✓	✓	✓	Yes
Reynolds et al, 2011	Dunlap et al (2000)	10 (1, 2, 3, 4, 5, 6, 7, 8, 10, 11)	✓	✓	✓	✓	✓	✓	Yes
McClellan & Grey, 2012	McClellan et al (2005)	11	✓	✓	✓	✓	✓	✓	Yes
Crates & Spicer, 2012	-	11	✓	✓	✓	✓	✓	✓	Yes

Study	Source Articles	Content Standards (11) (see Table 2.2)	Content Met (at least 8)	Format Standards				Format Met (all)	Included or not
				Person-focused	Longitudinal	Stakeholder involvement	Systems change		
Wills et al, 2013	Baker & Shephard, 2005 & Allen et al, 2005	9 (1, 2, 3, 4, 5, 6, 7, 10, 11)	✓	✓	✓	✓	✓	✓	Yes
Wardale et al, 2014 (a)		10 (1, 2, 3, 4, 5, 6, 7, 8, 10, 11)	✓	✓	✓	✓	✓	✓	Yes
Wardale et al, 2014(b)	Wardale et al, 2014 (a)	10 (1, 2, 3, 4, 5, 6, 7, 8, 10, 11)	✓	✓	✓	✓	✓	✓	Yes
Rose et al, 2014		9 (1, 2, 3, 5, 6, 7, 8, 10, 11)	✓	X	X	✓	X	X	No

3.5 Results

Studies meeting the above criteria were summarised into a table with information about author, date of publication, country where research was conducted, participants involved, sample size, design and data collection; outcome measures; reliability and results. This information is summarised in Table 3.3, in chronological order starting with the oldest studies first.

Although there is extensive research on generic challenging behaviour training available, there is limited literature published regarding outcomes of PBS training. This review identified 17 studies; five from Ireland, four from the US, four from the UK, three from Australia and one from Canada. Table 3.4 shows the length, format, and content of the training.

The studies vary as to whether they focused on staff outcomes (nine), service user outcomes (four), or both (four). In order to consider the studies and their outcomes in more detail, the 17 studies have been split into those including results for staff and those including results for service users.

3.5.1 Staff Studies

Participant numbers for the 13 studies that included staff outcomes ranged from 11 – 386 and included the following types of participant: staff from a residential and vocational support unit; staff from community services, including residential, day-support and specialist; direct-care staff; students undertaking a university diploma; students undertaking a distance learning course; nurses; behaviour specialists; teachers; allied health professionals; service managers; staff from a forensic secure unit and unregistered staff from a specialist health resources.

In terms of design, only one study of the 13 studies considering staff outcomes used a control group; Reid et al (2003) used this for the first phase of their study, although this was not randomised. Five studies included reliability data (Baker, 1998; Reid et al, 2003; McGill et al, 2007; Lowe et al, 2007b; Crates & Spicer, 2012). Ten studies used a repeated measure design (Baker, 1998; Grey et al, 2002; Reid et al, 2003; Freeman et al, 2005; McGill et al, 2007; Lowe et al, 2007b; Gore & Umizawa, 2011; Wills et al, 2013; and Wardale et al, 2014a & 2014b), with one of these studies having three data collection points (McGill et al, 2007). Three studies used a post-training evaluation only for their staff measures (La Vigna et al, 2002; Dench, 2005; Crates & Spicer, 2012).

Outcome measures used in each study to evaluate different variables were identified and are described below; these are summarised in Table 3.5.

Table 3.3: Details of Studies Included in the Systematic Review

Author(s)	Year	Country	Participants	Sample Size	Design and Data Collection	Outcome Measures	Reliability	Results
Baker ***	1998	US	Managerial and direct care staff from residential and vocational support unit Service users on whom training was focused	At least 16 staff 5 service users	Repeated measures design: review of Functional Assessment (FA) and Behaviour Support Plan (BSP) pre and post training; incident reports completed pre and post training, during a 2-month period	FA and BSP were reviewed against 'critical elements' (4 for FA and 6 for BSP) Incident reports	Coding was checked for 50% of FA and BSP (93-96% agreement) Challenging Behaviour (CB) reliability was assessed via comparison with house logs (100% agreement)	The number of FA and BSP with at least 4 critical elements significantly increased In 2 months post training, CB for 2 service users reduced to 0; others reduced to between 11-28% of baseline
Grey et al **	2002	Ireland	Residential and day service staff	34	Repeated measures design; before, during and after course	Challenging Behaviour Attributions scale (CHABA) (Hastings, 1997)	None reported	Positive changes in attribution from positive reinforcement to negative reinforcement and self-stimulation

Author(s)	Year	Country	Participants	Sample Size	Design and Data Collection	Outcome Measures	Reliability	Results
La Vigna et al **	2002	US	Professionals attending Summer Institute between 1989 – 1999 from 8 countries	126	Post training postal questionnaire	Self-report questionnaire re acquisition and use of skills	None reported	Acquisition of skills high in 5/5 key areas; use of skills high in 4/5
Reid et al **	2003	US	Supervisors of direct care staff	Pilot phase 12	Repeated measures within and between groups design	Observation of role-play demonstration of 2 supervisory skills	Inter-observer agreement checks on 46% of observations (95% agreement)	Pre-training 0-33% met criteria; post training 100%
				Implementation 386	Post training evaluation of skill	Paper exercise evaluation (17 modules); role-play evaluation (9 modules); and on-the-job checks (6 modules)	None reported	85% completed all training and achieved 'mastery' level
Freeman et al **	2005	US	Professionals working in human services	11	Repeated measures design, PBS experts score BSP pre and post training	37-item PBS Checklist (based on Horner et al, 2000)	None reported	10 PCP and 9 BSP produced; scores on both checklists increased from

Author(s)	Year	Country	Participants	Sample Size	Design and Data Collection	Outcome Measures	Reliability	Results
			with bachelor's degree					48% pre training to 73% post training
McCleane et al *	2005	Ireland	Service users whose staff were on the training course	138	Repeated measures design; staff recordings of target behaviour at 3 time points, baseline, intervention and follow-up (between 4-8 week periods)	Behaviour recordings	Inter-rater reliability checks for 1 week during baseline only for 21 cases (92% agreement)	Significant improvement in 77%, at average follow-up of 22 months (3 months – 5.5 years)
Dench ***	2005	Ireland	Students undertaking the course Service users – focus person for the course	38 staff 25 service users	Post training assessment of FA and BSP Staff recording of target behaviour at baseline and post training; pre and post questionnaire re quality of life	Behaviour Assessment Report and Intervention Plan Evaluation Instrument (Willis & La Vigna, 1990) Behaviour recordings; Quality of Life	None reported	91.7% scored across all 7 categories on the Evaluation Instrument CB reduced to less than 30% of baseline within 3 months for 56%; questionnaire did not demonstrate changes in quality of life;

Author(s)	Year	Country	Participants	Sample Size	Design and Data Collection	Outcome Measures	Reliability	Results
						Questionnaire (Schalock & Keith, 1993)		
McGill et al **	2007	UK	University Diploma students	79	Repeated measures design: questionnaires completed at 3 time points, beginning middle and end of course	Self-Injury Questionnaire (SIBUQ) (Oliver et al, 1996) ; CHABA (Hastings, 1997) ; Emotional Reactions to Challenging Behaviour scale (ERCB) (Mitchell & Hastings, 1998); vignettes	5% of data repeated entry to check reliability (95% agreement) Independent rating of 15-23% of vignettes (91% agreement)	SIBUQ – very significant increase of correct attributions; CHABA – no change in attributions; ERCB – negative emotional responses reduced; knowledge – significantly increased
Lowe et al **	2007 b	UK	Registered and non-registered staff in specialist health services	275	Repeated measures design: self-report questionnaire and knowledge test pre and	ERCB; CHABA; Challenging Behaviour Staff Perceptions Questionnaire (CBSPQ)	Inter-assessor reliability for 25% (86% agreement)	ERCB – limited impact on fear/anxiety or depression/anger scale; CHABA - initial changes in attributions after training reverted

Author(s)	Year	Country	Participants	Sample Size	Design and Data Collection	Outcome Measures	Reliability	Results
					post training and at 1 year follow-up	(Hastings & Brown, 2002) ; Confidence in Coping with Patient Aggression (CCPA) (Thackery, 1987); knowledge questions		to baseline over time; CBSPQ and CCPA – lasting increases occurred in confidence for both groups; knowledge – significant increase for both groups
Grey & McClean *	2007	Ireland	Service users whose staff were on the training course	60	Non-randomised matched control group (service users whose staff were not trained), repeated measures design; CCB completed by one staff member pre and post training ; staff completion of behaviour observation records for target group	Behaviour observation recordings (target group only); Checklist of Challenging Behaviour (CCB) (Harris et al, 1994); prescription of psychotropic medication	43% of CCB were completed by 2 staff to check inter-rater reliability (correlation coefficient range 0.8-0.92)	Significant differences between the groups post-training in frequency, management difficult and severity of behaviour (on CCB) For 66% of the target group the frequency of CB reduced to below 30% of baseline after 3 months post

Author(s)	Year	Country	Participants	Sample Size	Design and Data Collection	Outcome Measures	Reliability	Results
					only, pre and post training			implementation (behaviour recordings) No significant reduction in units of medication prescribed for either group
Gore & Umizawa *** + see note below	2011	UK	Teaching staff	33 staff 37 children	Repeated measures design: all measures completed by participants pre and post training; CCB post data collection was 1 month after training	CHABA; ERCB; CCB	None reported	CHABA – no significant differences; ERCB – significant decrease on fear/anxiety scale, no other significant differences CCB – one month after training significant decrease in frequency but not in severity or management difficulty for staff (difference between staff and family carer results)

Author(s)	Year	Country	Participants	Sample Size	Design and Data Collection	Outcome Measures	Reliability	Results
Reynolds et al * + see note below	2011	Canada	Children whose parents and staff attended training	35	Repeated measures design, pre and post training	Aberrant Behaviour Checklist (ABC) (Aman & Singh, 1986)	None reported	Significant improvements reported by staff in the total ABC score
McClellan & Grey *	2012	Ireland	Service users whose staff were on the training course	61	Repeated measures design: 3 data collection points – baseline, post training and follow-up (average 26 month follow-up)	Challenging Behaviour Rating Scale (CBRS) (based on Checklist of Challenging Behaviour, Harris et al, 1994); behaviour recordings	CBRS completed by 2 staff at baseline and post training (Pearson's correlation coefficient range 0.8-0.9)	Significant reduction in frequency, management difficulty and severity in the CBRS; behaviour records showed an average decrease of 61% at 3 months
Crates & Spicer ***	2012	Australia	Professionals working in disability services and the service users that they completed the training for	32 staff 32 service users	Post training evaluation of assessment ; repeated measures design for behaviour; 2 data collection points – baseline and 3 months post-training;	Assessment and Intervention Plan Evaluation Instrument (AIEI) (La Vigna et al, 2005)	Inter-rater reliability checks for marking of 62% of reports – 85%; mean reliability for behaviour occurrence and severity was at least 84%	Mean score on AIEI was 79.5% Reduction in occurrence of behaviour at 3 months for 29/32 (mean change was 49.6% of baseline); 27/30

Author(s)	Year	Country	Participants	Sample Size	Design and Data Collection	Outcome Measures	Reliability	Results
								showed reduction in severity at 3 months (mean change was 30.8% of baseline)
Wills et al**	2013	UK	Care staff from residential and day services	38	Repeated measures design, pre and post training	PBS multiple choice questionnaire; Revised Causal Dimension Scale II (McAuley et al, 1992); Helping Behaviour Scale (Jones & Hastings, 2003); emotional responses to challenging behaviour questionnaire (based on Rose & Rose, 2005); Optimism Pessimism Scale	None reported	Significant increase in PBS knowledge; significant decrease in controllability in causal attributions; significant increase in proactive helping; significant increase in optimism

Author(s)	Year	Country	Participants	Sample Size	Design and Data Collection	Outcome Measures	Reliability	Results
						(Mores & Grant, 1976)		
Wardale et al **	2014 (a)	Australia	Health professionals; direct support workers; managers; team leaders	234	Repeated measures design, pre and post training, except for BSP-QE only used post training	Knowledge acquisition test (based on O'Neill et al, 1997); CHABA; Evidence Based Practice Attitude Scale (EBPAS-50) (Aarons et al, 2010) ; Behaviour Support Plan Quality Evaluation Guide (BSP-QE) (Brown-ing-Wright et al 2003) used to score each BSP	None reported	Significant increase in PBS knowledge; significant differences for all CHABA subscales except for Emotional; significant differences for 2 of the EBPAS-50 subscales (Openness and Fit increased); participants produced PBS plans with mean score of 13.73

Author(s)	Year	Country	Participants	Sample Size	Design and Data Collection	Outcome Measures	Reliability	Results
Wardale et al **	2014 (b)	Australia	Care staff from a secure facility for offenders with learning disabilities	6	Repeated measures design, pre and post training, except for BSP-QE only used post training	Knowledge acquisition test (based on O'Neill et al, 1997); BSP-QE	None reported	Participant knowledge increased (significance not reported); participants produced PBS plans with scores of 18-20 (<i>good</i>)

+A variety of measures were used in relation to family carers, but these are not reported here, as this review only considers outcomes for staff and service users

* Service user outcomes only; ** Staff outcomes only; ***Both staff and service user outcomes.

Table 3.4: Length, Format and Content of Training in Studies Included in Systematic Review

Study	Length and Format of Training	Content
Baker (1998)	3x 3-hour sessions one month apart; inter-session assignments	Completing FA; developing and implementing BSP for focus person
Grey et al (2002)	9 days over 6 months; inter-session assignments	Multi-element PBS; completing FA and implementing BSP for focus person
La Vigna et al (2002)	2-weeks; intensive practice-based assignments	Supervised use of multi-element model; completing FA and BSP for focus person
Reid et al (2003)	Experimental group: not specified Implementation group: 5 days over 5 weeks (day 4 was on-the-job training)	Experimental: 2 modules (teaching staff a PBS-related skill and carrying out a staff observation) Implementation: Skills related to PBS; 26 module curriculum
Freeman et al (2005)	10 hours/ week over 1 year; online and monthly classes; online and practice-based assignments; portfolio	Person-centred planning, FA and BSP for focus person
McClean et al (2005)	9 days over 6 months in 5 blocks; 4 inter-session assignments	Multi-element PBS; completing FA and implementing BSP for focus person and Periodic Service Review (PSR)
Dench (2005)	9 days over 9 months in 6 blocks; 5 inter-session assignments	Multi-element PBS; completing FA and implementing BSP for focus person and PSR
McGill et al (2007)	57 days over 2 years; part-time University Diploma; practice-based assignments to implement in workplace	ABA: Active Support (AS); FA; multi-element PBS
Lowe et al (2007b)	80 hours teaching over 10 consecutive days: portfolio; mentored by unit managers; 5 on-the-job observations	AS; PBS

Study	Length and Format of Training	Content
Grey & McClean (2007)	9 days over 6 months in 5 blocks; 4 inter-session assignments	Multi-element PBS; completing FA and implementing BSP for focus person and PSR
Gore & Umizawa (2011)	2x 4-hour sessions; session 1 repeated twice, delivered separately to staff and family carers; session 2 was joint	Functions of behaviour; communication; recording behaviour; developing proactive and reactive strategies
Reynolds et al (2011)	3 full days and 2 half-days training over 7 weeks; training took place in teams of staff and family carers, focused round one child	Behavioural theory; functional assessment; designing BSP; data collection; crisis intervention
McClellan & Grey (2012)	9 days over 6 months in 5 blocks; 4 inter-session assignments	Person-focused training in multi-element PBS; completing FA and implementing BSP and PSR
Crates & Spicer (2012)	4 consecutive days then an additional 9 days over 9 months involving 3 practice-based assignments	Multi-element model; functional assessment; positive programming; focussed support; reactive strategies
Wills et al (2013)	5 half-days per week for 5 consecutive weeks	Understanding challenging behaviour; challenging environments; environmental change; positive programming; secondary prevention
Wardale et al (2014a)	4 days over 6-8 weeks	PCP; FA (informant and observation methods and hypotheses); developing and implementing PBS plans; reactive strategies
Wardale et al (2014b)	3 days across 5 weeks	PCP; FA (informant and observation methods and hypotheses); developing and implementing PBS plans; reactive strategies

Table 3.5: Staff Outcome Measures from Studies in Systematic Review

Variable Measured	Outcome Measures	Studies Reporting
Attitudes	Evidence Based Practice Attitude Scale (Aarons et al, 2010)	Wardale et al (2014a)
Attributions	Challenging Behaviour Attributions questionnaire (CHABA) (Hastings, 1997)	Grey et al (2002); Lowe et al (2007b); McGill et al (2007); Gore & Umizawa (2011); Wardale et al (2014a)
	Revised Causal Dimension Scale II (McAuley et al, 1992)	Wills et al (2013)
Confidence	Confidence in Coping with Patient Aggression Instrument (CCPAI) (Thackery, 1987)	Lowe et al (2007b)
	Challenging Behaviour Staff Perceptions Questionnaire Self-efficacy (CBSPQ) (Hastings & Brown, 2002)	Lowe et al (2007b)
Emotions	Emotional Reactions to Challenging Behaviour scale (ERCB) (Mitchell & Hastings, 1998)	Lowe et al (2007b) & McGill et al (2007); Gore & Umizawa (2011)
	Emotional responses to challenging behaviour questionnaire (based on Rose & Rose, 2005)	Wills et al (2013)
Helping	Helping Behaviour Scale (Jones & Hastings, 2003)	Wills et al (2013)
Knowledge	Knowledge-based questionnaire	Lowe et al (2007b); Wills et al (2013); Wardale et al (2014a); Wardale et al (2014b)

Variable Measured	Outcome Measures	Studies Reporting
Knowledge	Vignettes on behavioural function	McGill et al (2007)
	Self-Injury Understanding (SIBUQ) (Oliver et al, 1996)	McGill et al (2007)
Optimism	Optimism Pessimism Scale (Mores & Grant, 1976)	Wills et al (2013)
Skills	Questionnaire re skills and confidence in using tools	La Vigna et al (2002)
	Completion of FA and implementation of BSP	Baker (1998); Freeman et al (2005); Dench (2005); Crates & Spicer (2012)
	Carolina Curriculum on Positive Behavior Support	Reid et al (2003)
	Behaviour Support Plan Quality Evaluation Guide (BSP-QE) (Browning-Wright et al, 2003)	Wardale et al (2014a); Wardale et al (2014b)

Attitudes

One study looked specifically at staff attitudes. Wardale et al (2014a) used the Evidence Based Practice Attitude Scale (Aarons et al, 2010) which evaluates attitudes towards evidence-based practice. They found significant differences for two of the 12 subscales, Openness and Fit (all significant differences reported were as expected, unless otherwise specified).

Attributions

Six studies considered changes in attributions, five of these using the CHABA (Hastings, 1997). Grey et al (2002) saw positive changes in attribution, however they also reported limitations with the CHABA in relation to content validity; Lowe et al (2007b) reported initial changes in attribution after training reverted to baseline over time, and McGill et al (2007) found varied change across measures. Gore & Umizawa (2011) found no significant differences, and Wardale et al (2014a) found significant differences for all the subscales except Emotional. Wills et al (2013) used the Revised Causal Dimension Scale II (McAuley, 1992) and adapted this to refer to a case study describing a service user with learning disabilities and challenging behaviour; they reported little change in mean scores on the internality and stability dimensions, but the mean controllability score decreased significantly

Confidence

Lowe et al (2007b) measured confidence via the CCPA (Thackery, 1997) and also used the CBSPQ (Hastings & Brown, 2002); both of these showed significant increases in confidence.

Emotions

Four studies looked at this area, three of them utilising the ERCB (Mitchell & Hastings, 1998); Lowe et al (2007b) reported little lasting effect on emotions; McGill et al (2007) reported that Depression/Anger scores were significantly reduced, and Gore & Umizawa (2011) found a significant decrease in the Fear/Anxiety scores. Wills et al (2013) also considered emotions using a questionnaire based on Rose & Rose (2005); after training, negative emotions reduced by one point on the scale and empathy increased by nearly one point; there was little change in mean scores of positive emotion.

Helping

One study (Wills et al, 2013) evaluated changes in helping behaviour from staff, using the Helping Behaviour Scale (Jones & Hastings, 2003), which looks at the help offered by staff to service users engaging in self-injurious behaviour; this was adapted to include help in relation to aggressive behaviour also. Following training, they found a significant increase in proactive helping, and a significant decrease in unhelpful behaviour.

Knowledge

This was measured in a number of different ways in different studies. Four studies used a PBS knowledge questionnaire: Lowe et al (2007b) used a questionnaire based on the course work which covered policy, PBS strategies, PCP and Active Support; this demonstrated significant increases in knowledge, both immediately after the training and over time. Wills et al (2013) used a multiple choice questionnaire and found significant increase in knowledge, and Wardale et al (2014a & 2014b) used a test based on O'Neil et al (1997); they noted an increase in scores although they did not report significance. McGill et al (2007) reported on the SIBUQ (Oliver al, 1996) and showed a highly significant increase in correct attributions, demonstrating increased knowledge regarding self-injury. This study also used vignettes on behavioural function, where responses were rated as correct, partially correct or incorrect; post-training the students demonstrated more correct analysis with regard to the scenario with an attention function, but not with the escape/avoidance function.

Optimism

One study considered changes in staff optimism, using the Optimism Pessimism Scale (Mores & Grant, 1976), which measured how optimistic staff were about the possibility of reducing a

service user's likelihood of engaging in challenging behaviour. Wills et al (2013) found a significant increase in optimism following training.

Skills

There were a number of different ways of assessing skills gained in PBS. The most common method in these studies was production and/or evaluation of FA or BSP. In 1998 Baker, reported on FA and BSP which were produced following three 3-hour training sessions; although the quality of these was not formally assessed, each were reviewed to ensure that they met a number of criteria. His definition was that FA should contain: description of problem behaviour; predictors of the behaviours; testable explanations for the function; direct observation to confirm this explanation and BSP should contain at least four of six elements: definition of the behaviour; link to FA; emphasizing lifestyle change; including interventions addressing setting events; including interventions addressing consequences; emergency or reactive strategies. All three groups receiving training produced FA and BSP that met these criteria.

Dench (2005) measured the FA and BSP produced by students against an adapted version of the Assessment Report and Intervention Plan Evaluation Instrument (Willis & La Vigna, 1990) which lists 49 separate competencies under seven headings. Students achieved an average of 91.7% across all units, with one area, principles of skill-teaching, being lower than the others. Crates & Spicer (2012) also evaluated FA and BSP via a version of the same instrument, the Assessment and Intervention Plan Evaluation Instrument (AIEI) (La Vigna et al, 2005) which has 140 criteria in 12 areas of content. They reported a mean score of 79.85% over the two training sessions evaluated using the AIEI. La Vigna et al (2002) also required trainees to complete a FA and BSP, but they did not report specifically on any evaluation process of the standard of these; their short paper only reported on trainees' self-evaluation of their skills and confidence in using the tools.

Freeman et al (2005) reported on a state-wide plan in Kansas for embedding PBS into social care organisations, via a longitudinal course and although data were still being collected at time of report, nine BSP had been produced and had been assessed via a tool based on that devised by Horner et al (2000). This tool contains 37 items that are scored (in place, partially in place and not in place). Average scores moved from 48.31% pre-training, to 72.79% post-training.

Wardale et al (2014a & 2014b) both reported on improvements in writing BSP as a result of training; they used the Behaviour Support Plan Quality Evaluation Guide (BSP-QE) (Browning-Wright et al, 2003) to evaluate this. This tool is described as being based on six key concepts of PBS including behaviour function and changing behaviour via changing environmental features. Participants in the first study developed PBS plans scoring a mean total 13.73, which is rated 'under-developed' and likely to require modification; however they noted that this score needed to be seen in context of the fact that some items were not taught to the participants. The second

study reported scores of 18-20, indicating a rating of 'good' and likely to be helpful in effecting changes in behaviour.

Reid et al (2003) reported outcomes in relation to trainees developing an understanding of PBS and of the skills associated with it, rather than actually producing FA or BSP. They reported on training for supervisors of direct care staff where evaluation was carried out via paper exercise, role-play and on-the-job checks; 85% (328) supervisors completed the training successfully and performed all on-the-job skills at mastery criteria.

3.5.2 Service User Studies

Participant numbers for the eight studies that included service user outcomes, ranged from 5-138. In terms of design, only one study used a control group (Grey & McClean, 2007), and three studies included reliability data (Baker, 1998; Grey & McClean, 2007; McClean et al, 2005). Six studies used a repeated measures design, pre and post training (Baker, 1998; Crates & Spicer, 2012; Dench, 2005; Gore & Umizawa, 2011; Grey & McClean, 2007; Reynolds et al, 2011). In addition, McClean et al (2005) and McClean & Grey (2012) had a third data collection at follow-up.

Outcome measures used in each study to evaluate different variables were identified and are described below; these are summarised in Table 3.6.

Frequency of behaviour: Five studies looked at this using real-time behaviour recordings. Baker (1998) found that in the two months post training, challenging behaviour for two service users reduced to zero and for the other three reduced to between 11-28% of baseline. McClean et al (2005) reported significant improvement (defined as reduction to between 0 and 30% of the baseline rate of behaviour) in 77% of the group at an average follow-up of 22 months; in 2007 Grey & McClean reported 66% of the group significantly improved using the same definition; and in 2012 McClean & Grey reported an average decrease of 61% at three months. Dench (2005) reported reductions to less than 30% of baseline within three months for 56% of the group, and Crates & Spicer (2012) reported a mean change for the whole group to 49.6% of the baseline at three months.

Severity of behaviour: two studies reported specifically on severity: Reynolds et al (2011) used the Aberrant Behaviour Checklist (Aman & Singh, 1986) and reported significant decrease in the staff scores for this measure. Crates & Spicer (2012) described the use of an episodic severity measure (La Vigna & Willis, 2005b) which was based on the real-time behaviour recording; they reported a mean reduction to 30.8% of the baseline.

Frequency, severity and management difficulty of behaviour: three studies used the Checklist of Challenging Behaviour (CCB) (Harris et al, 1994) which allows for separate scoring on each of these three elements of behavioural measurement. Using this measure Grey & McClean (2007) reported significant differences between treatment and control group in all three areas.

Gore & Umizawa (2011) reported a significant decrease in frequency one month after training, but not in severity or management difficulty. McClean & Grey (2012) used the Challenging Behaviour Rating Scales which is drawn from the CCB and reported significant reduction in each of the three areas.

Quality of life: only one study considered the impact of PBS training on the quality of life of service users (Dench, 2005); the Quality of Life Questionnaire (Schalock & Keith, 1993) but it showed no changes.

Table 3.6: Service User Outcome Measures from Studies in Systematic Review

Variable Measured	Outcome Measures	Studies Reporting
Frequency of behaviour	Service incident reports	Baker (1998)
	Behaviour recordings	McClean et al (2005); Dench (2005); McClean & Grey (2012); Crates & Spicer (2012)
Frequency, severity and management difficulty of behaviour	Checklist of Challenging Behaviour (Harris et al, 1994)	Grey & McClean, (2007); Gore & Umizawa (2011); McClean & Grey (2012)
Severity of behaviour	Aberrant Behaviour Checklist (Aman & Singh, 1986)	Reynolds et al (2011)
Quality of life	Quality of Life Questionnaire (Schalock & Keith, 1993)	Dench (2005)

3.6 Discussion

This review presents data from 17 studies of PBS staff training. Although all the studies are based on PBS training, there are considerable differences between these studies. The length of training is a fundamental difference, with a range from three half-day sessions, to a two-year university diploma course. The studies also vary considerably in their measures, some measuring staff outcomes only, some service user and some both. Even those studies all measuring staff outcomes have a range of measures, some focus on emotional responses, others on knowledge acquisition, and some on skills acquisition. Despite these differences, however, there are a number of observations that may be made.

3.6.1 Limitations to Studies

It is perhaps useful to begin by considering some of the limitations to the literature as there is a range of different research designs used and some of these are less robust. Only one study (Grey & McClean, 2007) used a control group throughout the process (Reid et al, 2003 used it in the first phase), and neither of these were randomised or used blind assessors; the lack of a control group reduces confidence in results, as it cannot be ruled out that any changes following training are not in relation to some other organisational initiative. In addition, less than half the studies included reliability data in relation to how data were collected on the various measures, so this undermines confidence in these data, in particular for the real-time recording of behavioural challenges, which is based on individual staff observational records, completed at the time of an incident. Data entry was also poorly addressed with only one study reporting on this (McGill et al, 2007).

As well as lack of reliability measures for data collection, none of the studies measures reliability in terms of adherence to training protocol; treatment integrity is an issue that should be reported, and in these studies PBS training could be regarded as the treatment. Treatment integrity data are essential to the internal validity of any research outcomes based on intervention as without it, it is difficult to measure the effectiveness of an intervention. Crates & Spicer (2012) refer to the trainers being monitored in terms of adherence to the La Vigna training model, but no data is presented regarding this.

In terms of experimental design, the most common design was repeated measures, although three studies only presented post-training data in relation to staff measures (Crates & Spicer, 2012; Dench, 2005; La Vigna et al, 2002). Only collecting post-training data makes it impossible to evaluate following training. One of these studies (La Vigna et al, 2002) uses a weak design, with results based on a self-report postal survey completed by participants following training; this makes it very difficult to have any confidence in their results.

Follow-up data collection is also limited in many of the studies so that it is impossible to evaluate maintenance of training effect. Three studies had follow-up data collection (McClean et al, 2005; McClean & Grey, 2012; McGill et al, 2007).

Although all studies met the criteria of PBS training, some of them met a specific standard to a lesser degree than others. Lowe et al (2007b) for example, was ten days of consecutive teaching and then individual support for practical application in the workplace. It is therefore likely that the 'person-focused' element was far less strong than in for example McClean et al (2005), where the training involved a focus service user upon whom all the work of the course was based and for whom students wrote a full functional behavioural assessment and behaviour support plan. These differences may also have impacted on results.

3.6.2 Staff Outcomes

A number of studies only reported changes in staff knowledge, attitudes, attributions, or emotional responses (Gore & Umizawa, 2011; Grey et al, 2002; Lowe et al, 2007b; McGill et al, 2007; Wills et al, 2013). While there is merit in exploring these variables and significant changes were found following training, changes in staff behaviour were not measured. Thus, it is difficult to extrapolate these results to changes in staff practice. It might be logical to assume that changes in staff beliefs, for example, about challenging behaviour may lead to changes in support to individuals. However, there is little evidence to support this assumption, particularly since associations between staff beliefs and self-reported behaviour are weak (Jones & Hastings, 2003). Stokes & Baer (1977) referred to this as a 'train and hope' model (delivering training and hoping that staff will implement positive changes back in the service). In addition, the results were mixed, particularly for attributions and emotions, with only some studies showing significant changes, despite the sometimes considerable length of the training programme. This is in contrast to the more impressive results in relation to changes in challenging behaviour and raises the question as to why the staff measures appear to be less positive; for example is PBS training actually most effective in reducing behavioural challenges, rather than having an impact on staff knowledge, attitudes and emotions? Although there is substantial emphasis on values within the definition of PBS, it appears that this does not transmit as successfully as perhaps the 'technology' side of PBS, which addresses behaviour change. There are some clear links here with the findings from the review of PBS intervention effectiveness in chapter 2, where the lack of evidence for lifestyle changes was evident, but there was much better evidence for reductions in challenging behaviour.

Of the 13 studies focusing on staff outcomes, eight reported on staff skills. Of these, five studies reported on skills applied in practice (where PBS strategies are implemented and results for service users are reported). The remaining three studies (Freeman et al, 2005; Wardale et al, 2014a & 2014b) focused on related skills (writing BSPs). Of the five studies that focused on skills applied in practice, four supported trainees to write and implement an FA and BSP for an individual (Baker, 1998; Crates & Spicer, 2012; Dench, 2005; La Vigna et al, 2002). In addition, Reid et al (2003) used on-the-job checks and assessed role-play to teach a range of PBS skills. These studies however used a weak design in terms of their staff measures, with only post-training evaluation of skills, rather than pre and post; this leaves less confidence in results and lack of evidence in relation to improved skills.

It seems likely that skills taught in relation to direct practice are more likely to have an impact on practice than those that are not. It is therefore likely that research focusing on actual staff behaviours while providing support, rather than knowledge, feelings of efficacy, or causal

attributions may have more of an impact on practice and on support to service users. The literature in relation to Active Support is a helpful comparison (e.g. Mansell & Beadle-Brown, 2012) as it stresses on-the-job coaching as a necessary element of effective staff training and has demonstrated outcomes in achieving change of practice in staff's own workplace. Following Active Support training, staff practice has been shown to have improved, for example, increased levels and quality of assistance to service users, increased opportunities for choice-making provided to service users, increased contact to service users (Beadle-Brown et al, 2012; Jones et al, 1999; Mansell et al, 2008; Stancliffe et al, 2007). The Active Support model of training has some links with the person-focused model of PBS training, where skills are applied in practice, for a specific person; rather than just undertaking theory-based training participants are supported to apply their learning in real life and in direct practice with an individual service user. It appears that the studies using this approach, which supported staff to systematically apply their new skills in practice (e.g. Baker, 1998; Dench, 2005; Crates & Spicer, 2012) are more likely to impact on staff practice, and therefore also potentially on service user behaviour, than those which more tangentially refer to applying skills in practice (e.g. Gore & Umizawa, 2011; Lowe et al, 2007b; McGill et al, 2007). Unfortunately there are no studies that robustly evaluate staff skills in practice which also evaluate service user outcomes, and this seems to be an unfortunate lack. Studies that have thoroughly evaluated service user outcomes, for example McClean et al (2005) or Grey et al (2007) would also have benefitted from presenting staff outcomes. Perhaps the same difficulty would have applied as with Dench (2005) and Crates & Spicer (2012), which are both based on the same La Vigna multi-element model training programme, in that these do not include staff data prior to training. However given the substantial training programme undertaken and the fact they present service user data which is arguably more difficult and time-consuming to achieve, then it is disappointing that data about staff skills and practice could not have been included also. Given this lack, it is difficult to make any clear conclusions about the PBS training having an impact on staff practice, leading in turn to an impact on support to service users.

3.6.3 Service User Outcomes

Of the eight studies including measures of outcomes for service users, four are based on work done in Ireland by The Brothers of Charity and the Callan Institute. Dench (2005), Grey et al (2007), McClean et al (2005), and McClean & Grey (2012) all showed significant reductions in challenging behaviour following longitudinal person-focused training, demonstrated both by real-time behaviour recordings and by use of the CCB. In addition, both Baker (1998) and Crates & Spicer (2012) showed considerable reductions in challenging behaviour via their longitudinal training. Baker demonstrated a reduction from baseline for five individuals following a three-session PBS training course. (Although this training was shorter than the Irish studies,

each session was a month apart, and follow-up data was taken two months after the final training session, so it seems fair to describe this as longitudinal training also). Crates & Spicer utilised a training-for-trainers format in their Tasmanian study, based on the La Vigna et al (2005) model, and providing a training model similar to the Irish studies. It is difficult to come to conclusions based on the limited number of studies, but it may be that there are different levels of practical application within person-focused training, and that where there is a more systematic approach to apply the learning from PBS training in practice, then positive changes in levels of service user challenging behaviour are also more likely.

One of the key findings from this review is the lack of evidence for quality of life changes following PBS training, a factor also highlighted in the previous chapter with regard to PBS intervention. Only one study of the 17 reported directly on quality of life outcomes for service users; Dench (2005) attempted to evaluate outcomes for services users following a longitudinal PBS training course. However, despite positive anecdotal evidence of lifestyle changes, these were not able to be measured and reported; she notes this may have been due to the Quality of Life Questionnaire (Schallock & Keith, 1993) being insufficiently sensitive to measure these changes.

This appears to be a major challenge to the field of PBS training; why is it so difficult to evidence quality of life changes, when this is a key element of PBS, and should be a major focus of any training? It may be that the emphasis in both PBS training and PBS interventions is more specifically on changing behaviour and there is an expectation that quality of life changes will follow automatically, almost as an added extra. Perhaps there has been an underlying assumption that challenging behaviour is the main block to good quality of life, and that if service users can be supported by PBS to move on from their challenging behaviour, then good quality of life will automatically follow. However that assumption is undermined by these findings, as despite some impressive reductions in behaviour, this review demonstrates quality of life changes either did not follow changes in behaviour or were not measured. It may be that in addition to PBS training focusing on reducing behaviour, additional training is required to address quality of life.

Another potential option is that PBS training is having an impact on quality of life, but that this is difficult to evaluate because of the method of delivery, the person-specific focus, the stakeholder involvement and the non-scientific approach, may make it more difficult to collect data and to evidence this in a robust way, suitable for research purposes. It may also be that because quality of life changes are so person-specific, that this is difficult to evidence through training, as the measures to evaluate quality of life changes would need to be specifically defined and developed for each person (e.g. as in Clarke et al, 2002) and therefore do not easily lend themselves to synthesis and comparison.

3.7 Conclusions

There are a number of implications for research and practice arising from this discussion of PBS training. In terms of looking to the future, there is a need to be creative and innovative in how training is delivered, both to maximise resources and to achieve the coverage of training needed. Use of information technology, e-learning and video training are measures which have demonstrated some element of success (Macurik et al, 2008 & Sailor et al, 2000), and may be useful in this context. Further exploration of their use in relation to PBS training would be appropriate.

In addition to creative use of technology, coverage could also be increased by utilising a training-for-trainers model. Crates and Spicer (2012) demonstrated that ‘second generation’ training could be as effective as that delivered directly (e.g. La Vigna et al, 2005). This is an encouraging model in terms of achieving high quality training with reasonable coverage. Rotholz & Ford (2003) also utilized this approach in that trainees who successfully completed the PBS training for supervisors were eligible to enrol in a two-day trainers’ course and were assessed in their training skills. (This study has not been discussed in detail as the data were already reported in Reid et al, 2003). These elements of training may be useful as a means of furthering an organisation’s ability to provide PBS training to their staff.

There is a need for considerably more research on the impact of PBS training for service users in terms of frequency of behaviour and also severity. It would be useful to explore what supports good generalisation and maintenance of benefits from PBS training; thus, studies including long-term follow-up would be particularly helpful. Although it seems likely that staff training which supports the practice and application of PBS skills within the workplace will be more effective in producing positive changes in the lives of service users, research is needed to support this assumption.

It would be useful for research to examine what makes training effective for staff and service users, for example, considering issues around organisational environments that promote PBS training and practices. These might include the impact of providing brief PBS training for supervisors of trainees undergoing in-depth PBS training; the impact that good contextual fit (Albin et al, 1996) has on implementation of PBS plans; and how organisational systems, such as the Periodic Service Review (La Vigna et al, 1994) may be used to support implementation.

In all of these areas, it would be ideal to see use of more robust research designs such as powerful single case or group experimental designs, including randomised controls and use of blind assessors, in order to isolate the influences of different variables and allow clearer conclusions. More of an emphasis on the use of established measures for ensuring the reliability of the data presented would be helpful in order to synthesise research findings and more easily make comparisons between studies.

Finally, as noted, it has been difficult to show systematic links between PBS training and improvement in quality of life for service users, therefore more research is needed to specifically explore quality of life outcomes for service users following PBS training. Research into the impact of staff training in PBS has focused more on outcomes for staff than service users. The eight studies that did consider outcomes for service users demonstrated significant reductions in challenging behaviour; however, improved quality of life for those individuals was not evidenced by the training and this is a concerning omission. Based on this review it appears that the field of PBS training may need to adapt or develop, in order that it not only has an impact on challenging behaviour, but can also demonstrate quality of life changes, particularly since this focus on quality of life is one of the key elements that defines PBS. It would therefore be useful for future PBS research to focus specifically on evaluating impact on the lives of service users resulting from staff training in PBS.

3.8 Chapter Summary

This chapter has reported on a systematic review of PBS training. First the development of PBS training was discussed and a definition of PBS training was presented which identified key features of training format as well as PBS content. Based on this definition a search was carried out and 17 studies were identified that reported on outcomes from training in PBS. The outcomes from these studies were summarised and discussed with comparisons made of their reported data. A major finding from this review was the lack of evidence of impact of PBS training on the quality of life of service users; in contrast to impact on challenging behaviour which has been well-documented, there is very little evidence for impact on quality of life.

4 Introduction to Study

4.1 Purpose of Study

The purpose of the study was to evaluate outcomes of Positive Behaviour Support (PBS) training for first level social care managers. This study aimed to measure the impact that this training had, not only on the individual managers who received the training, but also on their staff teams, and on the service users they supported. There are difficulties within social care in achieving a well-trained and skilled work-force, particularly for staff working with individuals with challenging behaviour where turnover is often high and time available to free up care staff for training is limited.

This study therefore measured the impact of PBS training on those who did not directly receive it but who had it cascaded down to them by their manager. It was judged that if this could be shown to be an effective method of developing the skills of direct care staff, then it would be an area for development within social care organisations in terms of how they approach training their staff.

The study also measured the impact of PBS training on the lives of service users, both in terms of reduction of challenging behaviour and increase of quality of life. Although some research has previously considered impact of PBS training on the challenging behaviour of service users (see previous chapter), this has been limited in terms of long-term follow-up. This study attempted to address this by providing follow-up data (20 months after baseline) for service users whose manager attended the training.

In addition this study measured the impact of the PBS training on service users' quality of life, looking at quality of support and opportunities to participate. PBS has at its core a commitment to not only reducing the level of challenging behaviour for individuals, but also to improving their lives. Given the limited evidence of quality of life changes in relation to training in PBS (see Section 3.6.3), it was felt that was an appropriate area for this study to address. This included: engagement in activities, participation in leisure, and use of community facilities.

The study aimed to address the question – *can providing managers with training in PBS have long-term, positive outcomes in the following areas: for managers, in terms of how they support their staff with regard to challenging behaviour; for staff in terms of how they provide support to people with challenging behaviour; and for service users in terms of bringing about reductions in their challenging behaviour and improvements in their quality of life?*

4.2 Theory of Process of Change

In providing training to one group of individuals (managers), and then exploring the impact of this training not only on them, but also on another group of individuals (their staff and the service users supported by their staff), this study was adopting a particular theory about the process of change likely to occur both in terms of challenging behaviour and also quality of life (QoL), although it was acknowledged that the two areas overlap and were unlikely to change in isolation from each other. The process of change is shown in Figure 2 and Figure 3, and further explanation of these figures is below.

4.2.1 Challenging Behaviour

In terms of the impact on challenging behaviour it was hypothesised that managers who undertook the training would gain increased knowledge and understanding of PBS and in particular would learn about potential reasons for challenging behaviour occurring. This was expected to help them identify functions of behaviour for the specific service users they support and as the training was person-focused, managers would be supported to apply this learning to particular individuals. This would allow managers to develop person-specific PBS plans⁵, based on the functions of behaviour for the specific service user. Through this process they would also be supported to cascade this learning to their staff teams via a number of management systems such as team meetings and supervision. It was expected that staff who received this input from their manager would have a better understanding of challenging behaviour and would begin to have more positive attributions about specific service users that they supported; this in turn would lead to increased helping behaviour to service users and more motivation to follow the PBS plans. From the service users' perspective, it was expected that they would then experience more positive contact from staff, which would help develop relationships. Based on staff's better understanding of behaviour and their implementation of function-based PBS plans, it was expected that service users' needs were more likely to be met, resulting in less challenging behaviour.

4.2.2 Quality of Life

In terms of the process of change in quality of life it was hypothesised that the managers who undertook the training would learn skills in practice leadership, as course content included a focus on monitoring staff's direct support of service users via a process of observation, feedback,

⁵ Throughout this study 'PBS plans', refers to the overall behavioural plan of support; within this there may be a number of individual specific plans, for example a communication plan or an activity plan, or a reactive plan. For ease of differentiation, these individual plans are referred to in this study as 'PBS protocols'

role-modelling and coaching. It was expected that this would lead to increased use of practice leadership by the manager and that on the basis of this staff would receive better direction and support from their manager and would therefore be clearer about their role in supporting service users' engagement and participation in activity. This increased direction from the manager was also expected to lead to a more organised and focused staff team who would work together better to provide increased assistance to service users. In addition, managers would learn the importance of ensuring PBS plans focused not only on reducing challenging behaviour, but also on improving quality of life (the majority of new PBS protocols written as part of the training would be proactive, with a quality of life focus). The implementation of these PBS plans would generate a range of opportunities and activities with a focus on improving quality of life. From the service users' point of view, it was expected that both the increased assistance and increase in quality of life opportunities would result in increased engagement in activities and more participation in leisure and community activities, as well as in domestic and home-based tasks. Via this increased participation and engagement, it was felt that service users were likely to be happier and to experience an improvement in their quality of life.



Figure 2. Theory of Process of Change for Challenging Behaviour



Figure 3. Theory of Process of Change for Quality of Life

5 Method

5.1 Chapter Outline

This chapter provides an overview of the method used to carry out this study. This includes the study design, participants, procedure, measures, data collection, and reliability checks. There is a description of how data were analysed, along with an overview of how missing data were handled, including an outline of the ‘intention to treat’ approach and the use of imputed data. This included results from reliability checks and data regarding dropout. Finally an updated version of the theory of process of change is presented in relation to the measures used in this study.

5.2 Design

The study was a non-randomised control group study with both between-group and within-group comparisons. Training lasted a year and was carried out in two cohorts over two years; there were two cohorts of the experimental group and also two cohorts of the control group. The study was non-randomised due to the necessity of the needs of services; managers of services supporting service users with challenging behaviour that was causing the most severe difficulties were prioritised for access to the training. To have done otherwise and to have randomised access to the training between these high priority services and other services experiencing lesser difficulties would have been difficult to justify operationally, as those in most need of this training would potentially have been denied priority access to it.

However, it was recognised that there was a potential impact on this study in relation to non-randomisation; the impact of non-randomisation was therefore considered for each of the participant groups. Baseline measures were carried out for each of the groups (managers, staff, and service users), and analysis of these demonstrated whether there were any differences between groups. This area is discussed in more detail in the data analysis section, 5.10.

5.3 Participants

Participants in this study were 72 (of which 50 were experimental and 22 control) frontline managers in social care working within a community-based service provider; the staff managed by these managers; and 72 individuals with learning disabilities supported by the services managed by these managers.

5.3.1 Recruitment

The experimental group for this study was recruited from frontline managers enrolled on a Positive Behaviour Support (PBS) training course offered within the social care organisation in which they worked. These managers were nominated for the course by senior management and then completed an application form. They were then interviewed on an individual basis along

with their line manager in order to discuss the work of the course in more detail and to ensure they understood the commitment of time and resources required. Once managers were accepted onto the training course, they were then approached to ask for their consent to participate in the study; the process for obtaining consent is outlined below. Managers in the control group were also recruited by nomination by their line managers, senior managers within the organisation. They were frontline managers who managed similar services as the managers in the experimental group, and were eligible for attending the training – only had not been prioritised for attending the training on this occasion. Once these managers were nominated then they were contacted, and managers and service users' consent was sought as described below.

Staff teams were recruited on the basis that their manager had agreed to be part of the study. They were then asked if they were also willing to be part of the study and consent was sought as described below.

The managers in the experimental group selected a service user with whom to complete the work of the course. Service users in the experimental group were therefore recruited by the fact they were selected to be the 'focus person' for this training – although they were offered the opportunity to benefit from this training without being part of the study. For the control group, once a manager gave consent to participating in the study, then they were asked to select an appropriate service user to also participate in the study. Once this person was suggested, then the consent was sought as described below.

5.3.2 Consent

Consent from managers was obtained before they were included in the study. Potential participants were sent a letter, explaining the purpose of the study and inviting them to participate (sample information sheet in appendix two); individual meetings were then held with any potential participant who wished further information. Participation was voluntary and no implicit or explicit coercion was used. Although line managers made the original nominations for managers in the control group, the line managers were not involved in this consent decision; in both groups the frontline managers themselves had the right to consent about their own involvement. If they consented they were asked to sign a consent form indicating their agreement to be involved in all aspects of the study.

Once the manager of a service gave consent, then each of the staff they managed was given a written information sheet outlining the purpose and nature of the study and describing what would be required of them should they take part. If they agreed, they were asked to sign the consent form which indicates consent to involvement in all aspects of the study. Even where a manager agreed to take part, it was not necessary for the whole staff team to also take part as carrying out the measures with some of the team was sufficient. This therefore allowed for some

staff within a team to say no, but for the team as a whole still to be included. If any staff member did not wish to participate, then no filming took place while they were directly supporting the service user. The information sheet for both staff and managers explicitly stated that should they wish to refuse consent, or to withdraw from the study at any time, there would be no negative impact on them or their careers as a result.

Some of the service user participants in this study had capacity and were able to give consent. Wherever this was a possibility, then augmented communication methods were used by the manager, along with staff who knew the individual well to try and help service users understand the process and what was being asked of them. For several individuals this included the use of visual communication such as graphics and symbols, and advice was sought from individuals who knew the service user best in order to find the best way to facilitate understanding and informed consent.

However, it was not possible to carry out the research only involving those who were capable of giving consent, as most of the service users with severe learning disabilities did not have capacity to give informed consent. For these individuals, the principles of the Adults with Incapacity Act Scotland (2000) (AWI) (Scottish Government (SG), 2000) were considered: did the study further knowledge; was it of benefit to the adult or others in a similar condition; did it entail little or no risk or discomfort; was it stopped if the adult objected; and was consent was sought from a person with relevant powers, a guardian or family member⁶. In the experimental group 41 people lacked capacity (82%) and for this group consent was given by a welfare guardian (who may also have been a family member) for 26, and consent was given by a family member for 15. In the control group, 19 people lacked capacity (86%), and consent was given by a welfare guardian (who may also have been a family member) for 12, and by a family member for seven.

5.3.3 Capacity

Issues of capacity were initially considered by the manager of the service and others who knew the service user well. Guidelines from the Scottish Government's publication *Communication and Assessing Capacity* (SG, 2008) were followed.

It was recognised that in line with the principles of the AWI Act that the starting point is presumption of capacity. Based on reading of the AWI Act and Code of Practice (Scottish Government, 2010), along with good practice guidelines from the Mental Welfare Commission

⁶ The AWI differs from the Mental Capacity Act in England in that there is no 'consultee' status within it, therefore only those who were legal welfare guardians or were family members were approached for consent.

(Mental Welfare Commission, 2010), capacity was taken to be the ability to understand information relevant to a decision and to appreciate the consequences of making this decision or not. To demonstrate capacity individuals needed to:

- Understand broadly the nature and purpose of the research
- Understand the benefits and risks of being involved, and be able to make a choice about this
- Retain the information long enough to weigh it up and make a decision
- Communicate that decision.

Where capacity was in doubt advice was sought from relevant others, for example, members of the multi-disciplinary team involved in the service user's support.

5.3.4 Demographic Information

Demographic information was available for both managers and service users; it was not sought for staff, as individual staff were not represented in this study, rather they were regarded as a group.

5.3.4.1 Characteristics of Managers

Details regarding managers are outlined in Table 5.1, where both cohorts of the experimental group (50 participants) and both cohorts of the control group are combined (22 participants). This includes age, gender, length of time as a manager, numbers of service user supported, number of staff managed, and number of hours of support managed.

Table 5.1: Managers' Characteristics at T1, Experimental and Control Groups

	Experimental (n=50)	Control (n=22)
	Mean (Range)	Mean (Range)
Age	40 (25-55)	42 (25-57)
Gender	Male 8 (16%) Female 42 (84%)	Male 5 (23%) Female 17 (77%)
Length of Time as Manager (in years)	4 (1-14)	3 (1-10)
Number of Service Users Supported	5.26 (1-22)	7.82 (1-25)
Number of Staff Managed	11.46 (3-24)	13.55 (4-25)
Number of Hours of Support Managed (per week)	322 (105-585)	391 (165-685)

5.3.4.2 Characteristics of Service Users

Details regarding service users are outlined in Table 5.2, where both cohorts of the experimental group (50 participants) and both cohorts of the control group are combined (22 participants).

This includes age, gender, presence of an autism diagnosis or not, ABS scores, ABC scores, whole time equivalent (WTE) staff ratios, observation staffing ratios and numbers of service user per setting.

Table 5.2: Service Users' Characteristics at T1, Experimental and Control Groups

	Experimental (n=50)	Control (n=22)
	Mean (Range)	Mean (Range)
Age	41 (18-63)	39 (19-61)
Gender	Male 35 (70%) Female 15 (30%)	Male 13 (59%) Female 9 (41%)
Autism Diagnosis	Yes 23 (46%) No 27 (54%)	Yes 10 (45%) No 12 (56%)
ABS Scores	133 (35-262)	148 (52-248)
ABC Total Scores	57 (13-99)	35 (5-104)
ABC Number of Severe Behaviours	6 (0-21)	3 (0-25)
Staff to Service User Ratio (WTE)	3.00:1 (0.5:1-6:1)	2.19:1 (0.79:1-4:1)
Staff to Service User Ratio (observation)	1.08:1 (0.5-2)	0.94:1 (0.5-1)
Number of Service Users per Setting	1.1 (1-4)	1.36 (1-4)

The demographic information indicated that most managers managed support for more than one service user; however it also showed that most service users were living on their own, particularly in the experimental group, where the average number of service users per setting was 1.1. This was due to the fact that most managers managed a number of small houses, often only with one service user in each. There was also a high staff to service user ratio, with the equivalent of three full-time staff for every service user.

5.4 Group Size

The sample size of the experimental group was influenced by operational constraints and how many managers could be freed up to attend the training; it was decided that 25 managers per cohort was an acceptable number. From previous experience of running the training, a 20% drop out rate was expected. This would therefore leave 40 managers in the experimental group after drop-out.

It was identified early on in the study that achieving identical numbers in both the experimental and control group would be difficult, due to the perceived burden and lack of benefit in being part of the control group. It was therefore suggested that a control group of around half the size may be acceptable, 22 which may reduce to 18 following drop-out

Required sample size was calculated using data from previous studies on challenging behaviour using the Aberrant Behaviour Checklist (Aman & Singh, 1986), as this is the primary

dependent variable in the current study. In a randomized controlled trial of the impact of PBS on challenging behaviour, Hassiotis et al (2009) estimated an effect size of 0.8. Effect size was therefore set at 0.8. With power set at 0.8 and $p < 0.05$ (one tailed) a sample size of 58 (40 in experimental group and 18 in control group) is sufficient.

5.5 Ethics

As this study included people who lack capacity, ethical approval was sought from the ‘Scotland A Research Ethics Committee’. This was granted (see appendix three).

5.6 Measures

5.6.1 Overview of Measures

There was a substantial number of measures used in this study, therefore for clarity these are summarised below in Table 5.3 for each of the participant groups; each measure was used at each data collection point for both groups, unless otherwise specified. Following that, each of the measures is listed alphabetically and described. Data collection and reliability for each measure are described in section 5.8.

Table 5.3: Measures Used in the Study for Each Participant Group

Managers	Staff	Service Users
Demographic information (T1 only)	PBS knowledge test	Demographic information (T1 only)
PBS knowledge test	Challenging Behaviour Attribution Scale	Adaptive Behaviour Scale (T1 only)
Challenging Behaviour Attribution Scale	Active Support Measure	Aberrant Behaviour Checklist
Periodic Service Review (T2 and T3 only; experimental group only)	Momentary Time Sampling	Behaviour recording forms
Practice Leadership questionnaire		Momentary Time Sampling
		Guernsey Community Participation and Leisure Assessment

5.6.2 Aberrant Behaviour Checklist

The Aberrant Behaviour Checklist (ABC) (Aman & Singh, 1986) is an informant-based scale designed to assess challenging behaviour in people with learning disabilities. The ABC was chosen for this study in preference to other measures of challenging behaviour, for example, the Checklist of Challenging Behaviour (Harris et al, 1994), as it is very well-used within the literature, and in particular has been used for other studies in relation to PBS, for example Allen et al (2011), Hassiotis et al (2009), and McGill et al (in press).

The ABC has demonstrated reliability and validity: average internal consistency 0.90; inter-rater correlation 0.63; test-retest reliability 0.98 (Aman & Singh, 1986). In this study the internal reliability of the ABC was tested at T1 and Cronbach's alpha for the ABC total and each of the sub-scales was as follows: 0.49 (Total), 0.25 (Irritability), 0.10 (Lethargy), 0.84 (Stereotypic Behaviour), 0.80 (Hyperactivity), 0.82 (Inappropriate Speech). Cicchetti and Sparrow (1981) report alphas of 0.40-0.59 as indicating fair reliability and above 0.60 as good reliability, although others regard anything under 0.50 as unacceptable reliability (George & Mallery, 2003).

The five sub-scales of the ABC are irritability, agitation, crying (15 items); lethargy, social withdrawal (16 items); stereotypic behaviour (seven items); hyperactivity, non-compliance (16 items); inappropriate speech (four items). It consists of 58 items which may be scored as follows: 0=not a problem; 1=the behaviour is a problem but slight in degree; 2=the behaviour is moderately serious; 3=the problem is severe in degree. Examples of these items include:

- Seeks isolation from others
- Cries and screams inappropriately
- Deliberately hurts himself/herself

5.6.3 Active Support Measure

The Active Support Measure (ASM) (Mansell et al, 2005a) is an observer-completed rating scale which scores the level of staff support in 15 areas relating to carrying out Active Support, for example, offering choice, facilitating participation, providing age-appropriate activities and communication. Examples of these items include:

- Graded assistance to ensure success
- Staff work as a co-ordinated team
- Written plans are in routine use

Each area can be scored on a 4-point scale, where 0=not applicable and increased scores indicate increasing evidence of positive staff practice. It is reported to have good levels of reliability, Cronbach's alpha between 0.779-0.915 (Beadle-Brown et al, 2012).

5.6.4 Adaptive Behaviour Scale

The Adaptive Behaviour Scale (ABS) (Nihira et al, 1993) is a widely used assessment tool for people with learning disabilities and allows description and classification for comparison purposes. It was devised in 1969 and extensively revised in 1993 for adults living in the community. It is divided into two parts: part one is to assess skills and functioning; part two is to assess challenging behaviour. For this study, only part one was used. This consists of 73 items across 10 domains: independent functioning; physical development; economic activity; language development; numbers and time; domestic activity; pre-vocational/vocational activity; self-direction; responsibility; socialisation.

The internal consistency of the ABS is reported to be high with Cronbach's alphas 0.82–0.99 and standard errors of measurement reported to be small (confidence intervals 0.42–2.60). Two-week test-retest reliability is also reported to be high (0.88–0.99), as is inter-scorer agreement (0.83–0.99). Validity is reported to be good at 0.74 (Nihira et al, 1993).

5.6.5 Behaviour Recording Forms

Behaviour recording forms (BRF) are a commonly used observational method to record behaviour, based on the ABC format (Emerson & Einfeld, 2011). These standard behavioural recording forms were completed by staff after any incident of challenging behaviour to record the antecedents to behaviour, the behaviour that occurred and the consequences, or follow-up, to the behaviour. Counting each recording form provided data with regard to frequency of behaviour.

The BRF were also used to calculate the 'episodic severity' (La Vigna & Willis, 2005b) of each incident by allocating a score based on four areas, each of which was scored on a scale of 1-4. These categories were: type of challenging behaviour (ranged from disruptive behaviours to major physical aggression); injury sustained (ranged from no injury to significant injury requiring hospital treatment); restrictions used (ranged from none, to high level physical restraint) and duration of incident (ranged from less than five minutes, to over one hour). In each a higher score indicated a higher level of severity; scores were then added together to give possible maximum severity score of 16. Staff completing the BRF scored each of these categories at the time of completion.

5.6.6 Challenging Behaviour Attribution Scale

The Challenging Behaviour Attribution Scale (CHABA) (Hastings, 1997) was developed to assess what staff may believe about the reasons for people with learning disabilities engaging in challenging behaviour. It was chosen for this study in preference to other measures of attributions, for example The Controllability of Beliefs Scale (Dagnan et al, 2004), as it is well-used within the literature and was therefore felt to be most useful for putting the results from this

study in a relevant context. The CHABA has been used in a range of other studies exploring the impact of PBS training or interventions, for example, Grey et al (2002); Lowe et al (2007b); McGill et al (2007); and Wardale et al (2014b). It has received some criticism: Grey et al (2002) noted that the subscales appeared to lack content validity, in particular the Learned Positive subscale which is intended to measure positive reinforcement, they suggested may actually reflect staff judgements about the intention of the individual. They therefore noted that accurate inferences could not be drawn from the CHABA in relation to how staff might respond to challenging behaviour.

The CHABA is reported to have moderate to good levels of reliability, Cronbach's alpha between 0.65-0.87 (Hastings, 1997). For this study the Cronbach's alpha at T1 was calculated and was found to be as follows: 0.74 (Learned Positive), 0.32 (Learned Negative), 0.56 (Biomedical), 0.55 (Emotional), 0.56 (Physical Environment), and 0.71 (Stimulation).

The CHABA is a self-report measure with 33 statements which give possible reasons for challenging behaviour, for example, 'because she/he is given things to do that are too difficult for her/him', and 'because someone she/he dislikes is nearby'. Staff are asked to score each of the statements on a five-point scale as follows: -2=very unlikely; -1=unlikely; 0=equally likely/unlikely; 1=likely; 2=very likely.

Each statement is related to a causal model for challenging behaviour. These are: learned behaviour, including learned positive (three items) and learned negative (three items); biomedical (six items); emotional (seven items); stimulation (six items); physical environment (eight items). Scores are then calculated for each sub-scale to allow measurement of level of belief in that causal model. A sub-scale score of less than zero indicates that the respondent does not consider the causal model to apply; a score of above zero indicates that the respondent considers that the causal model does apply.

5.6.7 Demographic Information Forms

For managers this contained basic demographic information such as age, gender; previous qualifications, operational information about their management role and about the service they managed. The managers' information also gave details on number of staff managed, number of whole-time equivalent (WTE) staff, number of hours of support provided, number of service users supported, and number of service users per setting. This allowed the calculation of a WTE staff: service user ratio; this was calculated for day-time only, so any waking night staff were not included. Demographic information was not taken for individual staff as the staff measures were reported as a group average, not for individual staff. There was no attempt to identify individual staff, and it was recognised that the actual staff members returning information may be different at each data collection point; this was not seen as an issue, as the study focused on the staff group as a whole in terms of the service they provided, not on individual staff. For service

users, information such as age, gender, presence of diagnosis of autism, service received (including number of hours, type of service, and other service users sharing the environment) was included.

5.6.8 Guernsey Community Participation and Leisure Assessment

The Guernsey Community Participation and Leisure Assessment (GCPLA) (Baker, 2000) is used to assess level and quality of involvement in the community and leisure activities. For this study the GCPLA was chosen over other quality of life measures for a number of reasons; firstly it was used in two other relevant PBS studies, Allen et al (2011) and Hassiotis et al (2009). It was also felt to be a finer measure than for example, the Life Experiences Checklist (Ager, 1990), and was therefore more likely to reflect the types of small changes that may be expected for the client group in this study.

There are 49 items within six categories of community access (services; public transport; indoor leisure; leisure, sport and recreation; social; facilities/amenities). Each of these is operationally defined and scored for frequency of access and support while accessing. Items in each category include:

- Services (doctor, dentist, hospital, police)
- Public transport (bus, train, taxi, boat)
- Indoor leisure (craft, games, TV, videos)
- Leisure, sport and recreation (festival, museum, sport, exercise)
- Social (disco, pub, party, restaurant)
- Facilities/amenities (local shop, high street store, post office, hairdresser)

Frequency is rated as follows: 0=never; 1=very occasionally; 2=3 monthly or more frequent; 3=monthly; 4=weekly; 5=daily. Support is scored as follows: 1=supervised; 2=with carers, but not supervised; 3=unaccompanied; 4=with peer group. This assessment can be administered either directly with the person with a learning disability via an interview, or can be completed by staff on their behalf if they would be unable to complete it.

In the current study only the frequency aspect of the GCPLA was used as it was judged to be more relevant to the client group (whose support from staff was on the whole unlikely to change to any great degree). For the elements of the GCPLA used in the current study it was found to have acceptable inter-rater reliability (0.83 for range and 0.84 for busy); acceptable test-retest reliability (0.83 for range and 0.84 for busy, completed by carers); internal reliability of 0.93 Cronbach's alpha for frequency of contact; scored an average of four on a scale of one to five for relevance in terms of content validity; and concurrent validity with diary recordings was high (correlation for the total categories was $\rho=0.682$) as were correlations with relevant elements of the Life Experiences Checklist (Ager, 1990) ($\rho=0.742$ and $\rho=0.552$).

5.6.9 Momentary Time Sampling

Momentary Time Sampling (MTS) was carried out to measure both the quality of staff support and quality of life for service users. MTS is a non-participant observation method which has been used in research to measure engagement in activity (Beadle-Brown et al, 2012) and to assess quality and quantity of staff support (Mansell et al, 2003).

MTS involves observing behaviour and recording at set intervals, using pre-determined codes. For this study, a 20s interval was used and both staff and service users' behaviours were recorded. 20s was set as the interval as it is judged to be sufficient for providing accurate information over short periods of observation (Mansell, 1985). Only one code could be recorded at each interval, therefore if two behaviours were seen to be happening at the same time, observers had to make a judgement about which was the primary behaviour; guidance was developed to clarify how to judge the primary behaviour and this was used during training and practice (see 5.8.8 below). Due to the nature of the codes this would not be possible for very many codes, but could occur, for example, if a service user was engaged in both social engagement (talking) and non-social engagement domestic (washing the dishes). The codes used were those of Jones et al (1999) as follows:

5.6.9.1 Staff behaviour

Assistance: implicit or explicit instruction to perform an activity; presentation of materials in the context of an activity; gestural, physical prompting, demonstration or guidance to an activity; guiding or arranging materials; praise was incorporated into assistance because they often occurred at the same time, and at times praise was indistinguishable from encouragement/assistance.

Restraint: physical or verbal disapproval without correction or physically preventing activity.

Other conversation: all other interactions, neither encouraging nor discouraging of activity.

Processing: doing something to an individual without assisting their participation.

No contact: not giving the individual any attention or interaction, doing something without them being present; withdrawing from the individual; ignoring the individual.

For the analysis, assistance and other conversation were combined into 'positive contact', and restraint and processing were combined into 'negative contact'. However, assistance individually was also a specific focus for analysis, as there is substantial focus on this in the literature, and it is regarded by many as a key staff behaviour in supporting better engagement for service users (e.g. Beadle-Brown et al, 2014; Beadle-Brown et al, 2015b; Mansell & Beadle-Brown, 2012; McGill et al, in press; Stancliffe et al, 2007; Totsika et al, 2010). In addition, no contact and positive contact were selected for some analysis, as it was expected that there would be changes in levels of overall contact from staff to service users, and especially of positive contact, following PBS training.

5.6.9.2 Service user behaviour

Social engagement: speech or attempts to speak; signs, gestures or attempts to gain attention; giving of attention to another person who is reciprocally interacting.

Non-social engagement – domestic: preparing for, doing, or clearing away a household or garden activity.

Non-social engagement – personal: preparing for, doing, or clearing away a self-help activity

Non-social engagement – other: personal: preparing for, doing, or clearing away a recreational or educational activity.

Challenging behaviour: self-injury, aggression, damage to property, stereotypy or any other inappropriate behaviour.

Disengagement: all other behaviour, including no activity, holding materials, walking/wandering, un-purposeful activity.

For most of the analysis, the four types of engagement were combined and labelled as ‘total engagement’; this allowed consideration of the wider behaviour of engagement, and also supported comparison with a range of literature in relation to engagement (e.g. Beadle-Brown et al, 2014; Beadle-Brown et al, 2012; Bradshaw et al, 2004; McGill et al, in press; Stancliffe et al, 2007; Totsika et al, 2010). On occasions it was useful to split service user engagement into ‘social engagement’ or ‘non-social engagement’ in order to consider specific changes within these categories; non-social engagement was a combination of the three non-social codes, domestic, personal, and other.

5.6.10 PBS Knowledge Test

The knowledge test was comprised of 10 questions based on the questions contained in a DVD training pack developed by the Institute of Applied Behavioral Analysis (La Vigna & Willis, 2009). This had a number of questions added to it, mainly in relation to functions of behaviour and also about definitions of challenging behaviour. The questions were multiple choice and the total possible score for the test was 20; this was an unstandardized test. Internal reliability was tested at T1 for this study and Kuder-Richardson value was found to be acceptable at 0.61. A summary of the areas covered by the questions is below in Table 5.4.

Table 5.4: Summary of Areas in PBS Knowledge Test

Question Areas	Points
Definition of challenging behaviour	3
The 4 functions of behaviour	4
Define functional assessment	3
Differentiate between proactive and reactive support plans	2
Identify different types of positive programming plans	3

Define focused support strategies	1
Identify an antecedent control strategy	1
Identify a stimulus change strategy	1
Identify the need for performance monitoring	1
Identify the reasons for using a PSR	1

5.6.11 Periodic Service Review

The Periodic Service Review (PSR) (La Vigna et al, 1994) is a quality management tool which allows measurement of the extent to which behaviour support plans are being implemented and other recommendations are being carried out. There is no reported validity or reliability data available. For this study, the PSR was developed using an organisational pro forma which had two sections: a section relating to behaviour support plans, and a section relating to operational issues. This pro forma was then personalised for each manager to reflect the person-specific plans in place for the service users and to reflect the particular operational issues pertinent for their service. This was done in workshop three of the PBS training; as the PSR was developed as part of the PBS training, no PSR was in place for the control group.

The process for drawing up the PSR was that each area of the behaviour support plan was clearly defined as a standard that could be scored as occurring or not, for example, *a review of records and a visual check of the communication board show that the person's interpersonal behaviour support plan has been carried out as described every day in the past week*. For every standard met, a point was scored and if a standard was not met, then no point was scored.

Similarly, a range of operational procedures were defined as essential for supporting the process of implementing the behaviour support plan. The PSR included the following: standards around supervision, team meetings, debriefing, staff training, and staff induction. In addition, some services had standards particular to them which were felt to be important for that service, for example, *the staff rota will be prepared and in the service for the 1st of the month; there will be no more than 3 staff off sick in the preceding week; weekly contact will be made with the service user's mother and this contact will be recorded*.

Following the check of the PSR, the total score was converted into a percentage and displayed on a graph in the staff area of the service, so that all staff members received visual feedback of how the service was progressing against its own standards. La Vigna et al (1994) describe a PSR score of 85% as being that likely to indicate consistent implementation of behavioural recommendations (p. 9).

5.6.12 Practice Leadership Questionnaire

A practice leadership questionnaire was developed based on the Staff Experiences and Satisfaction Questionnaire (SESQ) (Beadle-Brown et al, 2005) which assesses staff's job satisfaction, attitudes towards their work, and knowledge. The SESQ was chosen as it is a staff questionnaire with a social care basis, and although its use in research is limited and most of the studies which do use it, are considering the impact of Active Support (Beadle-Brown et al, 2012; Beadle-Brown et al, 2014; Deveau & McGill, 2013), this was felt to be sufficiently similar to the current study to be useful for comparison. For this study, only a small part of the SESQ was used, Question 11 from Section B, which is about practice leadership. In the current study, this is referred to as the practice leadership questionnaire. It focused on staff's experience of three main areas; supervision, team meetings, and observation by manager. Items are scored one point if they meet specific criteria, for example:

- If supervision includes discussion of service user engagement in activity
- If the main topic of supervision is not paperwork
- If the main focus of team meetings is service user engagement in activity
- If the manager observes staff practice at least monthly
- If the manager gives feedback that is at least usually useful

There are ten items, two with two parts, giving a total score of 12, referred to as the practice leadership index; this was then converted to a percentage for comparison.

5.7 Procedure

5.7.1 Format of Training

The training was delivered in two cohorts of 25 managers, each training course lasting a year, with six months between them. The training course consisted of a two-day introduction, followed by eight one-day workshops, six weeks apart. Training was held in five different locations throughout Scotland and groups were small, with high trainer-to-participant ratio; on average two trainers for five participants. This allowed for considerable one-to-one time to be spent with participants, supporting them to write behaviour support plans and to discuss any difficulties or issues that arose. There was individual homework for the managers to complete between each workshop: this consisted mainly of practical tasks, for example, holding a team meeting to teach staff about the function of behaviour for a focus person; role-playing a new support plan with staff; directly observing staff practice and giving them feedback. There was considerable emphasis on teaching the managers to cascade the learning to their staff teams, for example, in each workshop managers were supported to prepare for a team meeting to pass on the learning of that workshop to their staff team.

5.7.2 Aim of Training

The overall aim of the training programme was to teach frontline managers how to provide clear leadership and example in all aspects of implementing PBS, including functional assessment, behaviour support planning, implementation of plans, and recording and reviewing of these plans.

The specific competencies for managers to achieve by the end of the training programme were as follows:

- Identify the function of behaviour for the focus person
- Develop and implement a range of behaviour support plans for that person
- Manage and review staff practice with regard to these plans, via direct observation
- Teach staff how to plan and present activities to the focus person
- Measure staff's implementation of PBS
- Monitor the impact of PBS for the focus person

5.7.3 Development of Training

The content of the training was developed based on the multi-element model of PBS presented by La Vigna & Willis (2005a). This included defining and recording behaviour; functional assessment; proactive strategies such as ecological changes, teaching new skills and reinforcement strategies; reactive strategies.

A variety of tools and strategies were also added from other approaches. For example, in teaching managers about functional assessment, the Questionnaire about Behaviour Function (QABF) (Matson & Vollmer, 1995) was used; in teaching about the importance of meaningful day activity as part of the ecological area of the multi-element model, Active Support (Mansell et al, 2005b) techniques were presented using video and group exercise; and to cover the interpersonal part of ecological plans, a range of alternative communication systems were introduced and demonstrated.

The La Vigna approach to assessment was simplified and managers were not taught to carry out a full-scale behavioural assessment using the La Vigna assessment tool (La Vigna & Donnellan, 1989). Instead they were taught to consider evidence and come to a conclusion regarding the function of their focus person's challenging behaviour via a process of analysing time sampling, staff-completed QABF, and antecedent analysis utilising an organisationally-developed behavioural database.

In terms of behaviour support planning, the La Vigna et al (1989) model was used and managers were taught to write interpersonal and activity-based support plans in the ecological section; coping and tolerance or functionally equivalent plans in the positive programming section; a simple visual reinforcement in the focused support section; and one reactive plan. The

five plans were written in workshops three to seven; workshop one and two taught how to define, record and analyse behaviour; and workshop eight was used for the managers to give a Power Point presentation to the group reflecting on their learning from the training, and explaining how they had passed this learning on to their staff teams. The training sessions are outlined in Table 5.5, with brief details on the content of each session.

Table 5.5: Content of Training Workshops

Training Sessions	Content of Session
Two-Day Introduction to PBS	<ul style="list-style-type: none"> • Functional Assessment • Behaviour Support Planning • Active Support • Alternative Communication • Teaching New Skills • Differential Reinforcement • Reactive Strategies • Managing PBS
Workshop 1	Functional Analysis: Defining and Recording Behaviour <ul style="list-style-type: none"> • Devise person-specific BRF • Severity recording • Individualise time sampling • Use QABF • Prepare for team meeting
Workshop 2	Functional Analysis: Functions of Behaviour <ul style="list-style-type: none"> • Revise functions of behaviour • Analyse BRF, time sampling and QABF • Summarise antecedent and consequence information • Graph antecedent and consequence information • Agree theory on function of behaviour • Prepare for sharing information at team meeting
Workshop 3	Behaviour Support Planning: Activity <ul style="list-style-type: none"> • Agree BSP overview linked to function of behaviour • Revise Active Support • Write activity-based protocol • Devise protocol -specific running notes • Develop direct observation tool

	<ul style="list-style-type: none"> • Prepare for teaching new protocol to staff team
Workshop 4	Behaviour Support Planning: Interpersonal <ul style="list-style-type: none"> • Revise different types of alternative communication tools • Write communication protocol • Make up any visual resources to accompany protocol • Devise protocol -specific running notes • Develop direct observation tool • Prepare for teaching new protocol to staff team
Workshop 5	Behaviour Support Planning: New Skills <ul style="list-style-type: none"> • Revise theory of functionally equivalent and coping & tolerance plans • Write protocol to teach new skills • Devise protocol -specific running notes • Develop direct observation tool • Prepare for teaching new protocol to staff team
Workshop 6	Behaviour Support Planning: Focused Support <ul style="list-style-type: none"> • Revise theory of differential reinforcement • Write a reinforcement protocol • Develop visual resources to accompany protocol • Devise protocol -specific running notes • Develop direct observation tool • Prepare for teaching new protocol to staff team
Workshop 7	Behaviour Support Planning: Reactive <ul style="list-style-type: none"> • Revise theory of reactive support planning • Write reactive protocol • Devise protocol -specific running notes • Develop direct observation tool • Prepare for teaching new protocol to staff team
Workshop 8	Managers' Presentations

5.7.4 Trainers

The training was delivered by members of the organisation's Positive Behaviour Support Team (PBST). Team members had been in their role for an average of 27 months with a range from 18 months to 72 months. They all used PBS in their daily practice and they had all attended a 4-day training course in multi-element PBS delivered by Gary La Vigna from the Institute of Applied

Behavior Analysis; they also received regular internal PBS training from the writer as part of team development. This consisted of team training days held on a quarterly basis and covering all aspects of PBS, for example: carrying out a functional assessment, identifying antecedents and consequences, writing PBS plans, using PSR, and evaluating outcomes from PBS.

They had all been involved in providing this training on previous occasions. Their backgrounds varied, but all had been involved in services for people with learning disabilities in some kind of operational management or supervisory role before joining the PBST. Each workshop group had two members of the PBST facilitating training, although one person was the lead trainer for each group.

5.7.5 Training Integrity

There were a number of measures in place to try and ensure consistency of approach by the different trainers: all trainers used the same training pack and the same PowerPoint presentation; trainers met monthly for training preparation sessions before each workshop; and training was delivered in pairs. However, there were no checks for procedural validity of the training.

5.8 Data Collection and Reliability

In the experimental group data was collected for managers, staff, and service users at three time points: before the training, immediately after the training and at six months follow-up, as shown in Table 5.6. Participants in the control group had data collected at baseline, then immediately after the training of the experimental group; however follow-up data was not collected from the control groups in any of the areas. This is because the initial six month follow-up allowed statistical analysis of whether any changes were due to the training or not. Thereafter it was not necessary to continue to take data from the control group, as maintenance of any changes within the experimental group were the real area of interest. This also minimised the long-term burden of data collection on the control group who were not benefiting directly from the study.

Table 5.6: Data Collection Points for Current Study

	T1: pre training (experimental and control)	Training	T2: post training (experimental and control)	T3: six months follow-up (ex- perimental only)
Cohort 1	December 2011	Jan-Dec 2012	January 2013	June 2013
Cohort 2	May 2012	June 2013 – May 2014	June 2014	December 2014

Details of how data were collected using each different measure are outlined below. It is worth noting that for staff data a representative sample of staff completed the relevant measures at each time point and then their scores were averaged to give a composite staff score. With this process the staff score could include at T2 or T3 those who had not completed the measure at earlier time points; however this was felt to be appropriate since the aim was to assess the impact on the team as a whole of their manager undergoing the training, not for each of them as individuals..

As this was an internal evaluation, carried out by the writer who is an employee of the organisation where the study took place, and is also the manager of the team providing the training, it was acknowledged that there was potential for bias. A number of reliability checks were therefore used in order to try and ensure minimise this as much as possible; reliability for each measure was considered and different approaches to this are outlined below. Unless otherwise specified, data for each measure is collected at every time point; 5% of scoring and data entry for all measures were checked at each time point; any errors noted were rectified.

5.8.1 Aberrant Behaviour Checklist (ABC)

The ABC was completed by the manager during visits to their service by a member of the PBST.

5.8.2 Active Support Measure (ASM)

The ASM was completed based on a two-hour observation of support being provided to the focus person, either videoed or observed, between 4-6pm. It was completed by the member of PBST taking the video or doing the observation (for those focus people who had refused permission for video to be taken). It was completed at the time of the observation. The ASM asks for the number of staff and service users present at the time of the observation, so this information was used to calculate staff: service user ratio during observations.

The ASM was checked for reliability by a second observer who watched 20% of all the videos and re-scored the ASM. This person was a new member of the PBST so did not know the focus person and was therefore blind as to whether the video was from the experimental or control group, and also as to whether it was from before or after the training.

5.8.3 Adaptive Behaviour Scale (ABS)

The ABS was carried out at T1 only and was completed by the manager.

5.8.4 Behaviour Recording Forms (BRF)

Behaviour recordings forms (BRF) were used on each occasion during a specified four-week period at each of the three data collection points. Frequency was calculated by adding together all of the BRF for the four-week period. The BRF were completed by staff in the service after

any incident of challenging behaviour. There was a lead-in period of using the BRF prior to formal data collection in order to try to minimise any issues about accuracy of recording and reminders about completing forms were given to teams if necessary.

Reliability of BRF was confirmed by checking the occurrence of behaviour via a variety of methods. The manager checked staff daily running notes to ensure that no incidents of behaviour were referred to within these that were not recorded on the forms; at the team meeting following the four-week period the manager verbally checked with staff the number of incidents and checked this matched the number of completed behaviour recording forms; and once each week during the four-week period a member of the PBST phoned the service and asked how many incidents had occurred in the last two days; this information was then checked against the number of behaviour recording forms completed for these days (two days was chosen as a suitably short time period where staff were likely still to remember if an incident had taken place or not). None of these checks indicated that there were any incidents of behaviour which had not been recorded on BRF.

The BRF were also used to measure severity. Staff information on each form provided a severity score (based on the four areas: type of behaviour, injury caused, use of any restrictions, and duration of challenge). From all the scores within the four-week period, an average severity score for the period could then be calculated.

5.8.5 Challenging Behaviour Attribution Scale (CHABA)

The CHABA was completed by the manager, and also by each staff member in their team. It was administered to managers at a visit to their service carried out by a member of the PBST, and for staff it was completed individually at team meetings, administered by the manager, at each of the same time points.

5.8.6 Demographic Information Form

This was completed by the manager before being accepted onto the course. This was completed by the manager for the service users, after the service user (or their representative) had given consent to be part of the study.

5.8.7 Guernsey Community Participation and Leisure Assessment (GCPLA)

The frequency part of the GCPLA was completed at each of the time points by the manager during a visit to the service by a member of the PBST.

5.8.8 Momentary Time Sampling (MTS)

MTS was carried out based on a two-hour observation of support being provided to the focus person between 4-6pm, either videoed or observed. It was carried out by the member of the

PBST taking the video or doing the observation (for those focus people who had refused permission for video to be taken).

The members of the PBST were all trained by the writer to code using MTS. A full definition of the codes was written up, with examples of each, detailed definitions, and guidance as to how to choose the primary code if two behaviours were occurring together; codes were based on those from Jones et al (1999). A series of team training sessions took place where team members practiced coding via video, and then checked for reliability in their scores. Any discrepancies were discussed and explained, and then further coding practice took place. This process continued until the group achieved at least a 90% agreement on all codes. This group coding was repeated periodically throughout the life of the study, in order to ensure ongoing reliability.

MTS coding was checked for reliability by a second observer re-coding 20% of the videos; these were a random selection from both the experimental and control groups. These reliability checks were done by new members of the PBST who were trained in coding video but because they were new to the team did not know the focus people or their support staff, so they were blind as to whether the video was from a service in the control or experimental group, and also as to whether it was from before or after the training. Their training took place in the same way as for the whole team as described above. The level of agreement was calculated for each behaviour in each re-coded video for occurrence and non-occurrence using the formulae below (Murphy, 1987, p.228) and then an average level of agreement for each behaviour's overall occurrence and non-occurrence was calculated.

$$R_{occ} \% = \frac{\text{number of intervals observers agreed behaviour occurred}}{\text{number of intervals either observer scored behaviour occurred}} \times 100$$

$$R_{nonocc} \% = \frac{\text{number of intervals observers agreed behaviour absent}}{\text{number of intervals either observer scored behaviour absent}} \times 100$$

Cohen's Kappa was also calculated as this provides an estimate of agreement once levels of chance agreement are taken into account. Kappa was calculated for each behaviour for each of the re-coded videos and then an average Kappa value was calculated for each behaviour using an online calculator (<https://www.easycalculation.com/statistics/cohens-kappa-index.php>).

As the study progressed, it became clear that some of the observation data were noticeably different from comparative literature, particularly service user engagement and staff assistance, which were both markedly higher. This raised some concerns about the coding methods being used in the study by the team of coders; it was therefore agreed as an additional check that the writer would re-code a 5% sample of the videos, from a random sample from each time point and from both experimental and control group. The re-coding focused on the two areas of concern, service user engagement and staff assistance.

5.8.9 PBS Knowledge Test

The PBS test was completed by the manager, and also by each staff member in their team. It was administered to managers at a visit to their service carried out by a member of the PBST. For staff it was completed individually at team meetings, administered by the manager, at each of the same time points.

It was marked by a member of the PBST who was blind as to whether it was from a control or experimental service and also blind as to whether it was from before or after the training. Answers were multiple choices and therefore there was no dubiety about answers being correct or not.

5.8.10 Periodic Service Review (PSR)

The PSR was scored by a member of PBST at a visit to the service for the experimental group only. This was evaluated mainly by consulting records, for example support plan running notes or team meeting minutes; visual checks were also used where appropriate, for example, a visual check was carried out of any communication aids or daily planners, in order to ensure that these were up-to-date.

5.8.11 Practice Leadership Questionnaire

The practice leadership questionnaire based on the SESQ was completed by each staff member in relation to their manager. It was administered by a member of the PBST visiting the service and leaving envelopes which contained the questionnaires for staff to complete, as well as guidance notes for how to complete and return it. Addressed and sealed envelopes were provided so that staff could return the completed forms via the organisation's internal mail without having to go through their line manager.

5.9 Missing Data

In this study an 'intention to treat' approach was taken to missing data; this was for two reasons. Intention to treat (ITT) is a strategy used for the analysis of randomised controlled trials, which includes all participants in the analysis whether they completed the study or not. Although this study is not randomised, it is noted that the effectiveness of the intervention may be overestimated if an intention to treat analysis is not carried out (Hollis & Campbell, 1999). Missing data can affect analysis and lead to misleading conclusions. If participants who have dropped out are not part of the analysis, then there is a possibility that will bias the analysis. Guidance on this topic recommends that no participants be withdrawn from analysis for lack of adherence to a programme (Freidman et al, 1998).

Participants who dropped out of the training may have done so for reasons linked to the programme and therefore to ignore dropouts and exclude them from analysis, fails to consider this potential link. In addition, although they failed to complete the training, they were still part of

the original data and therefore should be followed up. Therefore in the interests of complete transparency and to ensure that the analysis and conclusions of this study are not biased, it was decided to adopt the ITT approach. This is generally accepted as the most statistically robust method of analysis (e.g. Little & Yau, 1996; Steiner & Geddes, 2001). However, it is noted that in an ITT analysis, treatment impact is reduced due to noncompliance or dropout (Gupta, 2011).

The other reason to adopt this approach was the impact on group size; taking the alternative approach to analysis, 'per protocol' (where only participants who have completed the training programme without any violation of protocol are included in analysis) would have substantially reduced the group size. In order to use an ITT approach, imputation of missing data is required; how this is done depends on the type of missing data.

5.9.1 Types of Missing Data

Depending on the reasons for missing data, these situations were dealt with in different ways. If the service user left the study, either through death or leaving the organisation it was not possible to collect ongoing data for them. In these circumstances (without a focus person on whom to focus the training), the manager also left the training and therefore it was impossible to collect any data for the manager or staff. It was not considered useful to follow-up either the staff team or manager, since with the death or departure of the service user, the staff team was re-allocated to work elsewhere and therefore the group of staff and manager whose data were collected at T1 were no longer a specific entity that could be followed up. In this situation an imputation approach was taken; the service user, staff and manager were included in the data analysis, but their data were imputed.

If the manager left the service, their data were imputed but data were still collected for their staff and service user. If the manager left the training, but not the service, then data were collected from them and their staff and service user in exactly the same way and at the same time as if the manager had still been on the training.

Other missing data were from participants who refused to be videoed or observed. Any observation or video sessions of less than 30 minutes were regarded as too short to use, as the shorter length may not be representative of a full session; therefore if any service user ended the session after less than 30 minutes, then this was regarded as refusal and the data were not used. This applied to both the MTS and ASM data.

In all of the above cases none of these data were per protocol (PP), as even where data were actually collected (rather than imputed), these were from participants who had not adhered to the full protocol; therefore none of these data was included in the PP analysis.

5.9.2 Imputation of Data

Imputation of data has an impact on the results and is generally regarded as producing a conservative effect (Hollis & Campbell, 1999). The different ways of imputing data are considered below with some comments about their potential impact:

- Last observation carried forward (LOCF) – this means that the last score for any measure is carried forward for the following scores. The disadvantage of this is that this does not take into account the general trend of the group’s scores, for example, if at a specific time point the group generally were demonstrating a downward trend in a particular measure, then this approach would not take this into account
- Growth Curve Analysis – this calculates the trend of change over time for each participant based on whatever data are available, and then imputes what the missing data would be if the participant continued in this trend. The disadvantage for this approach is that it requires at least two previous data points, so was not suitable for use in this study as many of the dropouts took place after only one data collection
- Group average score (GA) – this means that the average of the group’s score is used to replace missing data. This is suitable for use with data that are missing at random
- Implicit assumption of worst outcome (IAWO) – for this study this means that the missing data would be replaced with the poorest score within the group for that measure at that time point.

In order to take a conservative approach to data analysis, it was decided to use IAWO for any dropout that was potentially linked to a service user’s challenging behaviour. As this study was considering the impact of training in reducing service user challenging behaviour, if a service user’s behaviour had deteriorated to the extent that they were admitted to hospital, then this was regarded as essentially the worst outcome that could occur. To reflect this, they were therefore imputed the worst score that had occurred in the group for that measure.

For any other dropout not related to levels of challenging behaviour (for example the service user dying or a manager leaving the organisation), then GA was used; this was felt to be the most appropriate for this study as it takes into account the expected change in scores and therefore imputes data in the same direction as the rest of the group, which LOCF does not. Missing observational data was imputed as GA.

5.10 Data Analysis

5.10.1 Guidance for Data Analysis

Guidance was sought from the University’s statistics department on several occasions in relation to a number of issues regarding study design and data analysis and on two occasions

meetings took place with representatives from this department. Advice related to the non-randomised nature of the study and how best to address this and also some queries with regard to the possibility of using some participants from the control group at cohort one, as participants in the experimental group in cohort two; following advice, this potential design was abandoned.

Advice was also sought in relation to data analysis, in particular for how best to analyse data when there are two sources of variance (within subjects and between subjects) but when data are non-parametric and therefore a mixed factorial ANOVA cannot be used. The advice was to calculate the change in scores from T1-T2 leaving only one source of variation (between subjects), and then carry out significance tests on that, using parametric or non-parametric tests as appropriate; this approach has therefore been adopted in the current study.

5.10.2 Descriptive Statistics

Mean/median and standard deviation/range were calculated for all measures, and the central tendencies were graphed.

5.10.3 Overall Approach to Analysis

Data were tested for their suitability for parametric analysis using Kolmogorov-Smirnov tests for normally distributed data, and Levene's test to check for homogeneity of variance. Where data were found to not meet parametric requirements, non-parametric tests were carried out; otherwise parametric tests were used with outliers and extreme scores removed. SPSS version 22 was used. Degrees of freedom (*df*) or number of participants (*n*) is shown for all tests.

Where there was a reasonable number of studies which had previously used each measure, relevant literature was checked to see how the measures used in this study have been analysed in previous research, where this was specified; research most similar to the current study was selected for comparison. Providing the data met other parametric assumptions, where the literature was divided on type of tests, parametric tests were used as these are more powerful. Types of tests used in the literature are shown in Table 5.7.

Table 5.7: Type of Test Used with Each Measure in the Comparative Literature

Measure	Studies Using Parametric	Studies Using Non-parametric
ABC	MacDonald et al, 2013; McGill et al, in press; Reynolds et al, 2011	Beadle-Brown et al, 2012; Daynes & Baker, 2014; Lowe et al, 1995
ASM	McGill et al, in press; Totsika et al, 2010;	Beadle-Brown et al, 2012; Mansell et al, 2002
CHABA	Gore & Umizawa, 2011; McGill et al, 2007; Tierney et al, 2007	Grey et al, 2002; Lowe et al, 2007b; Mansell et al, 2008; Wardale, 2014a
GCPLA	Baker, 2000; Baker, 2007	-
MTS	Bradshaw et al, 2004; Mansell & Beadle-Brown, 2011; Mansell, McGill & Emerson, 2001; McGill et al, in press; Stancliffe et al, 2007; Totsika et al, 2010	Beadle-Brown et al, 2012; Jones et al, 1999; Jones et al, 2001b; Mansell et al, 2002; Perry et al, 2011

5.10.4 Significance Testing T1-T2

In order to avoid multiple testing in calculating significance between T1-T2 for both the experimental and control group, a mixed-factorial ANOVA was used with outliers and extreme scores removed. The two groups were considered as a between-subjects factor and the two time points were considered as a within-subjects factor in the analysis. When an ANOVA was not possible due to non-parametric data, then the change in each participant's score from T1-T2 was calculated, and then these data were tested to see if they met parametric requirements. If they did, then a parametric test (unrelated t-test) was used, with outliers removed; if not, a non-parametric test for unrelated data (Mann-Whitney U) was used. This approach of using the change in scores compensated to some extent for the significant differences occurring for some measures between control and experimental groups at T1 and T2; it was therefore also done in addition to ANOVA as part of the sensitivity analysis (see 5.10.9 below) for any measure where experimental and control groups were significantly different at T1.

5.10.5 Significance Testing T1-T3

Control group data were only collected up to T2, so any tests involving T3 data, were only in relation to the experimental group. Analysis was therefore more straightforward and related tests were used, either parametric or non-parametric as appropriate according to the data.

5.10.6 Significance Testing T2-T3

Tests in relation to T2-T3 were to consider if any changes from T1-T2 were maintained; this analysis was therefore only done where there had been a significant change from T1-T2. It was not felt necessary to use ANOVA here as this analysis was not about the three time points (T1, T2, and T3) differing; rather, it was to consider if there had been significant changes from T2-T3.

5.10.7 Significance Level

Due to repeat comparisons on the same measures, Bonferroni adjustment was made to the significance level where appropriate, in order to avoid a possible Type 1 error. This was calculated by using $p=0.05$ and dividing this by the number of comparisons made. For correlations, due to the number of comparisons being done, $p=0.01$ was used for all correlations. Two-tailed tests were used throughout.

5.10.8 Effect Size

Effect sizes (ES) for each measure were calculated and these are noted as small, medium or large at the foot of each table; where effect size is negligible it is not included. The methods of calculating effect sizes are shown in Table 5.8 and the categories of effect size are shown in Table 5.9 (these are taken from Kinnear & Gray, 2010; specific page numbers for each test are shown in the table).

Table 5.8: Methods of Calculating Effect Size in Current Study

Test	Effect Size Formula	Method of Calculation
T-test	-	www.easycalculation.com/statistics/effect-size-t-test.php
Mann-Whitney U	$r = \frac{z}{\sqrt{n}}$	Manually
Wilcoxon	$r = \frac{z}{\sqrt{n}}$	Manually
ANOVA	-	By SPSS

Table 5.9: Effect Sizes for Tests Used in Current Study

Test	t-test (p.184)	Mann-Whitney U (p.195)	Wilcoxon (p.198)	ANOVA (p.221)
Small	0.2-0.5	0.1-0.3	0.1-0.3	0.01-0.06
Medium	0.5-0.8	0.3-0.5	0.3-0.5	0.06-0.14
Large	0.8→	0.5→	0.5→	0.14→

5.10.9 Sensitivity Analysis

The primary analysis was intention to treat, that is, based on the full data set, with all participants included whether they completed the training or not, and including data imputed as described earlier. For the primary analysis, outliers and extreme scores were removed where parametric tests were carried out, and left in where non-parametric tests were used.

Sensitivity analysis (Thabane et al, 2013) was carried out in addition to the primary analysis in order to explore any difference in results based on different methods of analysis. The sensitivity analysis considered other options for analysis:

- **Intention to treat (ITT, outliers in/out):** imputed data was included, with outliers included or removed, dependent on if they were included or removed in the primary analysis
- **Intention to treat, multiple testing:** separate tests were undertaken for significant difference from T1-T2 for both the experimental and control groups individually; this was only done where the primary analysis was based on change in scores (see section 5.10.4)
- **Per protocol, outliers removed (PP outliers out):** any dropouts were excluded from the data, no imputed data were included, outliers were removed
- **Per protocol, outliers included (PP outliers in):** any dropouts were excluded from the data, no imputed data were included, outliers were included
- **Per protocol, multiple testing:** as ITT above, only on a PP basis.

It is important to maintain the same criteria for sensitivity analysis as for the main analysis; therefore the same significance level was used, including Bonferroni adjustment, as described above.

5.10.10 Individual Change

In order to look at individual change and not just change within the group as a whole, the change in ABC total score was calculated for each participant in both the experimental and control groups at to look at difference from T1-T2, and from T1-T3. This was only carried out on the ABC total score as that was the primary measure.

5.10.11 Reliable Change

It was important to know how much change had occurred in levels of challenging behaviour following the training. When distributions of scores on a specific measure overlap in the control and experimental groups, it is possible for the difference in scores to not be statistically reliable, despite showing a significant change between time points. A 'reliable change' approach helps address this issue and tells us whether the change in scores between time points is more than just fluctuation within a specific measure/instrument. A reliable change score of greater than

1.96 would be unlikely to occur without actual change (Remington et al, 2007), therefore a score of more than 1.96 demonstrates that the change between time points is likely to be a real change. Using this approach provides a quantifiable assessment for each participant, rather than just for the group as a whole; it demonstrates statistical reliability of individual change.

The formula used for reliable change (RC) (Jacobson & Truax, 1991) is as follows:

$$RC = x_2 - x_1 \div S_{diff}$$

$$(S_{diff} = \sqrt{2(SE)^2} \quad \text{and} \quad SE = s_1\sqrt{1 - r_{xx}})$$

As this formula was developed for a measure where improvement would be indicated by increased scores at T2, for the current study the formula was changed to reflect the fact that the ABC scores reduce if showing improvement, as follows:

$$RC = x_1 - x_2 \div S_{diff}$$

The formula is explained further in Table 5.10.

Table 5.10: Reliable Change Formula Definitions

Symbol	Definition
s_1	Standard deviation of control and experimental groups at T1
r_{xx}	Test-retest reliability of ABC
x_1	T1 score of participant
x_2	T2 score of participant

5.11 Results from Reliability Checks

5.11.1 Inter-observer Reliability for Observational Data

The results from the inter-observer reliability checks for 20% (n=32) of the videos are shown in Table 5.11. Percentage occurrence and non-occurrence, and kappa values are shown for each behaviour except restraint. Only non-occurrence can be calculated for restraint as it did not occur in any of the re-coded videos; since it occurred at a very low level (around 0.5% across all data collection) this is not entirely surprising. Suen & Ary (1989) suggest that a kappa value of 0.6 or higher is acceptable for observational research and values for all behaviours met this. The same observer also re-scored the ASM while watching the video and correlation was carried out on the two sets of scores; this found a significant positive correlation ($r=0.969$; $p=0.007$; $n=32$).

Table 5.11: Inter-observer Reliability Data for MTS Behaviours

Behaviours	% Occurrence	% Non-occurrence	Kappa
Assistance	87.06	85.87	0.868
Other	64.36	96.38	0.710
Restraint	-	100	-
Processing	85.71	97.26	0.909
No Contact	89.25	90.71	0.884
Social Engagement	89.25	97.62	0.903
Non-social Engagement	78.44	92.64	0.738
Disengagement	88.79	77.64	0.773
Challenging Behaviour	92.5	99.25	0.751

5.11.2 Re-Coding of Videos

The re-coding of 5% of the videos by the writer focused on assistance and engagement as these were the behaviours scoring higher than expected. There were 161 videos taken in total including all three time points (this is less than the possible 194 videos as some people preferred to be observed rather than videoed, some videos were too short to use, and some service users left the organisation). Eight videos were therefore re-coded (5%) and the percentage scores for assistance and engagement for the re-coded videos are in Table 5.12. Tests found a significant positive correlation between original coding and re-coding scores for assistance ($r=0.996$; $p<0.001$; $n=8$) and engagement ($r=0.998$; $p<0.001$; $n=8$).

Table 5.12: Percentage Engagement and Assessment in Re-coding of Videos

Video	Assistance Main Coder	Assistance Writer	Engagement Main Coder	Engagement Writer
1	35	33	95	96
2	64	62	71	73
3	24	23	72	73
4	23	20	66	64
5	11	11	95	96
6	48	45	59	60
7	29	31	99	97
8	6	5	21	19

5.11.3 Length of Observations

The average length of observations was calculated for each data collection time point both in minutes and in number of observations, and the results are shown in Table 5.13. Total number of service users with short observation times are also shown, along with reasons for observations not lasting the full 120 minutes. Analysis was done to test for significant difference between experimental and control group in terms of number of observations and no difference was found at T1 (U=428; p=0.133; r=0.18, a small effect size) or at T2 (U=395; p=0.058; r=0.22, a small effect size).

Table 5.13: Service Users' Number of Observations at T1 and T2 and Reasons for Short Observations

Time	Experimental			Control		
	No of Observations Mean (range)	No of Minutes Mean (range)	Reasons for Short Observations	No of Observations Mean (range)	No of Minutes Mean (range)	Reasons for Short Observations
T1	275 (90-360) <i>n</i> =49	91.66 (30-120)	(<i>n</i> =14) RTL = 10 SU = 0 SO = 1 SB = 3	245 (93-360) <i>n</i> =22	81.66 (31-120)	(<i>n</i> =5) RTL = 5 SU = 0 SO = 0 SB = 0
T2	264 (92-360) <i>n</i> =44	88 (30.66-120)	(<i>n</i> =10) RTL = 9 SU = 1 SO = 0 SB = 0	224 (92-360) <i>n</i> =22	74.66 (30.66-120)	(<i>n</i> =6) RTL = 4 SU = 0 SO = 1 SB = 1
T3	273 (92-360) <i>n</i> =39	91 (30.66-120)	(<i>n</i> = 9) RTL = 9 SU = 0 SO = 0 SB = 0			

RTL= requested to leave by service user; SU=service user unwell; SO=service user going out; SB=service user went to bed

5.11.4 Reliability for Scoring of Measures and Data Entry

A random sample of 5% of each measure (*n*=10) were independently re-scored by a second person and also checked for correct entry into SPSS. The observational data was only checked for data entry, as scoring had already been checked as described above. Errors found in the sample were corrected. Agreement is shown in Table 5.14.

Table 5.14: Data Entry Agreement in Current Study

Measure	Scoring % Agreement	Data Entry % Agreement
ABC	100	90
ASM	-	100
ABS	100	90
Behaviour Recording Forms	100	100
CHABA	100	90
GCPLA	100	100
MTS	-	100
PBS Knowledge Tests	100	100
PSR	100	100
Practice Leadership	100	90

5.12 Results from Missing Data

There were a number of service users and managers who dropped out from the study; all of the dropouts by T2 were in the experimental group (by T3 the control group were no longer being followed up, so there is no data available from then). Two service users died in the course of the study, and four left the organisation; three of these due to breakdown of their service following challenging behaviours, and one for unrelated reasons. Three managers left the training and three left the service before T2; a further 16 managers left the service after the training was complete, but before T3 data collection. The result of this dropout is that by T3, only 22 of the managers in the experimental group had completed the training and were still in the same service; therefore only 44% of the group were per protocol by T3. Summary of dropouts in each cohort, with reasons and information about imputation is shown in Table 5.15.

5.13 Theory of Process of Change

The theory of process of change presented in chapter 4 can be considered in terms of the measures used in this study. Figure 4 shows these measures in relation to each of the three groups: managers, staff and service users. Each of these groups had specific measures to be considered in relation to them, although some of these measures were used for both staff and managers, for example, the CHABA; MTS was also used for both staff and service users, but was measuring different behaviours for both groups, so different codes were used. Figure 4 indicates the expected direction of change for each measure, that is, whether a reduction or increase was expected. It also shows the ongoing expected link between changes in the managers' measures, leading to changes in the staff measures, leading in turn to changes in the service users' measures. In all of these measures it was anticipated that the experimental group would show significant changes in the expected directions from pre to post training, and that the control group would show no significant change. As this study carried out follow-up data collection six months after the post training data collection, it was expected that some of the impact of training would have reduced and that there may be a return close to baseline; it was not expected that measures would continue to improve. Figure 4 is therefore a summary of the expected results, presented here before the actual results which follow in the next chapter.

Table 5.15: Summary of Dropouts, Data Imputed and Type of Imputation

Dropout	Cohort One	Cohort Two	Last Data Collection	Data Imputed	Type of Imputed Data
Service user died	2	-	T2	All data at T3	GA
Service user left organisation (for unrelated reason)	-	1	T1	All data at T2 and T3	GA
Service user left organisation (due to behaviour)	2	1	T1	All data at T2 and T3	IAWO
Service user refused to be observed	3	2	Varied	MTS and ASM (for when refusal took place)	GA
Manager left training	-	3	T1	None	Data for manager, staff and service user data collected, but not PP
Manager left service (before training complete)	-	3	T1	Manager's data at T2 and T3	GA for manager; staff and service user data collected, but not PP
Manager left service (after training complete)	13	3	T2	Manager's data at T3	GA for manager; staff and service user data collected, but not PP

5.14 Chapter Summary

This chapter has outlined the method used in the study presented in this thesis. The study is a non-randomised control group design with both between group and within group comparisons. Data were collected for three groups of participants: managers, staff and service users. The range of measures used and the process for data collection and reliability checks were described. The approach to data analysis and to missing data was outlined; an intention to treat approach was taken in the study and data for dropouts was imputed.

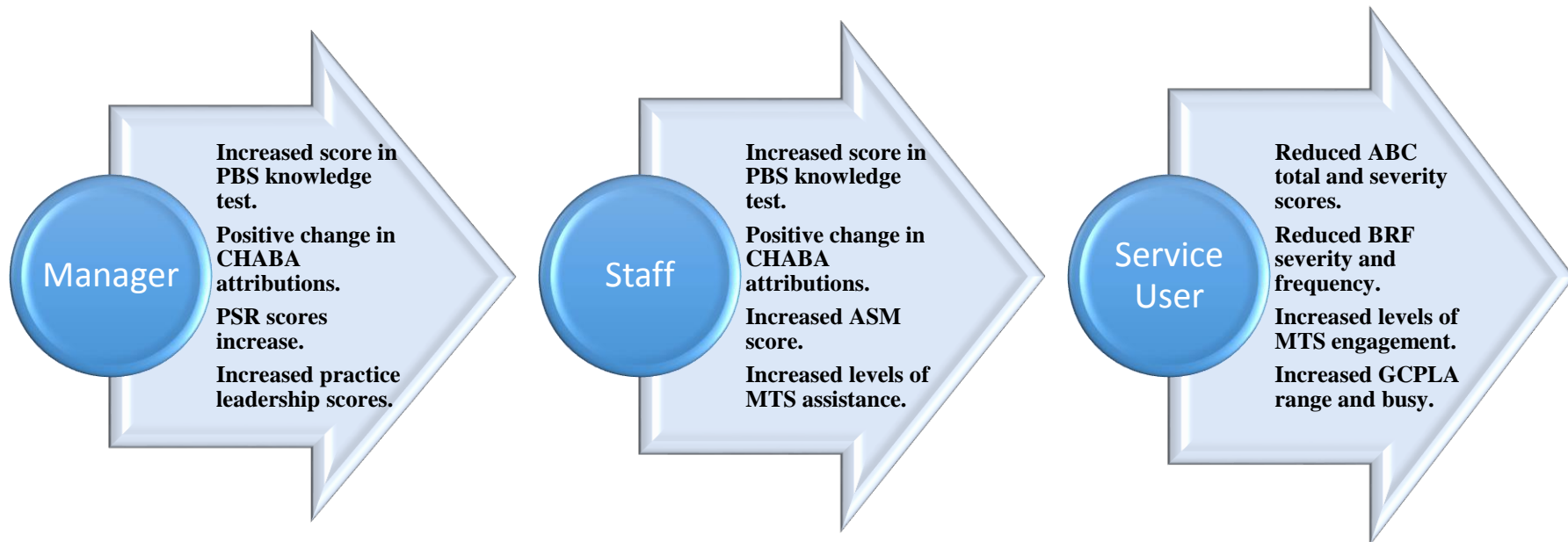


Figure 4. Theory of Process of Change in Relation to Measures Used in this Study

6 Managers' Results

6.1 Chapter Outline

This chapter details the results in relation to the managers' measures: the practice leadership questionnaire, the Positive Behaviour Support (PBS) knowledge test, the Challenging Behaviour Attribution Scale (CHABA) (Hastings, 1997), and the Periodic Service Review (PSR) (La Vigna et al, 1994). The results at baseline are considered first in order to identify any differences between the groups and between the cohorts at T1. The results following training are then presented with analysis of changes in each measure at T2; all tests of difference are followed up with sensitivity analyses in order to check the robustness of the results. This considers the impact of using an intention to treat (ITT) or a per protocol (PP) approach to data analysis, and also the impact of outliers on results. Finally there is a consideration of the results at T3, follow-up to training, to consider whether any changes following training have been maintained. The descriptive statistics for the sensitivity analyses are also presented in this section of the chapter.

6.2 Managers' Results at T1

6.2.1 Testing Differences between Cohort One and Cohort Two

This study involved two separate cohorts over two years; in both cohorts there was an experimental group and a control group. Both cohorts had 25 participants each in the experimental group; in the control group there were 12 participants in cohort one and 10 participants in cohort two. As there were two cohorts, analysis was carried out to establish if the two cohorts were significantly different, or if they could be combined and regarded as one group. Tests were carried out on each of the managers' measures to look at differences between cohorts one and two, both for the experimental group and also for the control group (the PSR was not included as it is only measured at T2 and T3). As can be seen in Table 6.1, none of the managers' measures showed any significant differences between cohort one and two, in either the experimental or control groups. The cohorts are therefore combined: both cohorts of the experimental are regarded as one group for all further analysis, and both cohorts of the control group are regarded as one group for all further analysis.

Table 6.1: Difference between Cohort One and Cohort Two at T1 for Managers' Measures for Experimental and Control Groups

Experimental Group						
	Cohort 1 Mean (SD)	Cohort 2 Mean (SD)	t	p	df	ES
Practice Leadership	62.27 (16.69)	64.56 (13.12)	0.539	0.592	48	0.16
PBS Knowledge	71.40 (13.11)	70.20 (14.03)	0.312	0.756	48	0.09
CHABA Learned Positive	0.93 (0.80)	1.09 (0.64)	0.776	0.441	48	0.22*
CHABA Learned Negative	0.90 (0.72)	0.92 (0.68)	0.067	0.947	48	0.02
CHABA Biomedical	0.40 (0.62)	0.51 (0.52)	0.701	0.487	48	0.20*
CHABA Stimulation	0.57 (0.60)	0.62 (0.73)	0.278	0.782	48	0.08
CHABA Emotional	0.92 (0.52)	1.04 (0.55)	0.781	0.439	48	0.23*
CHABA Physical Environment	0.52 (0.66)	0.41 (0.65)	0.226	0.822	47	0.07
Control Group						
Practice Leadership	58.65 (17.93)	57.77 (14.62)	U=48	0.456	n=22	0.17*
PBS Knowledge	61.67 (16.83)	57.50 (14.77)	0.611	0.548	20	0.27*
CHABA Learned Positive	1.08 (0.53)	0.90 (0.95)	0.564	0.579	20	0.25*
CHABA Learned Negative	0.83 (0.73)	0.83 (0.74)	0.003	0.998	20	0.00
CHABA Biomedical	0.38 (0.43)	0.28 (0.56)	U=56.5	0.821	n=22	0.05
CHABA Stimulation	0.39 (0.69)	0.35 (0.83)	0.487	0.632	19	0.22*
CHABA Emotional	1.01 (0.42)	1.14 (0.30)	0.774	0.448	20	0.35*
CHABA Physical Environment	0.54 (0.44)	0.45 (0.71)	0.392	0.699	20	0.18

* Small effect size

6.2.2 Testing Differences between Experimental and Control at T1

Tests were undertaken to establish if there were any significant differences between control and experimental groups at T1. This demonstrated the experimental group had significantly higher

scores in PBS knowledge; there was no other significant differences. This information is summarised in Table 6.2; the descriptive statistics for each measure are found in the individual sections in 6.3 below.

Table 6.2: Difference between Experimental and Control at T1 for Managers' Measures

Measures	t	p	df	ES
Practice Leadership	0.802	0.425	69	0.19
PBS Test	3.043	0.003†	70	0.73**
CHABA				
Learned Positive	0.550	0.584	69	0.13
Learned Negative	0.448	0.656	70	0.11
Biomedical	0.845	0.401	70	0.20*
Stimulation	1.291	0.201	70	0.31*
Emotional	0.719	0.474	70	0.17
Physical	0.225	0.823	70	0.05
Environment				

* Small effect size; ** Medium effect size; † Significant at $p \leq 0.05$

6.3 Managers' Results at T2: Impact of Training

6.3.1 Introduction

This section of the managers' results chapter considers the initial impact of training, that is, the difference from T1-T2. The various tests carried out are to establish if there are differences between the experimental and control groups from pre to post training. In the main this will be explored via ANOVA to consider any significant interaction between time and group; the other results from the ANOVA are presented for additional information in appendix four. For non-parametric data, the change in scores from T1-T2 will be calculated and then tests done on these change scores in order to examine any significant difference between experimental and control groups.

6.3.2 Practice Leadership

6.3.2.1 Descriptive Statistics

Table 6.3 shows the descriptive statistics for experimental and control groups at T1 and T2 for practice leadership. Figure 5 shows the graph of the mean percentage practice leadership score.

Table 6.3: Practice Leadership Descriptive Statistics T1 and T2

Practice Leadership	Experimental		Control	
	T1	T2	T1	T2
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Percentage Score	63.41 (14.90)	69.94 (17.88)	58.25 (16.13)	61.78 (15.89)



Figure 5. Mean Percentage Practice Leadership T1 and T2

6.3.2.2 Significance Testing

A mixed multi-factorial ANOVA was carried out on the percentage practice leadership score from T1-T2; this demonstrated there was no significant interaction between time and group ($F=1.030$; $p=0.314$; $df=1$; partial eta squared=0.015, a small effect size). As can be seen from appendix four, Table 13.1, neither time nor group was significant either.

6.3.2.3 Sensitivity Analysis

Three additional analyses were carried out in addition to the main analysis; these show that the results did not change dependent on method of analysis; regardless of method of analysis, there is no significant interaction between time and group for practice leadership. The results are described in Table 6.4.

Table 6.4: Practice Leadership Sensitivity Analysis T1-T2

Practice Leadership	F	p	df	ES
ITT, ANOVA, outliers in	0.330	0.568	1	0.005
PP, ANOVA outliers out	3.017	0.088	1	0.049*
PP, ANOVA outliers in	1.453	0.233	1	0.024*

* Small effect size

6.3.2.4 Practice Leadership Individual Areas Analysis

As there was no significant change in overall practice leadership scores from T1-T2, it was not felt necessary to do any further significance testing for individual elements of the measure. However, in order to later compare changes within the experimental group with relevant literature, data were analysed from the three main areas within the practice leadership questionnaire: team meetings, supervision, and observation by the manager, which had a number of different subsidiary areas of consideration. The means of each of these at T1 and T2, together with percentage change from T1-T2, are shown in Table 6.5.

Table 6.5: Practice Leadership Questionnaire Individual Areas Experimental Group T1-T2

n=50	T1	T2	% Change
Practice Leadership Index (range)	63 (31-89)	70 (25-94)	11
Team meetings			
Main focus is engagement	30	24	-20
Supervision			
Constructive feedback	60	68	13
Main focus is engagement	22	36	64
Observation by Manager			
Observation at least monthly	82	82	0
Main focus is engagement	26	48	85
Feedback usually	72	86	19
Models good support usually	40	64	60
Correction to support better usually	56	62	11
Can approach manager for support usually	78	80	3
Advice is at least quite helpful	64	72	13

6.3.3 PBS Knowledge

6.3.3.1 Descriptive Statistics

Table 6.6 shows the mean and standard deviation for PBS knowledge for the experimental and control groups at T1 and T2. Figure 6 shows the graph of the means.

Table 6.6: Managers' PBS Knowledge Descriptive Statistics T1 and T2

PBS Knowledge	Experimental		Control	
	T1	T2	T1	T2
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Percentage Score	71 (13.45)	83 (14.53)	60 (15.70)	67 (17.56)

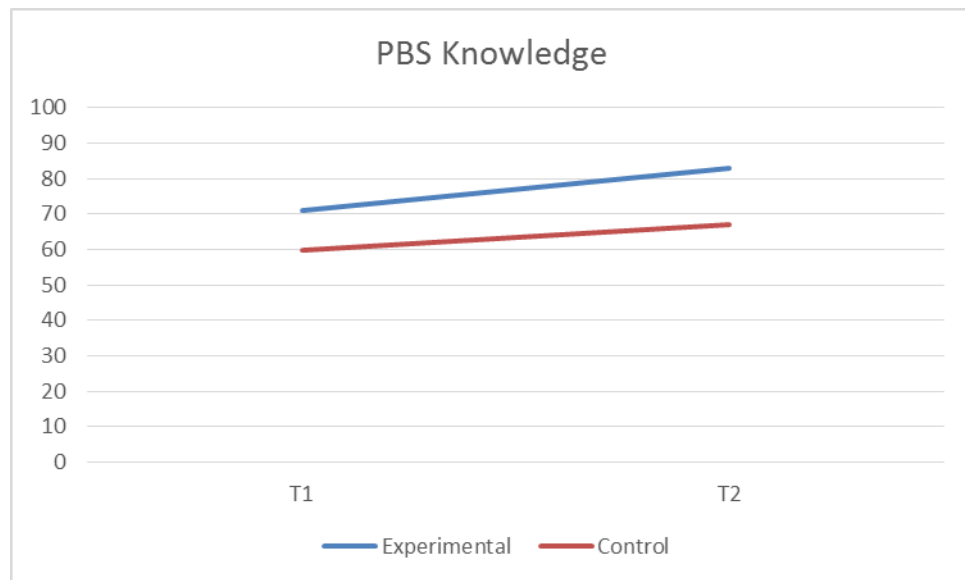


Figure 6. Managers' Mean Percentage PBS Knowledge T1 and T2

6.3.3.2 Significance Testing

A mixed multi-factorial ANOVA was carried out and showed that there was a significant interaction between time and group for PBS knowledge ($F=4.767$; $p=0.033$; $df=1$; partial eta squared=0.067, a medium effect size). As can be seen from appendix four, Table 13.2, both time and group were also significant.

6.3.3.3 Sensitivity Analysis

Three additional analyses were carried out in addition to the main analysis for PBS knowledge. As data did not meet parametric assumptions (with outliers in), then the change in scores from T1-T2 were calculated and the groups were compared on this basis. This approach was also useful as, to some extent, it helped to address the fact that the experimental group scored significantly higher than the control group at T1 for PBS knowledge. These analyses showed that the results changed dependent on method of analysis; when the change in scores was tested with outliers in, then the difference between experimental and control was not significant, using an ITT approach. However, under a PP approach, the difference was significant, whether outliers were included or not. The results are summarised in Table 6.7.

Table 6.7: Managers' PBS Knowledge Sensitivity Analysis T1-T2

PBS Knowledge	U	p	n	ES
ITT, outliers in	391.5	0.051	72	0.23*
PP, outliers out	244.5	0.005†	61	0.36**
PP, outliers in	248	0.004†	62	0.36**

* Small effect size; ** Medium effect size; † Significant at $p \leq 0.05$

6.3.4 Challenging Behaviour Attribution Scale (CHABA)

6.3.4.1 Descriptive Statistics

Table 6.8 shows the mean and standard deviation for the CHABA subscales for the experimental and control groups at T1 and T2.

Table 6.8: Managers' CHABA Descriptive Statistics T1 and T2

	Experimental		Control	
	T1	T2	T1	T2
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Learned Positive	1.01 (0.72)	0.94 (0.83)	0.10 (0.74)	0.94 (0.68)
Learned Negative	0.91 (0.69)	0.83 (0.75)	0.83 (0.71)	0.87 (0.74)
Biomedical	0.45 (0.57)	0.43 (0.72)	0.33 (0.48)	0.41 (0.53)
Stimulation	0.60 (0.66)	0.64 (0.79)	0.37 (0.73)	0.52 (0.73)
Emotional	0.98 (0.53)	0.88 (0.64)	1.07 (0.37)	0.99 (0.68)
Physical Environment	0.46 (0.65)	0.53 (0.69)	0.50 (0.57)	0.51 (0.62)

6.3.4.2 Significance Testing

In order to test difference in the CHABA subscales mixed multi-factorial ANOVA were used. This found that there was no significant interaction between time and group for any of the subscales, and effect sizes were negligible for all but BM, which was small. The results are summarised in Table 6.9. As can be seen from appendix four, Table 13.3 – Table 13.8, neither time nor group was significant either, for any of the subscales.

Table 6.9: Managers' CHABA Significance Testing T1-T2

	F	p	df	ES
Learned Positive	0.624	0.432	1	0.009
Learned Negative	0.301	0.585	1	0.004
Biomedical	1.626	0.207	1	0.024*
Stimulation	0.497	0.483	1	0.007
Emotional	0.600	0.441	1	0.009
Physical Environment	0.174	0.678	1	0.003

*Small effect size

6.3.4.3 Sensitivity Analysis

Three additional analyses were carried out in addition to the main analysis for each of the subscales of the CHABA, except for LN which had no outliers, so only one additional test was done; these show that the results did not change dependent on method of analysis. This is summarised in Table 6.10.

Table 6.10: Managers' CHABA Sensitivity Analysis T1-T2

Learned Positive	F	p	df	ES
ITT, outliers in	0.003	0.959	1	0.000
PP, outliers out	0.442	0.509	1	0.008
PP, outliers in	0.138	0.712	1	0.002
Learned Negative	F	p	df	ES
PP (no outliers)	0.007	0.931	1	0.000
Biomedical	F	p	df	ES
ITT, outliers in	0.249	0.619	1	0.004
PP, outliers out	0.664	0.419	1	0.012*
PP, outliers in	0.001	0.976	1	0.000
Stimulation	F	p	df	ES
ITT, outliers in	0.251	0.618	1	0.004
PP, outliers out	0.190	0.664	1	0.003
PP, outliers in	0.053	0.819	1	0.001
Emotional	F	p	df	ES
ITT, outliers in	0.019	0.891	1	0.000
PP, outliers out	0.288	0.594	1	0.005
PP, outliers in	0.005	0.944	1	0.000
Physical Environment	F	p	df	ES
ITT, outliers in	0.060	0.808	1	0.001
PP, outliers out	0.071	0.791	1	0.001
PP, outliers in	0.168	0.684	1	0.003

* Small effect size

6.3.5 Periodic Service Review (PSR)

6.3.5.1 Descriptive Statistics

The mean percentage score for the PSR at T2 for the experimental group was 66 (SD: 21.22).

6.3.5.2 Significance Testing

As the PSR was not measured at T1, and was only measured for the experimental group, then no significance testing could be done.

6.3.6 Correlations

Correlations were carried out for the managers' measures, and showed that for the experimental group at T2 there was a significant positive correlation between PSR and PBS knowledge; between PSR and practice leadership; and between knowledge and practice leadership. Knowledge and practice leadership were negatively correlated for the control group but this was not significant. The results are summarised in Table 6.11.

Table 6.11: Correlations for Managers' Measures at T2

Correlations	Experimental		Control	
	r	p	r	p
PSR/Knowledge	Rho=0.393	0.005†	-	-
PSR/Practice Leadership	0.555	p<0.001	-	-
Knowledge/Practice Leadership	0.438	0.001†	-0.351	0.110

† Significant at $p \leq 0.01$

6.3.7 Summary of Managers' Results at T2

6.3.7.1 Practice Leadership

Results from the practice leadership questionnaire show that although the practice leadership scores increased slightly for both groups, there was no significant change. This result remains the same regardless of how analysis is carried out. It is therefore possible to confidently assert that there was no significant change in practice leadership as assessed via the practice leadership questionnaire following PBS training.

6.3.7.2 Managers' Knowledge and Attributions

There was a significant increase in managers' PBS knowledge following training, this was found in most, but not all, of the alternative ways of analysing; it can therefore be tentatively concluded that there was a significant change in managers' knowledge following PBS training.

In the CHABA, four subscales decreased following the training for the experimental group; Stimulation and Physical Environment increased very slightly. For the control group all the subscales increased except for Emotional. However, none of these changes was significant for any subscale, even when analysed in a range of different ways. It can therefore be confidently concluded that there were no significant changes to managers' attributions about challenging behaviour following the PBS training.

6.4 Managers' Results at T3: Maintenance of Training Effects

6.4.1 Introduction

This section of the managers' results chapter considers the data gathered at follow-up, T3. The question being considered in this section is different from the previous results' section which considered initial impact from the training, that is, comparisons from T1-T2. This current section considers whether training effects were maintained and therefore analysis is done for T2-T3 to consider whether any changes occurring from T1-T2 were maintained; this analysis was only done for measures where there had been significant change from T1-T2 (**PBS knowledge**).

Analysis was also done from T1-T3; this was to consider the training's long-term effectiveness, and was done for all measures. So the questions being explored with the analyses are different depending on the time frame; from T1-T3 a significant difference (in the expected direction) would show the training's long-term effectiveness; a non-significant difference from T2-T3 would show that any changes following training had not been lost. Throughout this section, the PP analysis is based on data from only 22 managers, due to either managers leaving the course or the service, or service users leaving the organisation, and therefore the results from the PP analysis cannot be regarded as robust, as they represent less than half of the experimental group. In addition, it is worth noting that the within the ITT analysis, data for 25 managers were imputed, therefore to some extent these results also need to be treated with caution.

6.4.2 Practice Leadership

6.4.2.1 Descriptive Statistics

Table 6.12 shows the descriptive statistics for practice leadership for the experimental group at T1, T2 and T3, for the ITT dataset and also for the PP dataset, with and without outliers. Figure 7 shows the graphs of the mean percentage practice leadership score for the ITT dataset.

Table 6.12: Practice Leadership Descriptive Statistics T1-T3

PL (<i>n</i>)	T1	T2	T3
	Mean (SD)	Mean (SD)	Mean (SD)
ITT, outliers in (50)	63.41 (14.90)	69.94 (17.88)	65.39 (18.05)
ITT, outliers out (46)	63.86 (15.01)	72.40 (14.02)	68.91 (14.04)
PP (no outliers) (22)	62.60 (14.38)	76.20 (12.56)	74.34 (9.16)

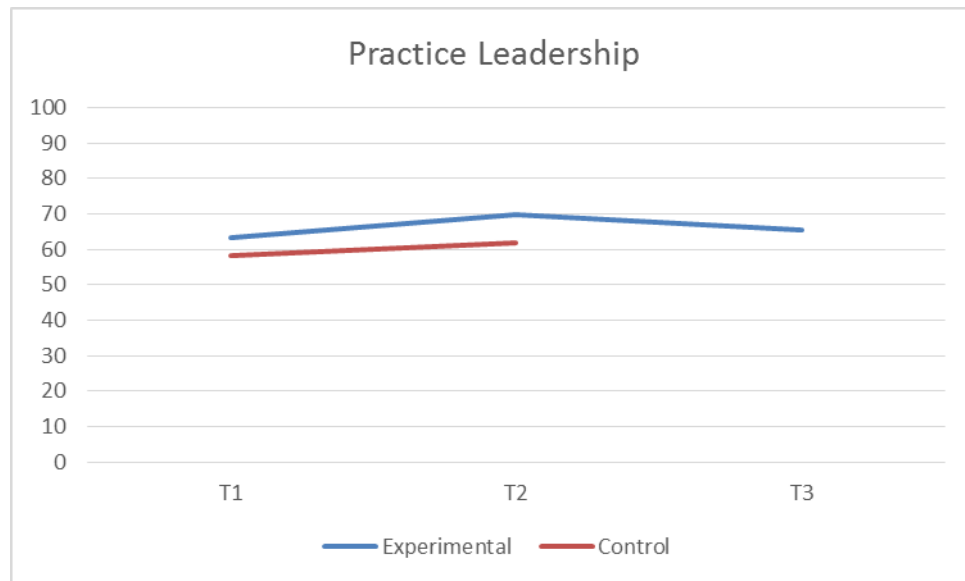


Figure 7. Mean Percentage Practice Leadership T1-T3

6.4.2.2 Significance Testing

A related t-test was carried out on the percentage practice leadership score from T1-T3; this demonstrated there was no significant difference between T1-T3 ($t=1.884$; $p=0.066$; $df=45$; $d=0.56$, a medium effect size).

In order to explore potential links between turnover of management and practice leadership, tests were carried out to check for any difference in practice leadership between the group whose manager had changed by T3, and the group whose manager remained the same. This showed that there was a significant difference between the groups ($U=175$; $p=0.009$; $n=50$; $r=0.37$, a medium effect size), with a mean practice leadership score for the unchanged group ($n=22$) of 72% and a mean practice leadership score of 60% for the group whose manager had changed ($n=28$).

6.4.2.3 Sensitivity Analysis

Two additional analyses were carried out in addition to the main analysis for practice leadership for T1-T3. These show that the results for T1-T3 change dependent on method of analysis; with the PP analysis there is significant difference from T1-T3 and a large effect size. However the PP analysis is based on data from only 22 managers, due to either managers leaving the course or the service, or service users leaving the organisation, and therefore this analysis cannot be regarded as robust; therefore it appears safe to conclude that there was no significant difference in practice leadership scores from pre training to follow-up. The results are described in Table 6.13, and descriptive statistics for the sensitivity analysis are above in Table 6.12.

Table 6.13: Practice Leadership Sensitivity Analysis T1-T3

PL T1-T3	t	p	df	ES
ITT, outliers in	0.677	0.501	49	0.19
PP (no outliers)	5.108	p<0.001	21	2.23***

*** Large effect size

6.4.3 PBS Knowledge

6.4.3.1 Descriptive Statistics

Table 6.14 shows the mean and standard deviation for PBS knowledge for the experimental group at T1, T2 and T3, for both ITT and PP datasets, with and without outliers. Figure 8 shows the graph of the mean percentage at the same time points for the ITT dataset.

Table 6.14: Managers' PBS Knowledge Descriptive Statistics T1-T3

PBS Knowledge (n)	T1	T2	T3
	Mean (SD)	Mean (SD)	Mean (SD)
ITT, outliers in (50)	71 (13.45)	83 (14.53)	84 (11.79)
ITT, outliers out (41)	70 (13.84)	85 (12.82)	86 (4.94)
PP (no outliers) (22)	69 (11.16)	88 (10.01)	88 (9.34)

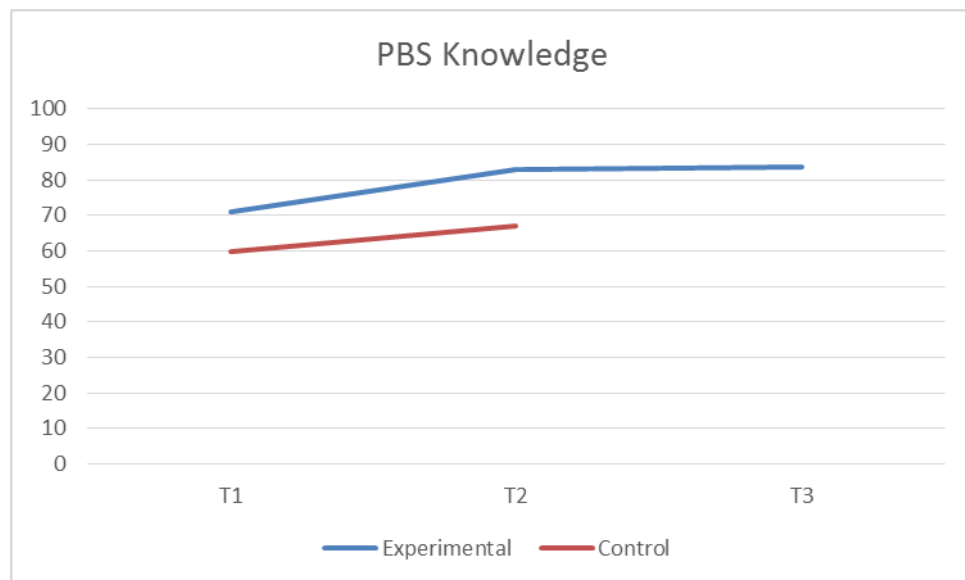


Figure 8. Managers' Mean Percentage PBS Knowledge T1-T3

6.4.3.2 Significance Testing

Tests carried out on PBS knowledge T1-T3 and T2-T3 showed that there was a significant difference from T1-T3 ($Z=4.330$; $p=0.00$; $n=50$; $r=0.612$, a large effect size) but not for T2-T3 ($Z=0.445$; $p=0.665$; $n=50$; $r=0.09$, a negligible effect size).

6.4.3.3 Sensitivity Analysis

Two additional analyses were carried out in addition to the main analysis for PBS knowledge at T1-T3 and T2-T3. These show that the results do not change dependent on method of analysis; whether PP or ITT analysis, T1-T3 shows a significant increase in PBS knowledge and a large effect size, and there is no significant difference between T2-T3. This demonstrates a robust conclusion as it does not change dependent on method of analysis. The results are summarised in Table 6.15 and descriptive statistics for the sensitivity analysis are above in Table 6.14.

Table 6.15: Managers' PBS Knowledge Sensitivity Analysis T1-T3 and T2-T3

PBS Knowledge T1-T3	Z/t	p	n	ES
ITT, outliers out	$Z=4.874$	$p<0.001$	41	0.76***
PP (no outliers)	$t=7.127$	$p<0.001$	21	3.11***
PBS Knowledge T2-T3	Z/t	p	n	ES
ITT, outliers out	$Z=0.900$	0.378	39	0.14*
PP (no outliers)	$t=0.179$	0.860	21	0.08

* Small effect size; *** Large effect size

6.4.4 Challenging Behaviour Attribution Scale (CHABA)

6.4.4.1 Descriptive Statistics

Table 6.16 shows the mean and standard deviation for the CHABA subscales for the experimental group at T1, T2 and T3 for both ITT and PP datasets, with and without outliers.

Table 6.16: Managers' CHABA Descriptive Statistics T1-T3

CHABA (<i>n</i>)	T1	T2	T3
	Mean (SD)	Mean (SD)	Mean (SD)
Learned Positive			
ITT, outliers in (50)	1.01 (0.72)	0.94 (0.83)	0.88 (1.06)
ITT, outliers out (48)	0.99 (0.72)	0.98 (0.72)	0.86 (0.83)
PP, outliers in (22)	0.98 (0.73)	1.12 (0.70)	0.83 (0.98)
PP, outliers out (20)	1.06 (0.61)	1.10 (0.72)	1.04 (0.69)
Learned Negative			
ITT, outliers in (50)	0.91 (0.69)	0.83 (0.75)	0.73 (0.78)
ITT, outliers out (49)	0.93 (0.69)	0.82 (0.76)	0.77 (0.75)
PP, outliers in (22)	0.76 (0.71)	0.91 (0.77)	0.70 (0.92)
PP, outliers out (21)	0.72 (0.70)	0.89 (0.78)	0.81 (0.78)
Biomedical			
ITT (no outliers) (50)	0.45 (0.57)	0.43 (0.72)	0.37 (0.78)
PP (no outliers) (22)	0.46 (0.51)	0.48 (0.71)	0.44 (0.87)
Stimulation			
ITT, outliers in (50)	0.60 (0.66)	0.64 (0.79)	0.43 (0.78)
ITT, outliers out (45)	0.61 (0.67)	0.61 (0.73)	0.40 (0.69)
PP (no outliers) (22)	0.51 (0.71)	0.77 (0.86)	0.40 (0.96)
Emotional			
ITT (no outliers) (50)	0.98 (0.53)	0.88 (0.64)	0.86 (0.70)
PP (no outliers) (22)	0.92 (0.54)	0.93 (0.67)	0.79 (0.85)
Physical			
Environment	0.46 (0.65)	0.53 (0.69)	0.42 (0.79)
ITT, outliers in (50)	0.46 (0.66)	0.51 (0.68)	0.46 (0.73)
ITT, outliers out (49)	0.38 (0.57)	0.52 (0.78)	0.41 (0.92)
PP (no outliers) (22)			

6.4.4.2 Significance Testing

In order to test difference in the CHABA subscales related t-tests were used. These found that there was no significant difference for T1-T3 for any of the subscales. Effect sizes at T1-T3 were small for all subscales except PE, which was negligible. The results are summarised in Table 6.17.

Table 6.17: Managers' CHABA Significance Testing T1-T3

T1-T3	t	p	df	ES
Learned Positive	1.064	0.293	47	0.31*
Learned Negative	1.168	0.249	48	0.38*
Biomedical	0.709	0.482	49	0.20*
Stimulation	1.618	0.113	44	0.49*
Emotional	1.174	0.246	49	0.34*
Physical Environment	0.006	0.995	48	0.00

*Small effect size

6.4.4.3 Sensitivity Analysis

Additional analyses were carried out in addition to the main analysis for each of the subscales of the CHABA for T1-T3 (some tests were not done as the main analysis had no outliers); these show that the results do not change dependent on whether ITT or PP analysis. Effect sizes were small for the subscales LP, LN and ST; the rest were negligible. This is summarised in Table 6.18 and descriptive statistics for the sensitivity analysis are above in Table 6.16.

Table 6.18: Managers' CHABA Sensitivity Analysis T1-T3

Learned Positive T1-T3	t	p	df	ES
ITT, outliers in	0.829	0.411	49	0.24*
PP, outliers out	0.112	0.912	19	0.05
PP, outliers in	0.744	0.465	21	0.32*
Learned Negative T1-T3	t	p	df	ES
ITT, outliers in	1.306	0.197	49	0.37*
PP, outliers out	0.491	0.629	20	0.22*
PP, outliers in	0.248	0.806	21	0.11
Biomedical T1-T3	t	p	df	ES
PP (no outliers)	0.098	0.923	21	0.04
Stimulation T1-T3	t	p	df	ES
ITT, outliers in	1.332	0.189	47	0.39*
PP (no outliers)	0.493	0.627	21	0.22*
Emotional T1-T3	t	p	df	ES
PP (no outliers)	0.730	0.473	21	0.32
Physical Environment T1-T3	t	p	df	ES
ITT, outliers in	0.357	0.722	49	0.10
PP (no outliers)	0.141	0.889	21	0.06

* Small effect size

6.4.5 Periodic Service Review

6.4.5.1 Descriptive Statistics

As the PSR was developed as part of the training, then there is no data for T1; comparisons here are therefore only between T2-T3. Table 6.19 shows the mean and standard deviation for the PSR for the experimental group at T2 and T3, for the ITT dataset and also for the PP dataset, with and without outliers.

Table 6.19: PSR Descriptive Statistics T2 and T3

PSR (<i>n</i>)	T2	T3
	Mean (SD)	Mean (SD)
ITT (50)	66 (21.22)	61 (20.50)
PP outliers out (21)	78 (15.00)	69 (19.23)
PP outliers in (22)	76 (16.42)	67 (21.12)

6.4.5.2 Significance Testing

The difference between T2-T3 for the PSR was tested and this found that there was no significant difference. Additional analyses were carried out under the PP approach and these were significant with a large effect size from T2-T3. However these results cannot be regarded as robust, as noted above. Results are summarised in Table 6.20.

Table 6.20: PSR Sensitivity Analysis T2-T3

PSR	<i>t</i>	<i>p</i>	<i>df</i>	ES
ITT (no outliers)	1.718	0.092	49	0.03
PP outliers out	2.119	0.047†	20	0.95***
PP outliers in	2.330	0.030†	21	1.02***

*** Large effect size; † Significant at $p \leq 0.05$

In order to explore potential links between turnover of management and PSR, tests were carried out to check for any difference between the group whose manager had changed by T3, and the group whose manager remained the same. This showed that there was no significant difference between the groups ($U=222$; $p=0.092$; $n=50$; $r=0.24$, a small effect size), with a mean PSR score for the stable management group ($n=22$) of 67% and a mean PSR score of 57% for the group whose manager had changed ($n=28$).

6.4.6 Correlations

Correlations were carried out for the managers' measures, and showed that there was a significant positive correlation between PSR and knowledge; between PSR and practice leadership; and between knowledge and practice leadership. The results are summarised in Table 6.21.

Table 6.21: Correlations for Managers' Measures at T3

Correlations	r	p
PSR/Knowledge	0.501	p<0.001
PSR3/Practice Leadership	0.521	p<0.001
Knowledge/ Practice Leadership	0.418	0.003†

† Significant at $p \leq 0.01$

6.5 Chapter Summary

6.5.1.1 Practice Leadership

Results from the practice leadership questionnaire show scores increased from T1-T2 for both groups, but there was no significant change. Scores decreased from T2-T3; however overall for the experimental group, there was a slight increase from T1-T3 but this was not a significant change. These results remain the same regardless of how analysis is carried out; it is therefore possible to confidently assert that there was no significant change in practice leadership as assessed via the practice leadership questionnaire following PBS training, either initially or at follow-up. There was a significant difference in practice leadership at T3 between the group whose manager had changed, and the stable management group.

6.5.1.2 Managers' Knowledge

In the experimental group, managers' PBS knowledge increased significantly from T1-T2; from T2-T3 there was a further slight increase, although this was not a significant change. This demonstrates that increased PBS knowledge had been maintained over time. The change from T1-T3 was also significant; this was demonstrated in all methods of analysis and the effect size was large; therefore there can be confidence in these results.

6.5.1.3 Managers' Attributions

For the CHABA, there were no significant differences in any subscale for either experimental or control group from T1-T2. For the experimental group, all the subscales decreased from pre training to follow-up. However, none of these changes were significant for any subscale, even when analysed in a range of different ways. It can therefore be confidently concluded that there

were no significant changes to managers' attributions about challenging behaviour following PBS training.

6.5.1.4 PSR

PSR was only measured for the experimental group, and only at T2 and T3; scores decreased from T2-T3, although this was not a significant change, except in the PP analysis which only includes 44% of the group, and therefore cannot be regarded as a robust conclusion. There was no significant difference in PSR score at T3 between the group whose manager had changed, and the stable management group.

6.5.1.5 Correlations

For the experimental group there was a significant positive correlation between PSR and PBS knowledge; between PSR and practice leadership; and between knowledge and practice leadership, both at T2 and at T3.

7 Staff Results

7.1 Chapter Outline

This chapter details the results in relation to the staff measures: Momentary Time Sampling (MTS), Active Support Measure (ASM) (Mansell et al, 2005a), Positive Behaviour Support (PBS) knowledge test, and the Challenging Behaviour Attribution Scale (CHABA) (Hastings, 1997). The results at baseline are considered first in order to identify any differences between the groups and between the cohorts at T1. The results following training are then presented with analysis of changes in each measure at T2; all tests of difference are followed up with sensitivity analyses in order to check the robustness of the results. This considers the impact of using an intention to treat (ITT) or a per protocol (PP) approach to data analysis, and also the impact of outliers on results. Where there were correlations carried in out in relation to the managers' measures, these are also presented in this chapter. Finally there is a consideration of the staff results at T3, to consider whether any changes following training have been maintained. The descriptive statistics for the sensitivity analyses are also presented in this section of the chapter.

7.2 Staff Results at T1

7.2.1 Testing Differences between Cohort One and Cohort Two

As previously noted, this study involved two separate cohorts over two years; in both cohorts there was an experimental group and a control group. Analysis was carried out to establish if the two cohorts were significantly different, or if they could be combined and regarded as one group. Tests were carried out on each of the staff measures to look at differences between cohorts one and two, both for the experimental group and also for the control group. As can be seen in Table 7.1, none of the staff measures showed any significant differences between cohorts one and two, in either the experimental or control groups. The cohorts are therefore combined: both cohorts of the experimental are regarded as one group for all further analysis, and both cohorts of the control group are regarded as one group for all further analysis.

Table 7.1: Difference between Cohort One and Cohort Two at T1 for Staff Measures for Experimental and Control Groups

Experimental Group						
Measures	Cohort 1 Mean (SD)	Cohort 2 Mean (SD)	t	p	df	ES
MTS Positive Contact	56.64 (29.79)	51.82 (25.98)	0.603	0.550	47	0.18
MTS Assistance	35.95 (28.48)	35.72 (21.25)	0.032	0.974	47	0.01
MTS Other Conversation	20.63 (21.58)	16.25 (16.70)	0.377	0.708	45	0.11
MTS Negative Contact	5.16 (7.86)	8.69 (18.30)	U=295	0.724	n=50	0.05
MTS Processing	3.90 (6.73)	6.98 (17.05)	U=295	0.727	n=50	0.05
MTS Restraint	1.26 (4.85)	1.71 (5.45)	U=265	0.244	n=50	0.16
MTS No Contact	38.23 (30.34)	39.96 (27.15)	0.233	0.817	47	0.07
CHABA Learned Positive	1.13 (0.72)	0.96 (0.61)	0.923	0.361	48	0.27*
CHABA Learned Negative	1.02 (0.54)	0.81 (0.61)	1.299	0.200	48	0.37*
CHABA Biomedical	0.38 (0.64)	0.40 (0.53)	0.266	0.791	47	0.08
CHABA Stimulation	0.39 (0.74)	0.25 (0.57)	U=267	0.381	n=50	0.12
CHABA Emotional	1.13 (0.47)	0.94, (0.63)	1.226	0.226	48	0.35
CHABA Physical Environment	0.26 (0.73)	0.28 (0.59)	0.271	0.788	47	0.08
ASM	56.09 (24.24)	61.42 (18.12)	U=271.5	0.568	n=50	0.08
PBS Knowledge	54.15 (13.86)	56.90 (8.75)	0.835	0.408	48	0.24*

Control Group						
Measures	Cohort 1 Mean (SD)	Cohort 2 Mean (SD)	t	p	df	ES
MTS Positive Contact	74.64 (34.00)	75.76 (22.36)	0.517	0.611	19	0.24*
MTS Assistance	42.87 (30.59)	64.17 (23.33)	1.805	0.086	20	0.81** *
MTS Other Conversation	31.77 (26.68)	11.40 (19.76)	1.998	0.059	20	0.89** *
MTS Negative Contact	25.36 (34.00)	24.20 (22.35)	U=51.5	0.582	n=22	0.18
MTS Processing	0.83 (2.07)	0.50 (1.58)	U=51	0.582	n=22	0.18
MTS Restraint ±	-	-	-	-	-	-
MTS No Contact	24.53 (33.96)	23.70 (22.58)	0.553	0.587	19	0.25*
CHABA Learned Positive	0.97 (0.69)	0.86 (0.72)	U=59	0.974	n=22	0.01
CHABA Learned Negative	0.91 (0.77)	0.99 (0.50)	0.287	0.777	20	0.13
CHABA Biomedical	0.50 (0.45)	0.38 (0.42)	1.091	0.289	19	0.50**
CHABA Stimulation	0.29 (0.52)	0.48 (0.54)	0.828	0.417	20	0.37*
CHABA Emotional	1.00 (0.46)	1.08 (0.38)	0.399	0.694	20	0.18
CHABA Physical Environment	0.34 (0.55)	0.61 (0.38)	1.316	0.203	20	0.59**
ASM	61.85 (21.33)	77.11 (11.57)	U=32	0.069	n=22	0.39**
PBS Knowledge	54.88 (8.14)	53.89 (14.32)	U=58.5	0.923	n=22	0.02

* Small effect size; **Medium effect size; *** Large effect size; ±Restraint = 0 for control group at T1 in both cohorts, so no comparisons done

7.2.2 Testing Differences between Control and Experimental at T1

Tests were undertaken to establish if there were any significant differences between control and experimental groups at T1. This demonstrated that in the experimental group MTS positive contact was significantly lower and MTS processing, restraint and negative contact were significantly higher than in the control group; also the ASM scored significantly higher in the control group. There were no significant differences for MTS assistance, other conversation, or no contact, for any of the subscales on the CHABA or for PBS knowledge. This information is

summarised in Table 7.2; the descriptive statistics for each measure are found in the individual sections in 6.3 below.

Table 7.2: Difference between Experimental and Control at T1 for Staff Measures

Staff Measures	t/U	p	df/n	ES
MTS				
Positive Contact	2.994	0.004†	70	0.72**
Assistance	2.553	0.013	70	0.61**
Other Conversation	U=544	0.941	n=72	0.01
Negative Contact	U=311.5	0.001†	n=72	0.37**
Processing	U=351	0.007†	n=72	0.32**
Restraint	U=396	0.006†	n=72	0.32**
No Contact	2.045	0.045	70	0.49*
CHABA Learned Positive	0.276	0.784	69	0.07
CHABA Learned	0.222	0.825	70	0.05
Negative	0.516	0.608	70	0.12
CHABA Biomedical	0.352	0.726	70	0.08
CHABA Stimulation	0.245	0.807	69	0.06
CHABA Emotional	0.352	0.726	70	0.08
CHABA Physical				
Environment				
ASM	U=388.5	0.048††	n=72	0.23*
PBS Knowledge	U=502	0.556	n=50	0.07

* Small effect size; **Medium effect size; † Significant at $p \leq 0.007$; †† Significant at $p \leq 0.05$

7.3 Staff Results at T2: Impact of Training

7.3.1 Introduction

This section of the staff results chapter considers the initial impact of training, that is, the difference from T1-T2. The various tests carried out are to establish if there are differences between the experimental and control groups from pre to post training. In the main this will be explored via ANOVA to consider any significant interaction between time and group; the other results from the ANOVA are presented for additional information in appendix five. For non-parametric data, the change in scores from T1-T2 will be calculated and then tests done on these change scores in order to examine any significant difference between experimental and control groups.

7.3.2 Momentary Time Sampling

7.3.2.1 Descriptive Statistics

Table 7.3 shows the mean and standard deviation for experimental and control groups at T1 and T2 for the range of MTS subscales. Assistance, positive contact, and no contact were chosen as behaviours of particular focus, as described in the method chapter. Figures 9-11 show the graphs of the means for these at T1 and T2 for the experimental and control groups.

Table 7.3: Staff MTS Descriptive Statistics T1 and T2

MTS	Experimental		Control	
	T1	T2	T1	T2
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Positive Contact	53.97 (27.21)	56.93 (26.60)	75.15 (28.64)	64.83 (33.00)
Assistance	35.53 (24.69)	35.83 (24.59)	52.55 (29.00)	43.46 (26.94)
Other Conversation	18.44 (19.22)	21.10 (21.99)	22.51 (25.48)	21.36 (25.09)
Negative Contact	6.93 (14.05)	4.05 (7.93)	0.68 (1.83)	5.17 (13.21)
Processing	5.44 (12.93)	3.12 (7.11)	0.68 (1.83)	5.12 (13.20)
Restraint	1.48 (5.11)	0.94 (3.98)	0	0.05 (0.213)
No Contact	39.10 (28.51)	39.03 (27.96)	24.15 (28.68)	29.99 (32.62)



Figure 9. Mean Percentage Assistance at T1 and T2

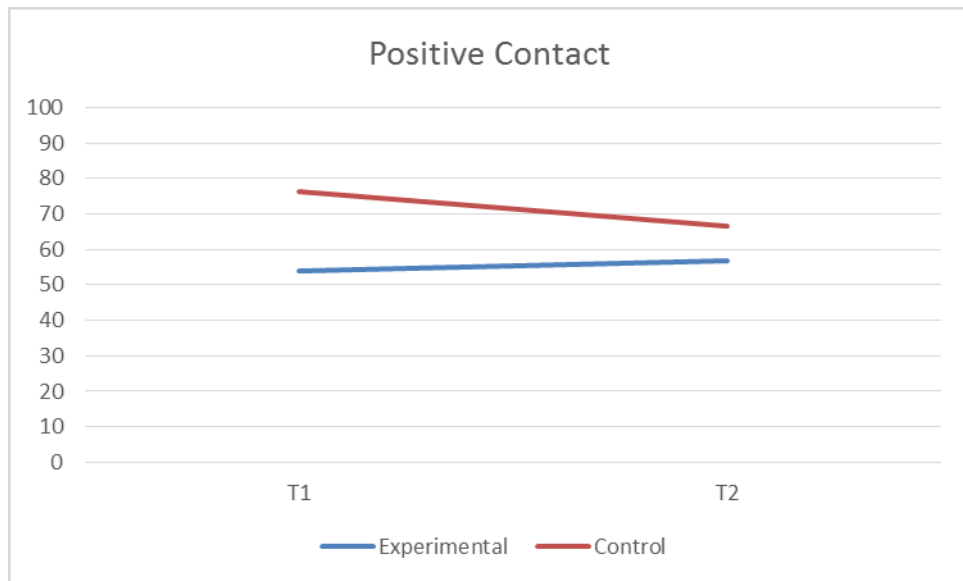


Figure 10. Mean Percentage Positive Contact at T1 and T2



Figure 11. Mean Percentage No Contact at T1 and T2

7.3.2.2 Significance Testing

In order to test difference in assistance, positive contact and no contact, mixed multi-factorial ANOVA was used. This found that there was no significant interaction between time and group for assistance ($F=0.985$; $p=0.324$; $df=1$; partial eta squared=0.014, a small effect size), for positive contact ($F=3.469$; $p=0.067$; $df=1$; partial eta squared=0.047, a small effect size), or for no contact ($F=0.687$; $p=0.410$; $df=1$; partial eta squared=0.01, a small effect size). As can be seen from appendix five, Tables 13.9 – 13.11, neither time nor group was significant either, for assistance, positive contact or no contact.

7.3.2.3 Sensitivity Analysis

Three additional analyses were carried out in addition to the main analysis for assistance, five additional analysis for positive contact and one for no contact, which had no outliers either with an ITT or PP approach. As part of this, the change in scores from T1-T2 were calculated for positive contact and tests were also done on these; this was to attempt to compensate to some extent for the fact that the control and experimental groups were significantly different at T1 for positive contact. These analyses show that none of these are significant regardless of how they are analysed (Bonferroni adjustment is used as there are three comparisons, so significance level is $p \leq 0.017$). The results are summarised in Table 7.4.

Table 7.4: Staff MTS Sensitivity Analysis T1-T2

Assistance	F	p	df	ES
ITT, outliers in	1.495	0.226	1	0.021*
PP, outliers out	0.795	0.376	1	0.014
PP, outliers in	1.396	0.242	1	0.024*
Positive Contact	t/F	p	df	ES
ITT, change score, outliers in	1.849	0.069	70	0.442*
ITT, change score, outliers out	1.377	0.173	69	0.33*
PP	F=5.821	0.019	1	0.091**
PP, change score, outliers in	2.401	0.020	58	0.63**
PP, change score, outliers out	1.995	0.051	57	0.53**
No Contact	F	p	df	ES
PP	2.979	0.090	1	0.049*

* Small effect size; ** Medium effect size

7.3.3 Active Support Measure

7.3.3.1 Descriptive Statistics

Table 7.5 shows the descriptive statistics for experimental and control groups at T1 and T2 for the ASM. Table 7.6 shows the percentage of participants achieving different categories of ASM scores (Mansell & Beadle-Brown, 2012); good (ASM score of at least 67%); mixed (ASM score of between 33% and 67%), or weak (ASM score of 33% or less). Figure 12 shows the graph of the mean percentage ASM score.

Table 7.5: ASM Descriptive Statistics T1 and T2

ASM	Experimental		Control	
	T1	T2	T1	T2
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Percentage Score	59 (21.35)	64 (21.76)	69 (18.87)	68 (18.24)

Table 7.6: Percentages of Group in Each Category of ASM Score at T1 and T2

ASM Score Categories	Experimental		Control	
	Percentage at T1	Percentage at T2	Percentage at T1	Percentage at T2
Good	38	48	64	68
Mixed	50	44	27	27
Weak	12	8	9	5

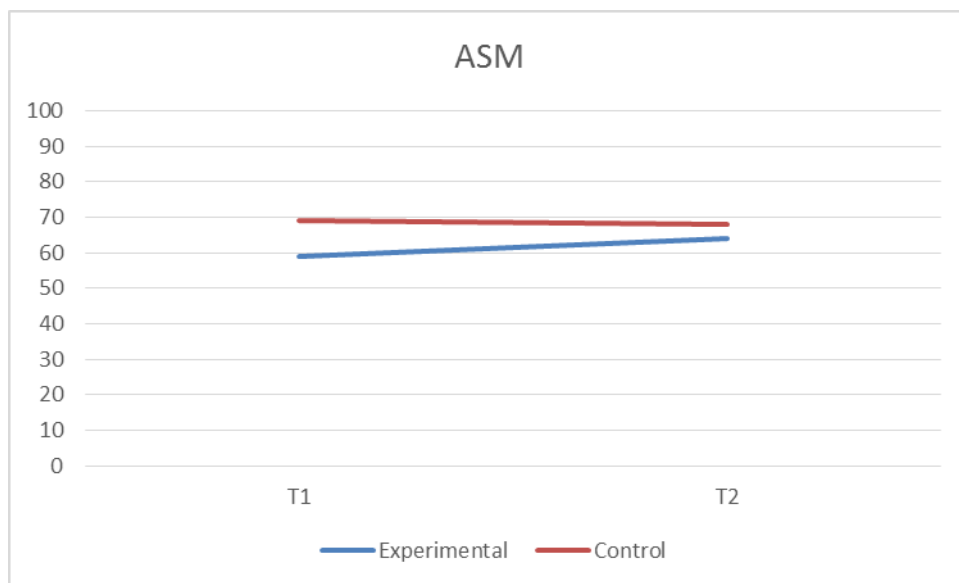


Figure 12. Mean Percentage ASM at T1 and T2

7.3.3.2 Significance Testing

A Mann-Whitney U test was carried out on the change in percentage ASM score from T1-T2 and this demonstrated there was no significant difference between the experimental and control groups ($U=460$; $p=0.270$; $n=72$; $r=0.13$, a small effect size). The use of change scores for analysis compensated to some extent for the fact that the control group scored significantly higher on the ASM at T1.

7.3.3.3 Sensitivity Analysis

Eight additional analyses were carried out in addition to the main analysis; these show that the results did not change dependent on method of analysis, with the exception of a multiple testing approach under the PP method. Using this analysis, the experimental group showed a significant difference in percentage ASM score from T1-T2; all other analyses showed no significant difference. The results are described in Table 7.7.

Table 7.7: ASM Sensitivity Analysis T1-T2

ASM	Test	p	df/n	ES
ITT, outliers out	U=412.5	0.468	n=68	0.09
ITT, ANOVA outliers out	F=0.572	0.452	1	0.009
ITT, multiple testing experimental	Z=1.470	0.141	n=50	0.21
ITT, multiple testing control	Z=0.131	0.896	n=22	0.02
ITT, multiple testing experimental	t=1.719	0.092	49	0.49*
ITT, multiple testing control	t=0.346	0.733	21	0.15
PP, outliers out	U=242	0.057	n=56	0.25*
PP, ANOVA outliers out	F=2.038	0.159	1	0.036*
PP, multiple testing experimental	Z=2.333	0.020†	n=38	0.33**
PP, multiple testing experimental	t=2.535	0.016†	37	0.83***

* Small effect size; ***Large effect size; † Significant at $p \leq 0.05$

7.3.3.4 ASM Individual Questions Analysis

In order to explore change in individual questions on the ASM, analysis was carried out on the change in mean from T1-T2 for each question. The results of this together with the means at T1 and T2 for each individual question are shown in Table 7.8. Bonferroni correction was used as although the questions within the ASM are distinct and separate, they are answered in relation to the same observational data. There was no significant difference between the experimental and control in any of the questions. Table 7.9 shows the percentage of each group scoring 'good' (a score of three) for each question for both the experimental and control groups at T1 and T2; this shows that the control group reduced scores from T1-T2 on 11 of the 15 questions; the experimental group scores did not reduce on any question.

Table 7.8: ASM Means per Question at T1 and T2, Percentage Change from T1-T2, and Comparison of Change Scores from T1-T2

ASM	Experimental			Control			Comparison of Change Scores in Experimental and Control from T1-T2	
	T1 Mean	T2 Mean	% Change T1-T2	T1 Mean	T2 Mean	% Change T1-T2	U	p
1. Age appropriateness of activities	1.95	2.15	10	2.29	2.40	5	228	0.810
2. Real rather than pretend activities	1.83	1.87	22	2.00	2.13	7	230	0.855
3. Choice of activities	1.57	1.67	6	2.00	1.73	-14	195	0.314
4. Demands presented carefully	1.57	2.03	17	2.10	2.20	5	190	0.255
5. Tasks appropriately analysed to facilitate involvement	1.67	1.95	17	2.19	2.20	0	221.5	0.688
6. Sufficient staff contact for service users	2.05	2.23	9	2.48	2.13	-14	203.5	0.413
7. Graded assistance to ensure success	1.63	2.08	28	2.19	1.53	-30	153.5	0.043
8. Speech level matches developmental level of service user	2.10	2.38	13	2.29	2.40	5	198	0.330
9. Interpersonal warmth	2.38	2.64	11	2.62	2.47	-6	176	0.129

ASM	Experimental			Control			Comparison of Change Scores in Experimental and Control from T1-T2	
	T1 Mean	T2 Mean	% Change T1-T2	T1 Mean	T2 Mean	% Change T1-T2	U	p
10. Differential reinforcement of adaptive behaviour	2.05	2.33	14	2.62	2.53	-3	180	0.158
11. Staff notice and respond to service user communication	2.31	2.32	0	2.67	2.13	-20	153.5	0.038
12. Staff manage challenges well	1.62	1.79	10	1.29	2.13	65	151	0.043
13. Staff work as a coordinated team to support service users	1.69	2.05	21	2.24	2.00	-11	147	0.031
14. Teaching is embedded in everyday activities	1.10	1.63	48	1.62	1.80	11	222.5	0.719
15. Written plans are in routine use	0.62	1.05	69	0.75	1.14	52	196	0.631

Table 7.9: Percentage Scoring ‘Good’ per Question at T1 and T2 for Experimental and Control

ASM Item	Experimental T1		Control T2	
	Percentage Scoring Good		Percentage Scoring Good	
	T1	T2	T1	T2
1. Age appropriateness of activities	34	36	59	36
2. Real rather than pretend activities	22	24	27	27
3. Choice of activities	12	20	32	14
4. Demands presented carefully	12	30	41	27
5. Tasks appropriately analysed to facilitate involvement	18	26	41	27
6. Sufficient staff contact for service users	30	36	73	32
7. Graded assistance to ensure success	22	34	41	18
8. Speech level matches developmental level of service user	18	34	27	45
9. Interpersonal warmth	28	52	68	41
10. Differential reinforcement of adaptive behaviour	14	34	59	45
11. Staff notice and respond to service user communication	28	36	77	27
12. Staff manage challenges well	12	24	23	32
13. Staff work as a coordinated team to support service users	6	28	36	27
14. Teaching is embedded in everyday activities	0	10	23	18
15. Written plans are in routine use	2	6	0	0

7.3.4 PBS Knowledge

7.3.4.1 Descriptive Statistics

Table 7.10 shows the mean and standard deviation for PBS knowledge for the experimental and control groups at T1 and T2. Figure 13 shows the graph of the means.

Table 7.10: Staff PBS Knowledge Descriptive Statistics T1 and T2

PBS Knowledge	Experimental		Control	
	T1	T2	T1	T2
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Percentage Score	55.53(11.56)	61.17(12.57)	54.43(11.08)	61.12(10.92)

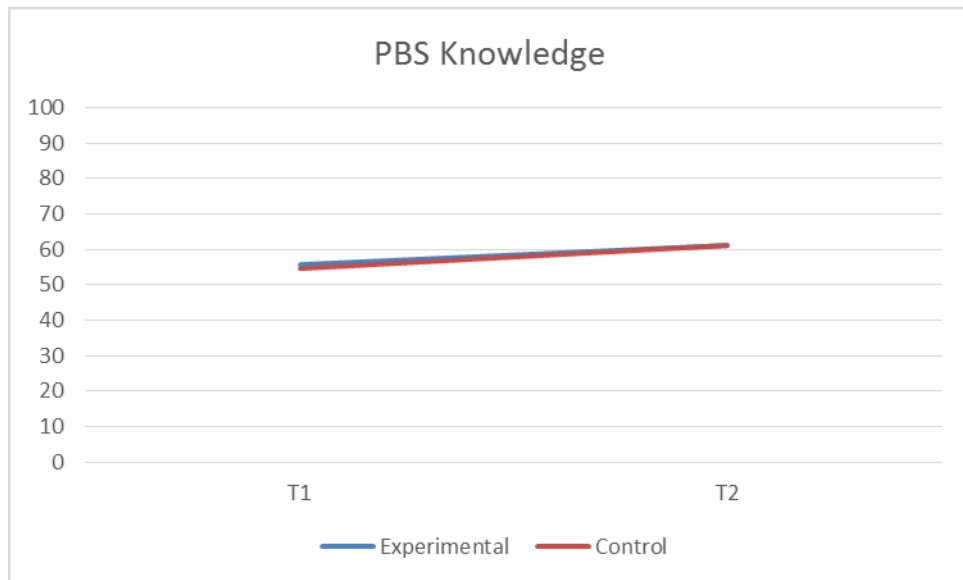


Figure 13. Staff Mean Percentage PBS Knowledge T1 and T2

7.3.4.2 Significance Testing

There was no significant interaction between time and group for PBS knowledge ($F=0.202$; $p=0.655$; $df=1$; partial eta squared=0.003, a negligible effect size). As can be seen from appendix five, Table 13.12, neither time nor group was significant either.

7.3.4.3 Sensitivity Analysis

Three additional analyses were carried out in addition to the main analysis for PBS knowledge; these show that the results do not change dependent on method of analysis. This is summarised in Table 7.11.

Table 7.11: Staff PBS Knowledge Sensitivity Analysis T1-T2

PBS Knowledge	F	p	df	ES
ITT, ANOVA outliers in	0.073	0.788	1	0.001
PP, ANOVA outliers out	0.537	0.467	1	0.010*
PP, ANOVA outliers in	0.025	0.874	1	0.000

* Small effect size

7.3.5 Challenging Behaviour Attribution Scale (CHABA)

7.3.5.1 Descriptive Statistics

Table 7.12 shows the mean and standard deviation for the CHABA subscales for the experimental and control groups at T1 and T2.

Table 7.12: Staff CHABA Descriptive Statistics T1 and T2

CHABA	Experimental		Control	
	T1	T2	T1	T2
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Learned Positive	1.04 (0.67)	0.93 (0.64)	0.92 (0.69)	0.78 (0.60)
Learned Negative	0.92 (0.58)	0.70 (0.62)	0.95 (0.64)	0.83 (0.48)
Biomedical	0.39 (0.58)	0.25 (0.49)	0.46 (0.42)	0.27 (0.55)
Stimulation	0.32 (0.66)	0.33 (0.66)	0.38 (0.53)	0.31 (0.65)
Emotional	1.03 (0.56)	0.93 (0.57)	1.04 (0.42)	1.04 (0.45)
Physical Environment	0.27 (0.66)	0.30 (0.62)	0.46 (0.49)	0.40 (0.66)

7.3.5.2 Significance Testing

In order to test difference in the CHABA subscales mixed multi-factorial ANOVA were used.

This found that there was no significant interaction between time and group for any of the subscales. The results are summarised in Table 7.13. As can be seen from appendix five, Table 13.13 – Table 13.18, neither time nor group was significant either, for any of the subscales.

Table 7.13: Staff CHABA Significance Testing T1-T2

	F	P	df	ES
Learned Positive	0.294	0.590	1	0.004
Learned Negative	0.030	0.863	1	0.000
Biomedical	0.184	0.669	1	0.003
Stimulation	0.426	0.516	1	0.006
Emotional	0.480	0.491	1	0.007
Physical Environment	0.471	0.495	1	0.007

7.3.5.3 Sensitivity Analysis

Three additional analyses were carried out in addition to the main analysis for each of the subscales of the CHABA; these show that the results do not change dependent on method of analysis. This is summarised in Table 7.14.

Table 7.14: Staff CHABA Sensitivity Analysis T1-T2

Learned Positive	F	p	df	ES
ITT, outliers in	0.006	0.939	1	0.000
PP, outliers out	0.042	0.839	1	0.001
PP, outliers in	0.011	0.917	1	0.000
Learned Negative	F	p	df	ES
ITT, outliers in	0.196	0.660	1	0.003
PP, outliers out	0.060	0.807	1	0.001
PP, outliers in	0.699	0.406	1	0.012*
Biomedical	F	p	df	ES
ITT, outliers in	0.076	0.784	1	0.001
PP, outliers out	0.138	0.711	1	0.002
PP, outliers in	0.032	0.859	1	0.001
Stimulation	F	p	df	ES
ITT, outliers in	0.135	0.714	1	0.002
PP, outliers out	0.130	0.720	1	0.002
PP, outliers in	0.004	0.950	1	0.000
Emotional	F	p	df	ES
ITT, outliers in	0.313	0.577	1	0.004
PP, outliers out	1.486	0.228	1	0.025*
PP, outliers in	1.075	0.304	1	0.018
Physical Environment	F	p	df	ES
ITT, outliers in	0.269	0.605	1	0.004
PP, outliers out	0.290	0.592	1	0.005
PP, outliers in	0.116	0.730	1	0.002

* Small effect size

7.3.6 Correlations

Correlations were carried out between staff measures, and also between some staff measures and managers' measures where relevant. Assistance was significantly positively correlated with ASM, at T1 and T2 for both the experimental and control groups. PSR was significantly positively correlated with ASM at T2 for the experimental group. Practice leadership (PL) at T2 was significantly positively correlated with ASM at T2 in the experimental group. None of the other correlations were significant. The results are summarised in Table 7.15.

Table 7.15: Correlations for Staff and Managers' Measures T1 and T2

Correlations	Experimental		Control	
	r	p	r	p
Assistance T1/ASM T1	0.566	p<0.001	0.762	0.000†
Assistance T2/ASM T2	0.501	p<0.001	0.619	0.002†
Staff Knowledge T1/Assistance T1	0.128	0.374	0.353	0.107
Staff Knowledge T2/Assistance T2	0.140	0.334	0.065	0.774
Staff Knowledge T2/Managers' Knowledge T2	0.314	0.026	-0.219	0.328
PSR T2/ASM T2	0.535	p<0.001	-	-
PSR T2/Assistance T2	0.213	0.137	-	-
PL T2/ASM T2	0.501	p<0.001	0.122	0.590
PL T2/Assistance T2	0.261	0.067	0.000	0.999

Practice leadership (PL) scores were also considered specifically in relation to ASM scores, to see if services within the different categories of ASM score scored differently in terms of practice leadership; in both groups, services with good ASM scores had the highest practice leadership scores. The results are shown in Table 7.16.

Table 7.16: Practice Leadership Scores for Each Category of ASM Score at T1 and T2

	Experimental		Control	
	PL Score T1	PL Score T2	PL Score T1	PL Score T2
Good ASM Services	68	77	62	63
Mixed ASM Services	64	68	52	63
Weak ASM Services	48	40	51	38

7.3.7 Summary of Staff Results at T2

7.3.7.1 Support to Service Users

Results from the MTS show that following the training there was no significant difference in staff assistance, positive contact or no contact to service users, with the exception of positive contact when analysed on a PP basis. The ASM results also show no significant difference except when analysed using multiple testing on a PP basis; the experimental group shows significant change with this approach. The PP analysis excludes data from 12 service users and their staff. For six of these service users there was no observation data; three of these individuals were imputed with an IAWO approach, because their support had broken down due to their challenging behaviour; the others were imputed with a GA approach because the service users had either left the organisation in positive circumstances, or refused to be observed. The PP analysis also excluded data from the six service users whose managers had either left the course (three) or left the organisation (three). It appears therefore that the approach to imputation makes some difference to significant change in terms of staff support to service users following the training; however due to the fact that significant change is only evident from a PP analysis, it is not possible to have confidence in these results. In summary, it is judged that there can be a confident conclusion regarding the lack of significant changes in assistance to, or contact with, service users, as reflected by these two measures.

7.3.7.2 Staff Knowledge and Attributions

There was a slight increase in PBS knowledge at T2 for both the experimental and control groups; however this change was not significant, regardless of how it was analysed.

In the CHABA, all subscales decreased following the training for both the experimental and control groups, with the exception of Stimulation and Physical Environment, which increased very slightly for the experimental group, and for Emotional which stayed the same for the control group. However, none of these changes were significant for any subscale, even when analysed in a range of different ways. It can therefore be confidently concluded that there were no significant changes to staff knowledge about PBS or to staff attributions about challenging behaviour following the PBS training for their managers.

7.4 Staff Results at T3: Maintenance of Training Effects

7.4.1 Introduction

This section of the staff results chapter considers the data gathered at follow-up, T3. The question being considered in this section is different from the previous results section which considered initial impact of the training, that is, comparisons from T1-T2. This section focuses

on changes from T1-T3, in order to consider the training's long-term effectiveness; this was done for all measures. As there were no significant changes on any measure from T1-T2, no T2-T3 tests were done, since the purpose of these would have been to consider whether changes occurring from T1-T2 were maintained. Throughout this section, the PP analysis is based on data from only 22 managers, due to either managers leaving the course or the service, or service users leaving the organisation and therefore the results from the PP analysis need to be treated with some caution as they represent less than half of the experimental group.

7.4.2 Momentary Time Sampling

7.4.2.1 Descriptive Statistics

Table 7.17 shows the mean and standard deviation at T1, T2 and T3 for MTS behaviours: total positive contact (assistance plus other conversation), total negative contact (processing plus restraint), and no contact, for the ITT dataset and also for the PP dataset, with and without outliers where appropriate. Figures 14-16 show the graphs of the means for assistance, no contact and total positive contact at T1, T2 and T3 for the experimental group ITT dataset.

Table 7.17: Staff MTS Descriptive Statistics T1-T3

MTS (<i>n</i>)	T1	T2	T3
	Mean (SD)	Mean (SD)	Mean (SD)
Total Positive Contact			
ITT, outliers in (50)	53.97 (27.21)	56.93 (26.60)	59.62 (20.40)
ITT, outliers out (47)	54.71 (27.71)	59.42 (25.26)	63.07 (15.48)
PP, outliers in (21)	53.25 (31.70)	64.30 (27.26)	61.27 (18.65)
PP, outliers out (18)	51.36 (31.61)	62.58 (26.76)	60.71 (9.97)
Assistance			
ITT, outliers in (50)	35.53 (24.69)	35.83 (24.59)	37.94 (15.63)
ITT, outliers out (48)	34.73 (24.86)	33.22 (21.08)	37.37 (15.61)
PP, outliers in (21)	35.42 (27.55)	46.46 (25.10)	43.65 (16.42)
PP, outliers out (19)	33.39 (28.19)	40.83 (18.67)	42.83 (16.84)
Other Conversation			
ITT, outliers in (50)	18.44 (19.22)	21.10 (21.99)	21.69 (17.81)
ITT, outliers out (46)	16.05 (15.80)	17.72 (17.94)	20.76 (16.97)
PP, outliers in (21)	17.83 (21.77)	17.86 (18.44)	17.62 (19.07)
PP, outliers out (20)	14.93 (17.68)	17.94 (18.92)	18.50 (19.12)
Total Negative Contact			
ITT, outliers in (50)	6.93 (14.05)	4.05 (7.93)	1.20 (2.48)

MTS (<i>n</i>)	T1	T2	T3
	Mean (SD)	Mean (SD)	Mean (SD)
PP, outliers in (21)	4.11 (7.15)	3.20 (6.26)	1.10 (2.88)
PP, outliers out (15)	1.33 (2.73)	1.32 (1.98)	0.11 (0.24)
Processing			
ITT, outliers in (50)	5.44 (12.93)	3.12 (7.11)	0.94 (2.35)
PP, outliers in (21)	2.78 (4.82)	3.03 (6.30)	1.04 (2.86)
PP, outliers out (15)	0.65 (1.24)	1.39 (1.95)	0.11 (0.24)
Restraint			
ITT, outliers in (50)	1.48 (5.11)	0.94 (3.98)	0.26 (0.76)
PP, outliers in (21)	1.32 (2.59)	0.17 (0.78)	0.06 (0.22)
PP, outliers out (18)	0	0	0
No Contact			
ITT, outliers in (50)	39.10 (28.51)	39.03 (27.96)	39.37 (20.29)
ITT, outliers out (45)	36.01 (28.93)	36.10 (25.35)	35.20 (11.89)
PP, outliers in (21)	42.66 (32.33)	32.50 (26.53)	38.11 (17.81)
PP, outliers out (20)	41.15 (32.40)	31.27 (26.60)	35.29 (12.60)



Figure 14. Mean Percentage Assistance T1-T3



Figure 15. Mean Percentage Positive Contact T1-T3

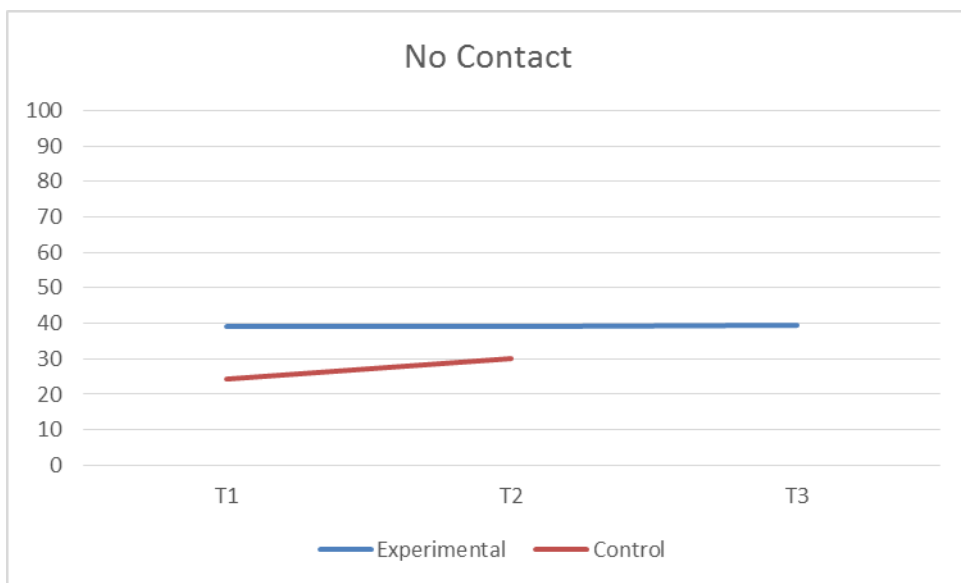


Figure 16. Mean Percentage No Contact T1-T3

7.4.2.2 Significance Testing

In order to test difference in assistance, positive contact and no contact, related t-tests were used. These found that there was no significant difference in assistance at T1-T3 ($t=0.616$; $p=0.540$; $df=49$; $d=0.18$, a negligible effect size) and no significant difference for positive contact at T1-T3 ($t=1.955$; $p=0.057$; $df=46$; $d=0.58$, a medium effect size). As data for no contact were not normally distributed a Wilcoxon Signed Ranks test was carried out and found no significant difference at T1-T3 ($Z=0.249$; $p=0.807$; $r=0.04$, a negligible effect size).

In order to explore potential links between assistance and management, tests were carried out to check for any difference in assistance between the group whose manager had changed by T3, and the group whose manager remained the same. This showed that there was a significant difference between the groups ($t=2.307$; $p=0.025$; $df=48$; $d=0.66$, a medium effect size), with a mean assistance for the unchanged group ($n=22$) of 43.45 and a mean assistance of 33.60 for the group whose manager had changed ($n=28$).

7.4.2.3 Sensitivity Analysis

Three additional analyses were carried out in addition to the main analysis for each of the measures except assistance T1-T3 which had no outliers either with a PP or ITT approach. The PP analysis cannot be regarded as robust as it relates to less than half the experimental group, mainly due to either managers or service users leaving. However, none of these measures show any significant difference, regardless of how they are analysed, and therefore the PP analysis agrees with the ITT analysis, the main difference being in terms of larger effect sizes with the PP analysis. The results are summarised in Table 7.18 and descriptive statistics for the sensitivity analysis are above in Table 7.17.

Table 7.18: Staff MTS Sensitivity Analysis T1-T3

MTS Assistance T1-T3	t	p	df/n	ES
PP, (no outliers)	1.379	0.183	20	0.62**
MTS Positive Contact T1-T3	t	p	df	ES
ITT, outliers in	1.306	0.198	49	0.37*
PP, outliers out	1.311	0.207	17	0.64**
PP, outliers in	1.256	0.224	20	0.56**
MTS No Contact T1-T3	t/Z	p	df	ES
ITT, outliers out	Z=0.613	0.546	n=45	r=0.09
PP, outliers out	0.914	0.372	19	0.42*
PP, outliers in	0.730	0.474	20	0.33*

* Small effect size; ** Medium effect size

7.4.3 Active Support Measure

7.4.3.1 Descriptive Statistics

Table 7.19 shows the descriptive statistics for the experimental group at T1, T2 and T3 for the ASM for the ITT dataset and also for the PP dataset, with and without outliers. Table 7.20 shows the percentage of participants with good, mixed and weak ASM scores at T1, T2 and T3, and Table 7.21 shows the practice leadership scores in relation to whether ASM scores were good,

mixed or weak. Figure 17 shows the graphs of the mean percentage score at the same time points for the ITT dataset.

Table 7.19: ASM Descriptive Statistics T1-T3

ASM (<i>n</i>)	T1	T2	T3
	Mean (SD)	Mean (SD)	Mean (SD)
ITT, outliers in (50)	59 (21.35)	64 (21.76)	68 (22.13)
PP, outliers in (21)	59 (21.61)	73 (18.28)	80 (14.26)
PP, outliers out (20)	62 (17.71)	65 (22.00)	82 (12.01)

Table 7.20: Percentage of Group in Each Category of ASM Score T1-T3

ASM Categories	T1	T2	T3
Good	38	48	58
Mixed	50	44	32
Weak	12	8	10

Table 7.21: Practice Leadership Scores for Each Category of ASM Score T1-T3

	PL Score T1	PL Score T2	PL Score T3
Good ASM Services	68	77	72
Mixed ASM Services	64	68	61
Weak ASM Services	48	40	43

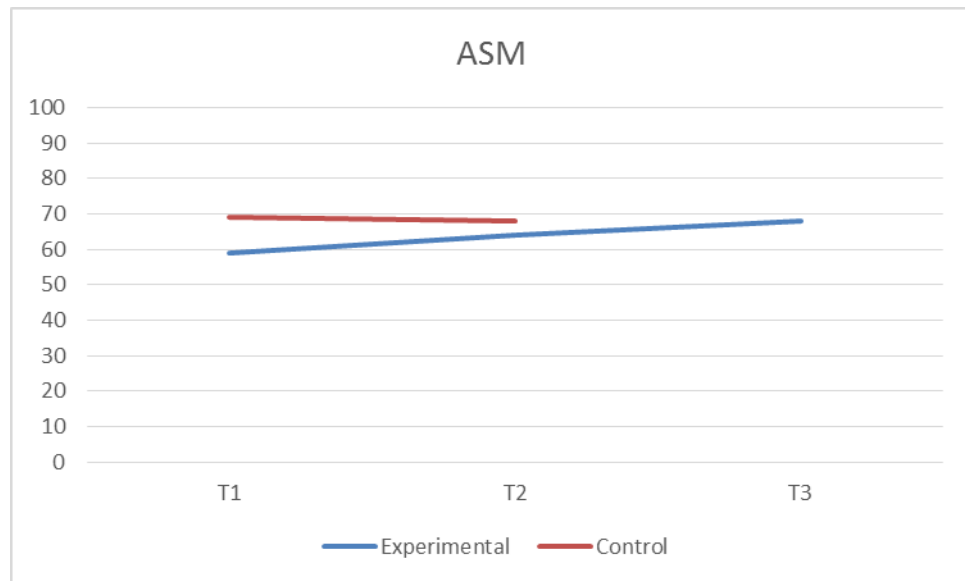


Figure 17. Mean Percentage ASM T1-T3

7.4.3.2 Significance Testing

A related t-test was carried out on the percentage ASM score from T1-T3. This demonstrated there was a significant difference from T1-T3 ($t=2.417$; $p=0.019$; $df=49$; $d=0.69$, a medium effect size).

In order to explore potential links between ASM and management, tests were carried out to check for any difference in assistance between the group whose manager had changed by T3, and the group whose manager remained the same. This showed that there was a significant difference between the groups ($U=141.5$; $p=0.001$; $n=50$; $r=0.46$, a medium effect size), with a mean ASM for the unchanged group ($n=22$) of 79% and a mean ASM of 59% for the group whose manager had changed ($n=28$).

7.4.3.3 Sensitivity Analysis

Two additional analyses for T1-T3 were carried out in addition to the main analysis. These show that the results did not change dependent on method of analysis; there is still a significant difference from T1-T3, although effect size is larger with the PP analysis. The results are described in Table 7.22 and descriptive statistics for the sensitivity analysis are above in Table 7.19.

Table 7.22: ASM Sensitivity Analysis T1-T3

ASM T1-T3	t	p	df	ES
PP, outliers in	4.126	0.001†	20	1.85***
PP, outliers out	3.810	0.001†	19	1.75***

***Large effect size; † Significant at $p \leq 0.05$

7.4.4 PBS Knowledge

7.4.4.1 Descriptive Statistics

Table 7.23 shows the mean and standard deviation for PBS knowledge for the experimental group at T1, T2 and T3 for the ITT dataset and also for the PP dataset, with and without outliers.

Figure 18 shows the graph of the means for the ITT dataset.

Table 7.23: Staff PBS Knowledge Descriptive Statistics T1-T3

PBS Knowledge (<i>n</i>)	T1	T2	T3
	Mean (SD)	Mean (SD)	Mean (SD)
ITT, outliers in (50)	55.53 (11.56)	61.17 (12.57)	58.93 (15.61)
ITT, outliers out (49)	56.17 (10.73)	61.47 (12.52)	59.07 (15.74)
PP, outliers in (22)	55.05 (12.42)	62.33 (10.58)	58.33 (16.59)
PP, outliers out (21)	56.52 (10.56)	63.09 (10.21)	58.62 (16.94)

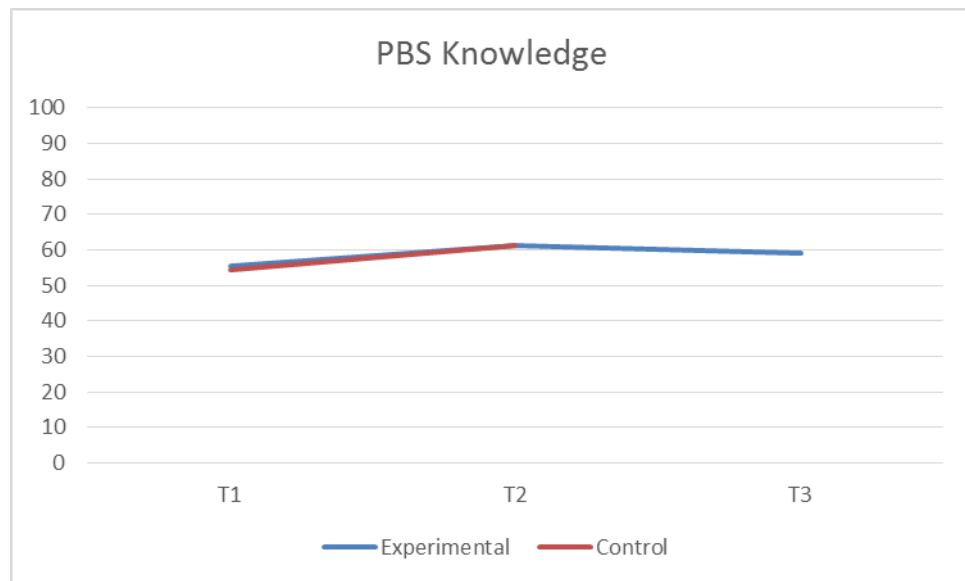


Figure 18. Staff Mean Percentage PBS Knowledge T1-T3

7.4.4.2 Significance Testing

There was no significant difference from T1-T3 for PBS knowledge ($t=1.122$; $p=0.267$; $df=48$; $d=0.32$, a small effect size).

7.4.4.3 Sensitivity Analysis

Three additional analyses were carried out in addition to the main analysis for PBS knowledge for T1-T3; these show that the results do not change dependent on method of analysis. This is summarised in Table 7.24 and descriptive statistics for the sensitivity analysis are above in Table 7.23.

Table 7.24: Staff PBS Knowledge Sensitivity Analysis T1-T3

PBS Knowledge T1-T3	t	p	df	ES
ITT, outliers in	1.320	0.193	49	0.38*
PP, outliers out	0.421	0.679	18	0.20*
PP, outliers in	0.678	0.506	20	0.30*

* Small effect size

7.4.5 Challenging Behaviour Attribution Scale (CHABA)

7.4.5.1 Descriptive Statistics

Table 7.25 shows the mean and standard deviation for the CHABA subscales for the experimental group at T1, T2 and T3 for the ITT dataset and also for the PP dataset, with and without outliers.

Table 7.25: Staff CHABA Descriptive Statistics T1-T3

CHABA (n)	T1	T2	T3
	Mean (SD)	Mean (SD)	Mean (SD)
Learned Positive			
ITT, outliers in (50)	1.04 (0.67)	0.93 (0.64)	0.88 (0.57)
ITT, outliers out (49)	1.05 (0.67)	0.96 (0.61)	0.87 (0.57)
PP, outliers in (22)	0.95 (0.63)	0.69 (0.72)	0.70 (0.61)
Learned Negative			
ITT, outliers in (50)	0.92 (0.58)	0.70 (0.62)	0.89 (0.59)
PP, outliers in (22)	0.88 (0.57)	0.42 (0.65)	0.77 (0.65)
PP, outliers out (13)	0.82 (0.22)	0.43 (0.51)	0.92 (0.63)
Biomedical			
ITT, outliers in (50)	0.39 (0.58)	0.25 (0.49)	0.48 (0.69)
ITT, outliers out (47)	0.41 (0.57)	0.28 (0.41)	0.48 (0.71)
PP, outliers in (22)	0.32 (0.53)	0.14 (0.56)	0.37 (0.76)
Stimulation			
ITT, outliers in (50)	0.32 (0.66)	0.33 (0.66)	0.37 (0.65)
ITT, outliers out (47)	0.32 (0.67)	0.36 (0.66)	0.39 (0.65)
PP, outliers in (22)	0.21 (0.57)	0.11 (0.77)	0.22 (0.65)
PP, outliers out (20)	0.19 (0.58)	0.12 (0.55)	0.26 (0.65)
Emotional			
ITT, outliers in (50)	1.03 (0.56)	0.93 (0.57)	0.97 (0.63)
ITT, outliers out (48)	1.06 (0.51)	0.98 (0.58)	1.03 (0.57)
PP, outliers in (22)	0.89 (0.66)	0.67 (0.64)	0.77 (0.72)
Physical Environ- ment			
	0.27 (0.66)	0.30 (0.62)	0.43 (0.68)
ITT, outliers in (50)	0.28 (0.66)	0.34 (0.55)	0.50 (0.57)
ITT, outliers out (48)	0.15 (0.64)	0.12 (0.66)	0.41 (0.62)
PP, outliers in (22)	0.15 (0.65)	0.14 (0.43)	0.39 (0.56)
PP, outliers out (20)			

7.4.5.2 Significance Testing

Related t-tests were used to test difference in the CHABA subscales. These found that there was no significant difference between T1-T3 for any of the subscales. The results are summarised in Table 7.26 and descriptive statistics for the sensitivity analysis are above in Table 7.25.

Table 7.26: Staff CHABA Significance Testing T1-T3

T1-T3	t	p	df	ES
Learned Positive	1.476	0.146	49	0.42*
Learned Negative	0.242	0.810	49	0.07
Biomedical	0.814	0.420	49	0.23*
Stimulation	0.384	0.702	49	0.11
Emotional	0.373	0.711	47	0.11
Physical Environ- ment	1.661	0.103	48	0.48*

* Small effect size; ** Medium effect size

7.4.5.3 Sensitivity Analysis

Additional analyses were carried out in addition to the main analysis for each of the subscales of the CHABA for T1-T3; some subscales had no outliers, or too many to take out, so fewer tests were done for these. Tests show that the results do not change dependent on method of analysis. This information is summarised in Table 7.27.

Table 7.27: Staff CHABA Sensitivity Analysis T1-T3

Learned Positive T1-T3	t	p	df	ES
PP, (no outliers)	1.444	0.164	21	0.63**
Learned Negative T1-T3	t	p	df	ES
PP, outliers in	0.601	0.554	21	0.26*
Biomedical T1-T3	t	p	df	d
PP, (no outliers)	0.330	0.744	21	0.14
Stimulation T1-T3	t	p	df	d
PP, outliers out	0.043	0.966	20	0.02
PP, outliers in	0.454	0.655	21	0.24*
Emotional T1-T3	t	p	df	d
ITT, outliers in	0.599	0.552	49	0.17
PP, (no outliers)	0.598	0.556	21	0.26*
Physical Environment T1-T3	t	p	df	d
ITT, outliers in	1.143	0.259	49	0.33*
PP, (no outliers)	1.385	0.181	21	0.60**

* Small effect size; ** Medium effect size

7.4.6 Correlations

Correlations were carried out between staff measures, and also between some staff measures and managers' measures where relevant. Assistance was significantly positively correlated with PSR scores, with practice leadership scores and with the ASM, but not with staff knowledge at T3. The ASM was significantly positively correlated with PSR scores and with practice leadership. The results are summarised in Table 7.28.

Table 7.28: Correlations for Staff and Managers' Measures at T3

Correlations	r	p
Assistance /Staff Knowledge	0.204	0.155
Assistance /PSR	0.548	p<0.001
Assistance/Practice Leadership	0.595	p<0.001
Assistance/ASM	0.615	p<0.001
ASM/PSR	0.476	p<0.001
ASM /Practice Leadership	0.507	p<0.001

7.5 Chapter Summary

7.5.1.1 Support to Service Users

Results from the staff MTS show that there were no significant differences in staff assistance, positive contact or no contact to service users either from T1-T2, or from T1-T3. Assistance and positive contact increased slightly but not significantly at each time point; no contact also increased slightly but not significantly. These results do not change regardless of method of analysis, whether including outliers or not, and whether using ITT or PP approach. This indicates that there can be confidence in the results.

The ASM did not change significantly from T1-T2; however, it increased significantly from T1-T3 and the results did not change dependent on method of analysis, although effect size is larger with PP analysis. This indicates that we can confidently conclude that there was an increase over time in staff providing active support to service users.

7.5.1.2 Staff Knowledge

There was a slight increase in PBS knowledge from T1-T2 and from T1-T3; however, neither of these was significant, regardless of method of analysis. It can therefore be confidently concluded that there was no significant changes to staff knowledge about challenging behaviour following PBS training for their managers.

7.5.1.3 Staff Attributions

In the CHABA, all subscales reduced from T1-T2 for both experimental and control group, but none of these changes were significant. From T1-T3 some subscales reduced slightly (Learned Positive, Learned Negative and Emotional), while others increased slightly (Biomedical, Stimulation and Physical Environment); however none of these changes were significant, regardless of method of analysis. It can therefore be confidently concluded that there was no significant changes to staff attributions about challenging behaviour following PBS training for their managers.

7.5.1.4 Correlations

The ASM was significantly positively correlated with assistance, with PSR scores, and with practice leadership, at T2 and T3. Assistance was significantly positively correlated with the ASM at T2, and with PSR scores, with practice leadership, and with ASM at T3.

8 Service User Results

8.1 Chapter Outline

This chapter details the results in relation to the service user measures: Adaptive Behaviour Scale (ABS) (Nihira et al, 1993), Aberrant Behaviour Checklist (ABC) (Aman & Singh, 1986), Behaviour Recording Forms (BRF), Momentary Time Sampling (MTS), and Guernsey Community and Participation Leisure Assessment (GCPLA) (Baker, 2000). The results at baseline are considered first in order to identify any differences between the groups and between the cohorts at T1. The results following training are then presented with analysis of changes in each measure at T2; all tests of difference are followed up with sensitivity analyses in order to check the robustness of the results. This considers the impact of using an intention to treat (ITT) or a per protocol (PP) approach to data analysis, and also the impact of outliers on results. Results are also presented from correlations carried in out in relation to the managers' and the staff's measures. Finally there is a consideration of the service users' results at T3, to consider whether any changes following training have been maintained. The descriptive statistics for the sensitivity analyses are also presented in this section of the chapter.

8.2 Service User Results at T1

8.2.1 Testing Differences between Cohort One and Cohort Two

As previously noted, this study involved two separate cohorts over two years, and in both cohorts there was an experimental and a control group. Analysis was carried out to establish if the two cohorts were significantly different, or if they could be combined and regarded as one group. Tests were carried out on every measure to look at differences between cohorts one and two, both for the experimental group and also for the control group. As can be seen in Table 8.2, none of the measures showed any significant differences between cohort one and two, in either the experimental or control groups. The cohorts are therefore combined: both cohorts of the experimental are regarded as one group for all further analysis, and both cohorts of the control group are regarded as one group for all further analysis.

Table 8.2: Difference between Cohort One and Cohort Two at T1 for Service User Measures for Experimental and Control Groups

Experimental Group						
Measures	Cohort 1 Mean (SD)	Cohort 2 Mean (SD)	t	p	df	ES
ABS	134.04 (45.99)	130.48 (65.46)	0.223	0.825	48	0.24*
ABC Total	54.72 (23.51)	60.16 (21.50)	0.854	0.397	48	0.25*
ABC Severe	5.20 (5.48)	6.88 (5.97)	1.458	0.151	47	0.42*
BRF Frequency	9.52 (8.81)	11.28 (9.13)	0.755	0.455	44	0.22*
BRF Severity	6.70 (1.92)	6.88 (1.99)	1.125	0.267	44	0.33*
GCPLA Range	20.76 (8.45)	19.24 (7.16)	0.686	0.496	48	0.20*
GCPLA Busy	12.60 (5.45)	12.08 (5.28)	1.227	0.226	42.6	0.37*
MTS Engagement	68.83 (27.04)	72.22 (23.06)	0.437	0.664	47	0.13
MTS Disengagement	26.92 (27.06)	25.71 (22.89)	0.825	0.413	45	0.25*
MTS Challenging Behaviour	4.25 (6.64)	2.07 (6.51)	U= 223.5	0.063	n=50	0.26*
Control Group						
Measures	Cohort 1 Mean (SD)	Cohort 2 Mean (SD)	t	p	df	ES
ABS	139.83 (63.14)	157.20 (54.32)	0.684	0.502	20	0.31*
ABC Total	30.58 (26.36)	40.80 (25.08)	1.956	0.065	19	0.9***
ABC Severe	2.83 (7.06)	3.40 (5.78)	U=46.5	0.381	n=22	0.20*
BRF Frequency	8.67 (13.11)	7.80 (6.11)	0.104	0.919	11.9	0.06
BRF Severity	7.18 (2.13)	6.20 (1.22)	0.873	0.394	19	0.40
GCPLA Range	23.08 (10.20)	16.50 (6.01)	1.794	0.088	20	0.8***
GCPLA Busy	12.42 (6.27)	11.50 (4.14)	0.395	0.697	20	0.18
MTS Engagement	83.74 (29.28)	89.61 (12.46)	U=45.5	0.346	n=22	0.205
MTS Disengagement	15.70 (28.46)	9.58 (12.89)	0.600	0.556	18	0.28*
MTS Challenging Behaviour	0.56 (1.18)	0.81 (1.97)	U=58	0.923	n=22	0.04

* Small effect size; *** Large effect size

8.2.2 Testing Difference between Control and Experimental T1

Tests were undertaken to establish if there were any significant differences between control and experimental groups at T1, both in terms of their characteristics and the measures used; Bonferroni adjustment was used to address the number of comparisons done on each set of data. This demonstrated that the experimental group had significantly higher ABC scores (both total and severe behaviours) lower engagement, higher disengagement, higher staff to service user ratio (both WTE and during observation), and lower numbers of service users per setting. There were no significant differences for ABS, BRF (frequency or severity), GCPLA (range or busy), MTS challenging behaviour, service user age, or presence of an autism diagnosis. This information is summarised in Table 8.3; the descriptive statistics for each measure are found in the individual sections in 6.3 below.

Table 8.3: Difference between Experimental and Control for Service User Measures and Characteristics T1

Measures	U	p	n	ES
ABS Scores	480	0.392	72	0.10*
ABC Total	249	p<0.0010.004†	72	0.43**
ABC Severe Behaviours	314.5		72	0.26*
BRF Frequency	395	0.057	72	0.22*
BRF Severity	534	0.850	72	0.02
GCPLA Range	549	0.990	72	0.001
GCPLA Busy	547.5	0.976	72	0.001
MTS Total Engagement	329	0.007††	72	0.32**
MTS Disengagement	348.5	0.014††	72	0.29*
MTS Challenging Behaviour	398	0.036	72	0.25*
Characteristics	U	p	n	ES
Age	518	0.695	72	0.05
Gender	490	0.369	72	0.11*
Autism Diagnosis	547	0.966	72	0.001
WTE Ratio	367.5	0.025†	72	0.26*
Observation Ratio	449	0.040†††	71	0.24*
Number of Service Users	360.5	0.019†	72	0.27*

* Small effect size; **Medium effect size; † Significant at $p \leq 0.025$; †† Significant at $p \leq 0.017$; ††† Significant at $p \leq 0.05$

8.2.3 Correlations at T1

ABC score was significantly negatively correlated with CGPLA busy for the experimental group and with total engagement for the control group. The Active Support Measure (ASM) and engagement were significantly positively correlated for both groups. No other correlations were significant. The results are shown in Table 8.4.

Table 8.4: Correlations at T1 for Managers', Staff, and Service User Measures

T1 Correlations	Experimental		Control	
	r	p	r	p
ABC/BRF	Rho=0.290	0.041	Rho=0.279	0.208
ABC Stereotypical Behaviour/MTS Challenging Behaviour	-0.078	0.591	0.274	0.217
ABC/Total Engagement	-0.162	0.262	Rho= -0.548	0.008†
ABC/Assistance	-0.033	0.820	-0.015	0.946
ABC/GCPLA Range	-0.242	0.091	-0.135	0.549
ABC/GCPLA Busy	-0.431	0.002†	1	0.60
ABC/Practice Leadership	-0.332	0.019	-0.489	0.021
ABC/Staff Knowledge	-0.167	0.246	0.114	0.614
ABS/Total Engagement	0.335	0.018	0.273	0.220
ABS <151/ Total Engagement	0.199	0.300	0.410	0.164
ABS/ASM	0.326	0.021	0.143	0.524
ABS/Assistance	0.173	0.230	0.130	0.564
ABS/GCPLA Range	0.208	0.146	0.123	0.586
Assistance/Engagement	0.327	0.020	0.493	0.020
Engagement/Practice Leadership	0.067	0.645	0.475	0.025
GCPLA Busy/ Practice Leadership	0.247	0.083	-0.034	0.881
ASM/Engagement	0.515	p<0.001	0.651	0.001†

† Significant at $p \leq 0.01$

8.3 Service User Results at T2: Impact of Training

8.3.1 Introduction

This section of the service user results chapter considers the initial impact of training, that is, the difference from T1-T2. The various tests carried out are to establish if there are differences between the experimental and control groups from pre to post training. In the main this will be explored via ANOVA to consider any significant interaction between time and group; the other results from the ANOVA are presented for additional information in appendix six. For non-parametric data, the change in scores from T1-T2 will be calculated and then tests done on these change scores in order to examine any significant difference between experimental and control groups.

8.3.2 Cohort Comparison

As the ABC is the main outcome measure, the experimental groups of each cohort were compared at T2 in order to establish if there was a significant difference between them in terms of change in ABC total score. There was no significant difference ($t=0.137$; $p=0.891$; $df=48$; $d=0.04$, a negligible effect size) so the cohorts were combined and treated as one group for all other analysis.

8.3.3 Aberrant Behaviour Checklist

8.3.3.1 Descriptive Statistics

Table 8.5 shows the mean and standard deviation for experimental and control groups at T1 and T2. Figures 19 and 20 show the graphs of these means for ABC total and for severe behaviours.

Table 8.5: Descriptive Statistics for ABC T1 and T2

ABC	Experimental		Control	
	T1	T2	T1	T2
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Total	57.44 (22.46)	33.38 (24.31)	35.23 (25.70)	32.64 (23.89)
Severe	6.04 (5.73)	1.78 (3.22)	3.09 (6.36)	2.77 (4.80)
Irritability	19.22 (7.88)	11.72 (8.63)	13.68 (9.31)	12.36 (8.94)
Lethargy	14.34 (8.55)	8.78 (7.69)	6.73 (5.82)	7.14 (6.68)
Stereotypical	4.18 (3.85)	2.98 (4.61)	2.77 (4.70)	3.09 (4.61)
Hyperactivity	15.52 (8.40)	8.48 (7.04)	9.36 (8.11)	7.36 (6.50)
Inappropriate Speech	4.04 (3.43)	2.5 (2.94)	2.23 (3.18)	2.18 (2.34)

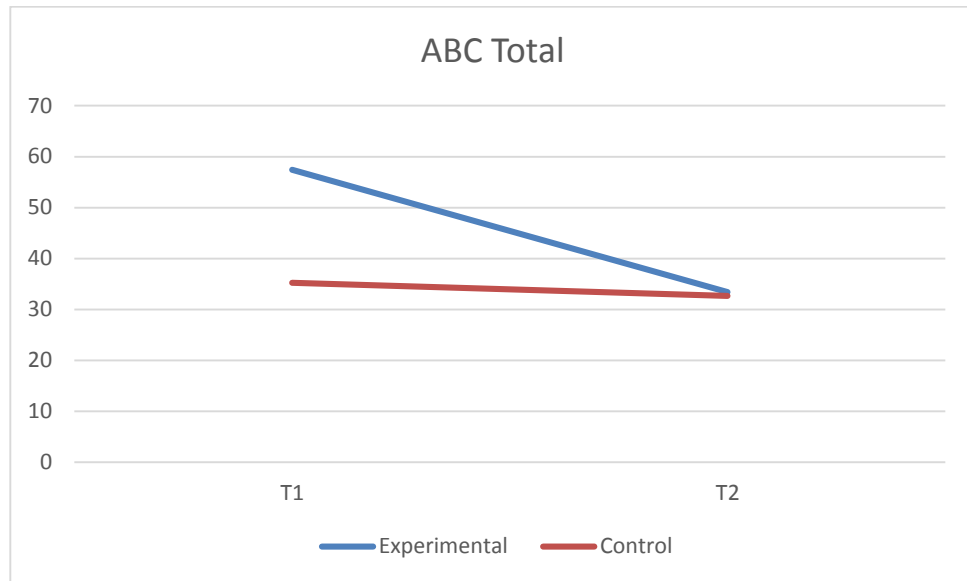


Figure 19. Mean ABC Total at T1 and T2

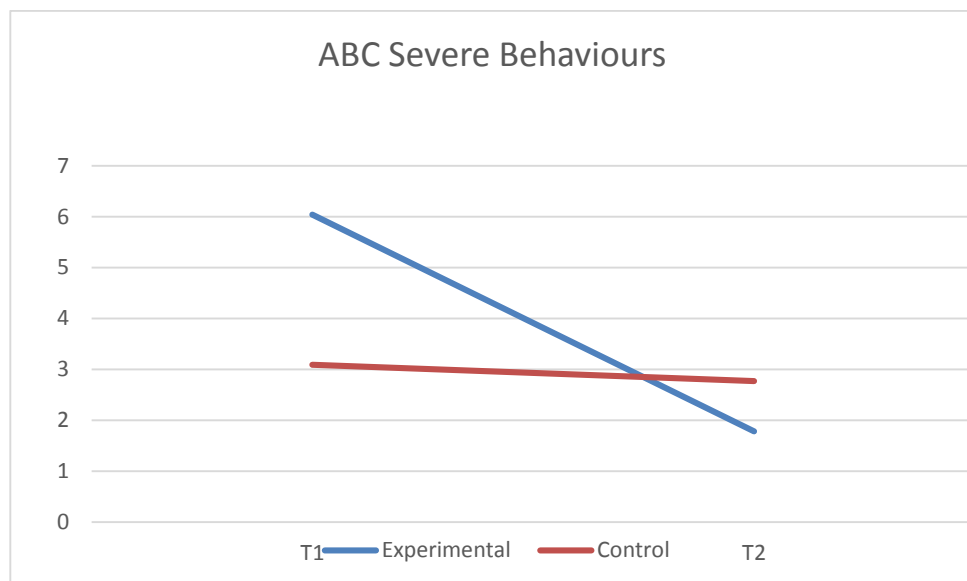


Figure 20. Mean ABC Severe Behaviours at T1 and T2

8.3.3.2 Significance Testing

In order to test differences in ABC total, a mixed multi-factorial ANOVA was used. Since there were 2 comparisons on the ABC, then the Bonferroni significance is $p < 0.025$. There was a significant interaction between time and group for ABC total ($F = 16.837$; $p < 0.001$; $df = 1$; partial eta squared = 0.201, a large effect size). As can be seen from appendix six, Table 13.19, both time and group were also significant. For ABC severe behaviours, a non-parametric test was carried

out on the change in scores between T1-T2 and this demonstrated a significant difference (U=271; p=0.001; r=0.407, a medium effect size).

8.3.3.3 Sensitivity Analysis

Seven additional analyses were carried out in addition to the main analysis for ABC total and three additional analyses for severe behaviours. As part of these additional analyses change in scores from T1-T2 were calculated for ABC total and tests were also done on these; this was to attempt to compensate to some extent for the fact that the control and experimental groups were significantly different at T1 (this test had already been done as the main analysis on severe behaviours as data did not meet parametric assumptions). These additional analyses showed that the results did not change regardless of method of analysis. For ABC total, whether analysed on an ITT basis, or on a PP basis, and whether outliers are included or not, there is a significant interaction between time and group for ABC total. The effect size was greatest with the PP analysis. For severe behaviours, the difference is also significant, regardless of how the analysis is done and effect size is also greater for the PP analysis. This information is summarised in Table 8.6.

Table 8.6: Sensitivity Analysis ABC Total and Severe Behaviours T1-T2

ABC Total	F/U	p	df/n	ES
ITT, outliers in	F=10.799	0.002†	1	0.134**
ITT, change score, outliers in	U=276	0.001†	n=72	0.40**
ITT, change score, outliers out	U=186.5	0.001†	n=67	0.42**
PP, outliers out	F=26.184	p<0.001	1	0.315***
PP, outliers in	F=17.035	p<0.001	1	0.221***
PP, change score, outliers in	U=181.5	p<0.001	n=62	0.48**
PP, change score, outliers out	U=113	p<0.001	n=57	0.52***
Severe Behaviours	Z/U	p	df	ES
ITT multiple testing experimental	Z=4.159	p<0.0010.857	n=50	0.59***
control	Z=0.180		n=22	0.04
PP, outliers in	U=189.5	p<0.001	n=62	0.48***
PP multiple testing experimental	Z=4.564	p<0.001	n=40	0.722***

* Small effect size; **Medium effect size; ***Large effect size; † Significant at p≤0.025

8.3.3.4 Individual Change

In order to look at individual change, i.e. not just change within the group as a whole, the change in ABC total score from T1-T2 was calculated for each participant in both the experimental and control groups. This is shown in Figures 21 and 22. This shows that overall positive change in the experimental group is borne out for the majority of participants (84%); in the control group, positive change is only shown in 55% of the group.

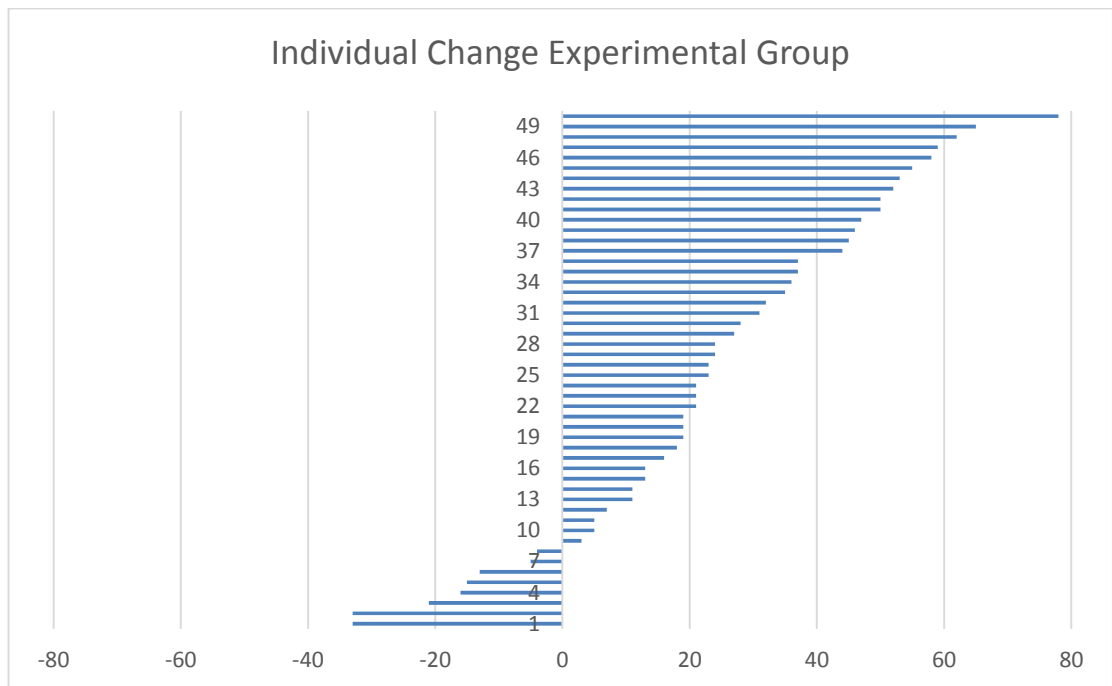


Figure 21. Individual Change in ABC Total Scores T1-T2 Experimental

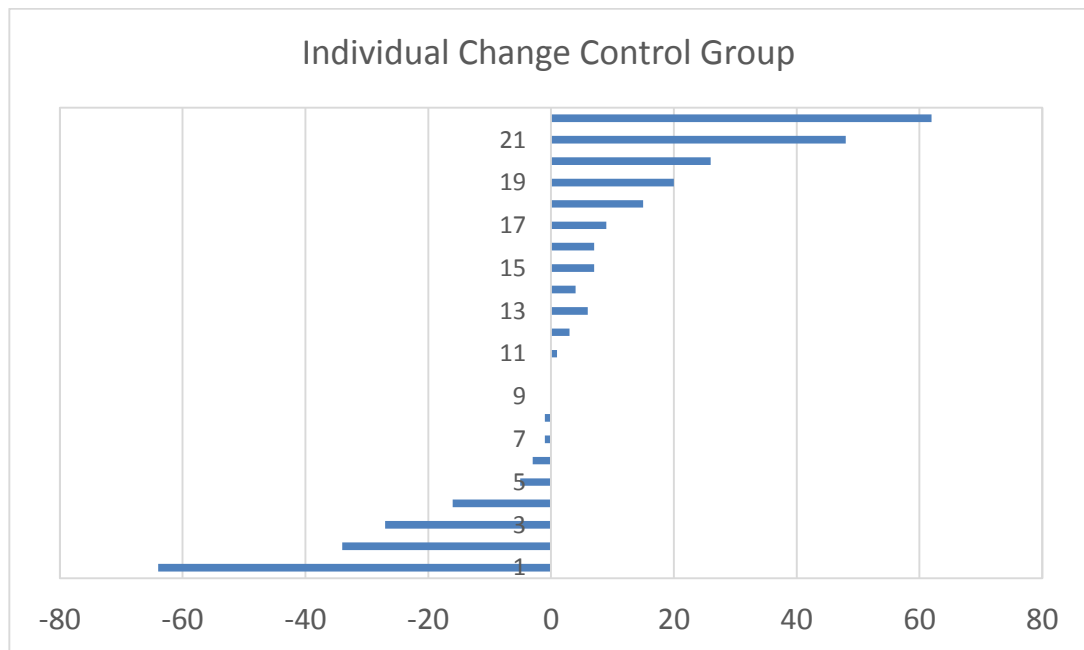


Figure 22. Individual Change in ABC Total Scores T1-T2 Control

8.3.3.5 Reduction to the Baseline

Another way of considering change in ABC score is to look at reduction to the baseline. This can be considered in three categories (McClellan et al, 2005): reduction to 70% or more of the baseline, reduction to 30-69% of the baseline, and reduction to less than 30% of the baseline. Percentages in each category from T1-T2 are shown for the experimental and control groups in Table 8.7. The average reduction was to 60% of the baseline for the experimental group and to 107% of the baseline for the control group.

Table 8.7: Reduction to the Baseline in ABC Total for Experimental and Control T1-T2

ABC Total	Experimental Percent (n)	Control Percent (n)
Reduction to 70% or more of baseline	30 (15)	64 (14)
Reduction to 30-69% of baseline	52 (26)	23 (5)
Reduction to less than 30% of baseline	18(9)	13 (3)

8.3.3.6 Reliable Change

Using the reliable change formula

$$RC = x1 - x2 \div Sdiff$$

$$(Sdiff = \sqrt{2(SE)^2} \quad \text{and} \quad SE = s1\sqrt{1 - rxx})$$

and the fact that a score of at least 1.96 is required in order to indicate that a reliable change has taken place; then using the formula below and the information in Table 8.8, it was calculated that the difference between T1 and T2 scores ($x1 - x2$) had to be at least 27.36.

$$x1 - x2 \div 13.96 = 1.96$$

The individual change graphs were therefore redrawn, with a red line showing the score of 27.36 which would indicate a reliable change. Results are shown in Figures 23 and 24. This shows that 21 participants (42%) in the experimental group achieved a positive reliable change in ABC total score; two (4%) also had a negative reliable change. In the control group, two participants (9%) achieved a positive reliable change, but two (9%) also had a negative reliable change.

Table 8.8: Reliable Change Data

Symbol	Definition	Value
$s1$	Standard deviation of control and experimental groups at T1	25.49
rxx	Test-retest reliability of ABC	0.85*
SE	$s1\sqrt{1 - rxx}$	9.87
$Sdiff$	$\sqrt{2(SE)^2}$	13.96

* 0.85 (taken from the average reported test-retest reliability in Lehotkay et al (2015), Ono et al (1996) and Zeilinger et al, (2011))

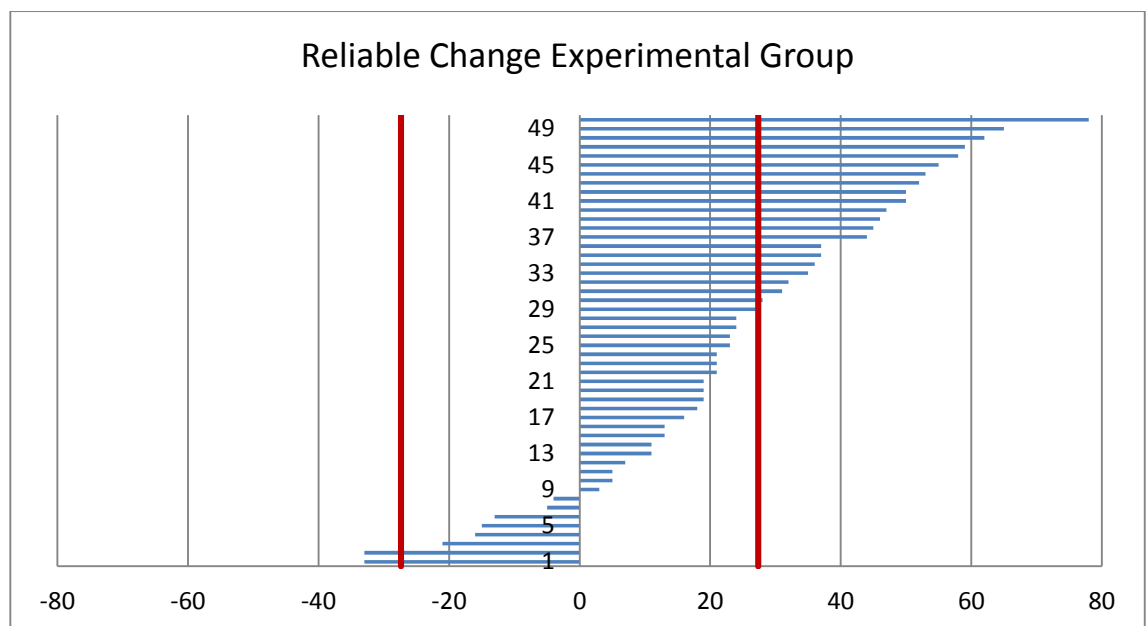


Figure 23. Reliable Change in ABC Total Scores T1-T2 Experimental

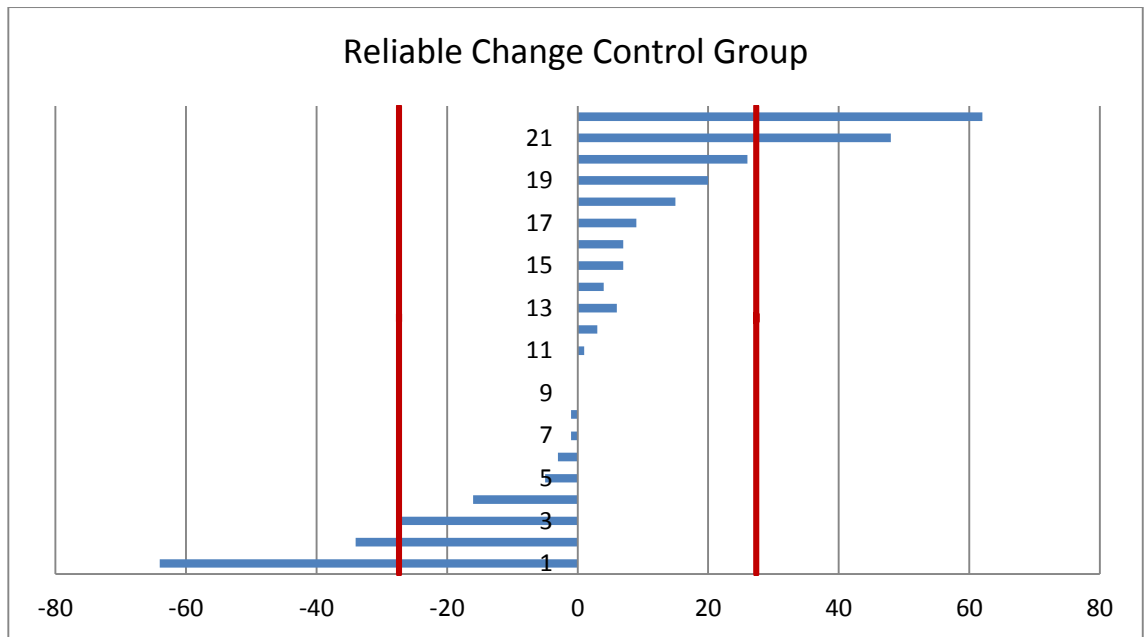


Figure 24. Reliable Change in ABC Total Scores T1-T2 Control

8.3.3.7 Differences between Trainers

As the experimental group were trained by six different trainers, it was necessary to check for any difference between these groups in terms of change in total ABC score. The descriptive statistics for each group are in Table 8.9. In order to check for any differences between trainers a one-way ANOVA was carried out on the change in ABC total scores from T1-T2; this found that there was no significant interaction between trainer and change in ABC total score ($F=1.07$; $p=0.990$; $df=49$, $\eta^2=0.012$, a small effect size).

Table 8.9: Descriptive Statistics for ABC Difference for Each Trainer

Trainers	Difference ABC T1-T2:Mean (SD)
1	25.09 (28.99)
2	21.00 (29.43)
3	29.29 (26.20)
4	21.00 (18.79)
5	25.00 (28.61)
6	23.50 (27.82)

8.3.4 Behaviour Recording Forms

8.3.4.1 Descriptive Statistics

Table 8.10 shows the descriptive statistics for experimental and control groups at T1 and T2 for both frequency and severity, taken from the Behaviour Recording Forms (BRF). Figures 25 and 26 show the graphs of the means for frequency and severity.

Table 8.10: Descriptive Statistics for BRF T1 and T2

	Experimental		Control	
	T1	T2	T1	T2
Frequency	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
	10.40 (8.92)	5.20 (6.62)	7.36 (9.72)	6.82 (9.96)
Severity	6.79 (1.94)	4.44 (3.46)	6.73 (1.81)	4.48 (3.02)

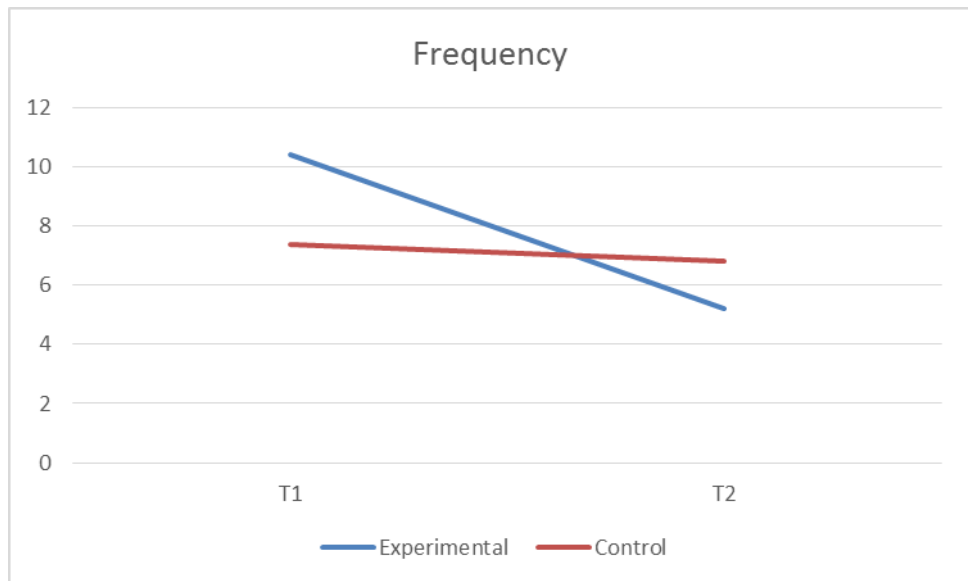


Figure 25. Mean Frequency at T1 and T2

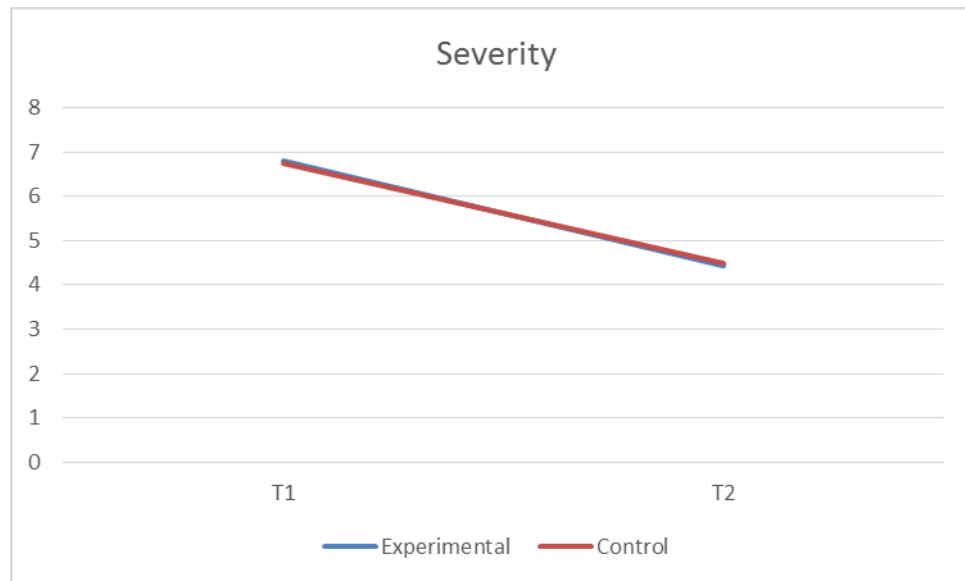


Figure 26. Mean Severity at T1 and T2

8.3.4.2 Significance Testing

For frequency, an unrelated t-test was carried out on the change in frequency from T1-T2 (change scores data met parametric assumptions, although actual scores did not) and this demonstrated a significant difference between experimental and control groups ($t=4.851$; $p=0.00$; $df=60.961$; $d=1.24$). For severity a non-parametric test was carried out on change scores. This showed no significant difference between control and experimental groups ($U=496.5$; $p=0.512$; $r=0.07$, a negligible effect size).

8.3.4.3 Sensitivity Analysis

Seven additional analyses were carried out in addition to the main analysis for both frequency and severity; these show that with the exception of the ITT ANOVA the results for frequency did not change dependent on method of analysis. However for severity the results do change dependent on method of analysis. The results from either the ANOVA or using the change in scores from T1-T2 were the same; however, using multiple testing, that is, if individual tests were done on both the control and experimental groups, then there was a significant decrease in severity for both groups. The results are described in Table 8.11.

Table 8.11: Sensitivity Analysis BRF Frequency and Severity T1-T2

Frequency	Test	p	df/n	ES
ITT, outliers in	t=3.398	0.001†	67.359	0.83***
ITT, multiple testing experimental	t=4.268	p<0.001	49	1.22***
control	t=0.870	0.394	21	0.38*
ITT, ANOVA, outliers out	F=4.178	0.045	1	0.060**
PP, outliers out	t=4.836	p<0.001	54.360	1.31***
PP, outliers in	t=4.575	p<0.001	56.638	1.22***
PP, multiple testing: experimental	Z=4.890	p<0.001	n=40	0.77***
PP, ANOVA, outliers out	F=12.066	p<0.001	1	0.191
Severity				
ITT, outliers out	U=446.5	0.322	71	0.12
ITT, multiple testing experimental	Z=4.178	p<0.001	n=40	0.59***
control	Z=2.486	0.013†	n=22	0.51***
ITT, ANOVA, outliers out	F=0.248	0.620	1	0.004
PP, outliers out	U=334.5	0.194	n=61	0.17
PP, outliers in	U=374.5	0.334	n=62	0.12
PP, multiple testing experimental	Z=4.101	p<0.001	n=40	0.65***
PP, ANOVA, outliers out	F=0.752	0.389	1	0.013

* Small effect size; ** Medium effect size; † Significant at $p \leq 0.025$

8.3.4.4 Reduction to the Baseline

Another way of considering change in frequency and severity is to look at reduction to the baseline. This can be considered in three categories, reduction to 70% or more of the baseline, reduction to 30-69% of the baseline, and reduction to less than 30% of the baseline. Percentages in each category from T1-T2 are shown in Table 8.12 for the experimental and control groups for both frequency and severity. The average reduction for frequency was to 59% of the baseline

for the experimental group and to 107% of the baseline for the control group. The average reduction for severity was to 66% of the baseline for the experimental group and to 71% of the baseline for the control group.

Table 8.12: Reduction to the Baseline in BRF for Experimental and Control T1-T2

BRF	Experimental Percent (<i>n</i>)		Control Percent (<i>n</i>)	
	Frequency	Severity	Frequency	Severity
Reduction to 70% or more	26	58	64	73
Reduction to 30-69%	22	10	9	0
Reduction to less than 30%	52	32	27	27

8.3.5 Momentary Time Sampling (MTS)

8.3.5.1 Descriptive Statistics

Table 8.13 shows the mean and standard deviation for experimental and control groups at T1 and T2 for MTS behaviours. Figures 27 - 29 show the graphs of the means for total engagement (the sum of social, domestic, personal and other engagement), disengagement and challenging behaviour.

Table 8.13: Service Users' MTS Descriptive Statistics T1 and T2

MTS	Experimental		Control	
	T1	T2	T1	T2
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Total Engagement	70.52 (24.93)	76.02 (26.561)	86.38 (22.90)	81.50 (28.76)
Social Engagement	26.76 (20.72)	27.46 (19.08)	28.49 (21.68)	26.07 (25.96)
Domestic Engagement	13.69 (17.07)	11.09 (12.81)	19.74 (23.81)	18.45 (21.69)
Personal Engagement	11.65 (11.62)	14.56 (14.34)	14.58 (12.86)	15.34 (14.65)
Other Engagement	18.43 (22.96)	22.92 (21.27)	23.56 (24.70)	21.64 (30.00)
Disengaged	26.32 (24.81)	21.11 (24.73)	12.95 (22.47)	18.32 (28.79)
Challenging Behaviour	3.16 (6.60)	2.86 (6.36)	0.68 (1.55)	0.18 (0.50)

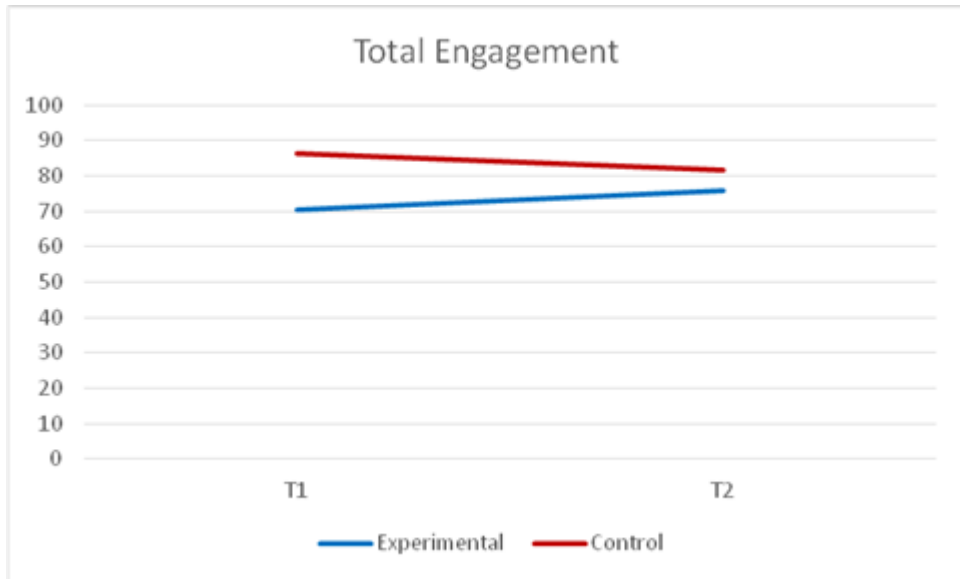


Figure 27. Mean Percentage Total Engagement T1 and T2

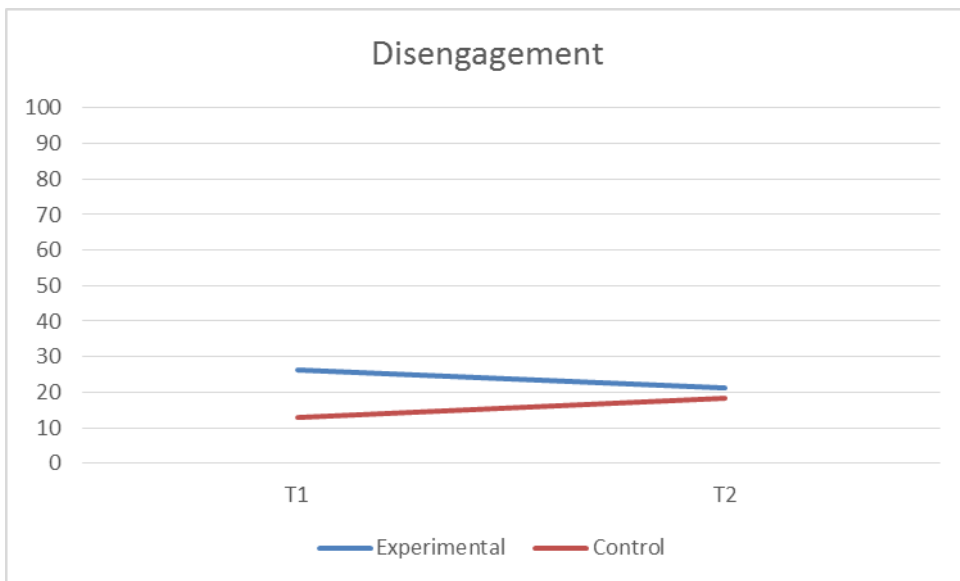


Figure 28. Mean Percentage Disengagement T1 and T2

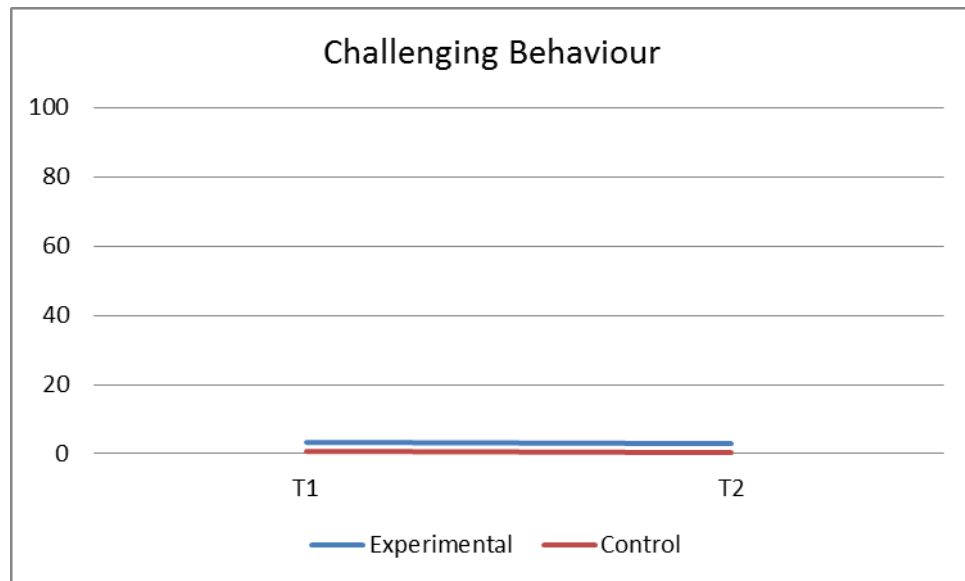


Figure 29. Mean Percentage Challenging Behaviour T1 and T2

8.3.5.2 Significance Testing

Non-parametric tests were carried out on the change in scores between T1-T2 and these demonstrated no significant difference between experimental and control groups for total engagement ($U=426$; $p=0.129$; $df=70$; $r=0.18$, a small effect size), for disengagement ($U=416$; $p=0.101$; $df=70$; $r=0.19$, a small effect size), or for challenging behaviour ($U=537$; $p=0.862$; $df=70$; $r=0.02$, negligible effect size). As change in scores from T1-T2 were used (due to data not meeting parametric assumptions) this also to some extent compensated for the fact that the control and experimental groups were significantly different at T1 for engagement and disengagement.

In order to explore the different types of engagement, tests were also done for changes in social and non-social engagement (a combination of the three non-social engagement codes: domestic, personal and other). These showed that there was no significant difference between the experimental and control groups in the change in social engagement from T1-T2 ($U=467$; $P=0.310$; $df=70$; $r=0.12$, a small effect size); and there was no significant interaction between time and group for non-social engagement T1-T2 ($F=0.921$; $p=0.341$; $df=1$; partial eta squared=0.013, a small effect size). As these were not major behaviours of interest, no further analysis was done on these.

8.3.5.3 Sensitivity Analysis

Six additional analyses were carried out in addition to the main analysis for total engagement, five additional analyses for disengagement and three additional analyses for challenging behaviour (some tests were not done if there were too many outliers to make the tests viable if they were removed). These analyses show that for total engagement and disengagement the results

change dependent on method of analysis. If analysed on an ITT basis, there is no significant difference between the control and experimental groups in terms of the change in score from T1-T2. However, if analysed on a PP basis using ANOVA there is a significant interaction between time and group for engagement and disengagement. Also, testing the experimental data separately demonstrates a significant increase in engagement and decrease in disengagement from T1-T2 for the experimental group. For challenging behaviour, there is no difference in significance, regardless of how the data are analysed. The results are summarised in Table 8.14.

Table 8.14: Services Users' MTS Sensitivity Analysis T1-T2

MTS Total Engagement	Test	p	df/n	ES
ITT, outliers out	U=273	0.492	n=61	0.09
ITT, multiple testing				
experimental	Z=1.323	0.186	n=50	0.19*
control	Z=0.022	0.983	n=22	0.00
ITT, ANOVA	F=3.663	0.060	1	0.054*
PP, outliers in	U=261	0.016†	n=60	0.31
PP, multiple testing	Z=2.671	0.008†	n=38	0.433**
PP, ANOVA	F=8.425	0.005†	1 n=54	0.139**
MTS Disengagement	Test	p	df/n	ES
ITT, multiple testing				
experimental	Z=1.415	0.157	n=50	0.20*
control	Z=0.131	0.896	n=22	0.03
ITT, ANOVA	F=2.421	0.125	1 n=63	0.038*
PP, outliers in	U=260	0.015†	n=60	0.31**
PP, multiple testing				
experimental	Z=2.590	0.010†	n=38	0.42**
PP, ANOVA	F=8.607	0.005†	n=50	0.152
MTS Challenging Behaviour	Test	p	df/n	ES
ITT, multiple testing				
experimental	Z=0.091	0.927	n=50	0.01
control	Z=1.483	0.138	n=22	0.32**
PP, outliers in	U=384.5	0.562	n=60	0.07
PP, multiple testing				
experimental	Z=1.167	0.243	n=38	0.19*

* Small effect size; **Medium effect size; ***Large effect size; † Significant at $p \leq 0.017$

8.3.5.4 Video versus Observation

As some MTS was done live at the time for service users who did not want to be videoed, tests were carried out to check for any differences between those videoed and those observed. This showed that there was no significant difference between the videoed group and the observed group for the change in scores from T1-T2 for engagement ($U=125.5$; $p=0.524$; $r=0.08$, a negligible effect size) for disengagement ($U=112.5$; $p=0.343$; $r=0.12$, a small effect size) or for challenging behaviour ($U=133.5$; $p=0.655$; $r=0.06$, a negligible effect size). For this analysis, the imputed MTS data was not included in order to check more accurately for any differences in the groups. The descriptive statistics are shown in Table 8.15; all of the observed service users were in the experimental group.

Table 8.15: Descriptive Statistics for Observed and Videoed Groups

	Videoed ($n=61$)		Observed ($n=5$)	
	T1	T2	T1	T2
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Engagement	74.46 (26.56)	78.35 (26.39)	85.48 (13.51)	95.90 (6.58)
Disengagement	22.99 (26.09)	20.15 (25.99)	13.03 (12.58)	2.41 (3.07)
Challenging Behaviour	2.54 (6.06)	1.50 (5.14)	1.50 (3.35)	1.69 (3.78)

8.3.6 Guernsey Community Participation and Leisure Assessment (GCPLA)

8.3.6.1 Descriptive Statistics

Table 8.16 shows the mean and standard deviation for the GCPLA range score and the other four sub-scales for the experimental and control groups at T1 and T2. Figure 30 shows the graph of the means for GCPLA range at T1 and T2.

Table 8.16: GCPLA Descriptive Statistics T1 and T2

	Experimental		Control	
	T1	T2	T1	T2
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Range	20.00 (7.79)	20.08 (8.47)	20.09 (9.01)	20.23 (8.38)
Busy	12.34 (5.32)	12.36 (5.04)	12.00 (5.31)	12.82 (5.58)
Home-based	4.52 (1.86)	4.18 (1.51)	4.27 (1.91)	3.68 (1.17)
Community	15.50 (6.94)	15.92 (7.57)	15.77 (7.58)	16.55 (7.74)
Leisure	11.90 (6.21)	11.92 (5.78)	11.73 (6.37)	13.18 (5.96)

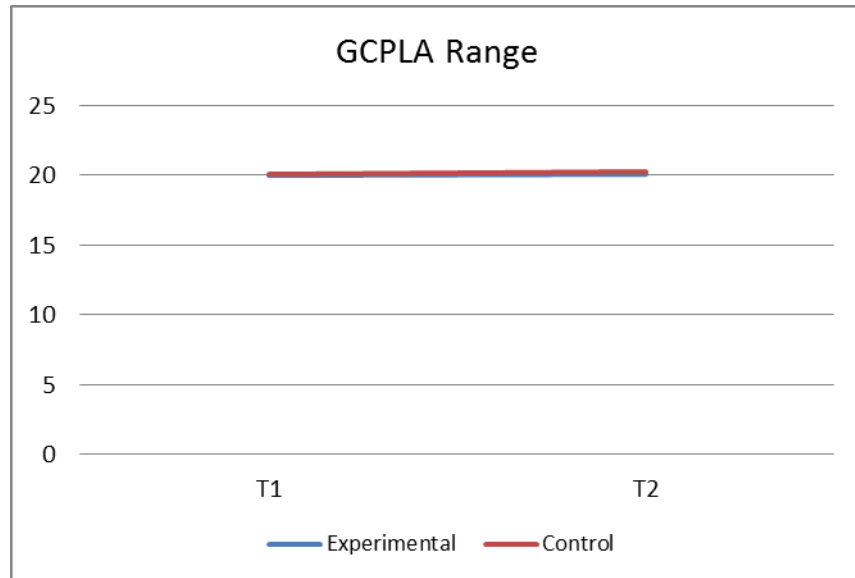


Figure 30. Mean GCPLA Range T1 and T2

8.3.6.2 Significance Testing

There was no significant interaction between time and group for range ($F=0.116$; $p=0.735$; $df=1$; partial eta squared=0.002, a negligible effect size), or for busy ($F=0.365$; $p=0.548$; $df=1$; partial eta squared=0.005, a negligible effect size). As can be seen from appendix six, Table 13.20 and 13.21, neither time nor group was significant either, for either range or busy.

8.3.6.3 Sensitivity Analysis

Three additional analyses were carried out in addition to the main analysis for both range and busy; these show that the results do not change dependent on method of analysis. This is summarised in Table 8.17.

Table 8.17: GCPLA Range and Busy Sensitivity Analysis T1-T2

GCPLA Range	F	p	df	ES
ITT, outliers in	0.00	0.992	1	0.000
PP, outliers out	0.116	0.734	1	0.002
PP, outliers in	0.019	0.891	1	0.000
GCPLA Busy	F	p	df	ES
PP, no outliers	0.060	0.807	1	0.001

8.3.7 Correlations

ABC total at T2 was significantly negatively correlated with a number of measures at T2 for the experimental group: practice leadership, Periodic Service Review (PSR), ASM, knowledge for both staff and managers, engagement and GCPLA range and busy. The results are summarised in Tables 8.18.

Table 8.18: ABC Correlations T2

ABC T2 Correlations	Experimental		Control	
	r	p	r	p
BRF	Rho=0.334	0.018	0.482	0.023
Practice Leadership	-0.567	p<0.001	-0.020	0.931
PSR	-0.474	p<0.001	-	-
Assistance	-0.315	0.026	-0.025	0.913
ASM	-0.497	p<0.001	-0.316	0.152
Managers' Knowledge	-0.471	0.001†	-0.206	0.358
Staff Knowledge	-0.441	0.001†	0.069	0.761
GCPLA Busy	-0.508	p<0.001	-0.325	0.139
GCPLA Range	-0.382	0.006†	-0.405	0.061
Engagement	-0.497	p<0.001	0.027	0.906

† Significant at $p \leq 0.01$

Other correlations were also considered to explore potential association between measures. Assistance and ASM were significantly positively correlated for both the experimental and control group, and ASM was also significantly positively correlated with both engagement and practice leadership only for the experimental group. Practice leadership was also significantly positively correlated with GCPLA range and busy, with the PSR scores and with managers' knowledge, for the experimental group only. The PSR was significantly positively correlated with managers' knowledge and with GCPLA range and busy for the experimental group only. The results are summarised in Table 8.19.

Table 8.19: Correlations T2 for all Measures

	Experimental		Control	
	r	p	r	p
Assistance/Practice Leadership	0.261	0.067	0.000	0.999
Assistance/PSR	0.213	0.137	-	-
Assistance/Engagement	0.328	0.020	0.504	0.017
Assistance/ASM	0.501	p<0.001	0.619	0.002†
Assistance/Staff Knowledge	0.140	0.334	0.065	0.774
ASM/Engagement	0.634	p<0.001	0.354	0.106
ASM/Practice Leadership	0.501	p<0.001	0.122	0.590
Engagement/Practice Leadership	0.282	0.047	0.221	0.322
Engagement/PSR	0.247	0.083	-	-
GCPLA Range/Practice Leadership	0.423	0.002†	0.443	0.039
GCPLA Busy/Practice Leadership	0.401	0.004†	0.329	0.135
PSR/Practice Leadership	0.555	p<0.001	-	-
PSR/Managers' Knowledge	Rho=0.393	0.005†	-	-
PSR/GCPLA Range	0.450	0.001†	-	-
PSR/GCPLA Busy	0.478	p<0.001	-	-
Managers' Knowledge/Practice Leadership	0.438	0.001†	-0.351	0.110

† Significant at $p \leq 0.01$

8.3.8 Summary of Service User Results at T2

8.3.8.1 Challenging Behaviour

Results from the ABC total and severe behaviours, are both significant, regardless of the method of analysis. Results from the BRF show that frequency is significant except with the ANOVA, and that severity is only significant when using multiple testing; when both experimental and

control are significant. Challenging behaviour was also analysed via MTS; however this occurred at a very low level for both the experimental and control groups, and there was no significant difference post-training. Due to the low level of occurrence during video observation and the high number of outliers and extreme scores, this is not judged to be a reliable assessment of challenging behaviours. In summary, it is judged that there can be a confident conclusion regarding significant changes in frequency and severity of service users' challenging behaviour post-training, as reflected by the ABC, the main outcome measure.

8.3.8.2 Quality of Life

Results from the MTS analysis show that total engagement is significant only when analysed using the PP approach. This approach excluded data from six imputed participants and from six service users whose managers had either left the course (three) or the organisation (three). The PP analysis demonstrated a significant change in engagement for service users post-training; however, there is less confidence in these results as they are only achieved with the PP analysis. In summary, it is judged that there can be a confident conclusion that there was no significant change in engagement from T1-T2.

GCPLA range and busy scores were not significant using any method of analysis, demonstrating that there was no significant improvement in service users' use of the community or leisure activities T1-T2.

8.4 Service User Results at T3: Maintenance of Training Effects

8.4.1 Introduction

This section of the service user results chapter considers the data gathered at follow-up, T3. The question being considered in this section is different from the previous results section which considered initial impact of training, that is, comparisons from T1-T2. This current section considers whether training effects were maintained and therefore analysis is done for T2-T3 to consider whether any changes occurring from T1-T2 were maintained; this analysis was only done for measures where there had been significant change from T1-T2 (**ABC** and **BRF**, for frequency only). Analysis was also done from T1-T3; this is to consider the training's long-term effectiveness, and was done for all measures. So the questions being explored with the analyses are different depending on the time frame; from T1-T3 a significant difference (in the expected direction) would show the training's long-term effectiveness; a non-significant difference from T2-T3 would show that any changes following training had not been lost. Throughout this section, the PP analysis is based on data from only 22 managers, due to either managers leaving the course or the service, or service users leaving the organisation and therefore the results from the

PP analysis need to be treated with some caution as they represent less than half of the experimental group.

8.4.2 Cohort Comparison

As ABC was the main outcome measure, the two experimental cohorts were compared at T3 in order to establish if there was a significant difference between them in terms of change in ABC total score from T1-T3. There was no significant difference ($t=2.003$; $p=0.051$; $df=46$; $d=0.59$, a medium effect size) so the cohorts were combined and treated as one group for all other analysis.

8.4.3 Aberrant Behaviour Checklist

8.4.3.1 Descriptive Statistics

Table 8.23 shows the mean and standard deviation for the experimental group at T1, T2 and T3 for the ITT and PP data sets, with and without outliers. Figures 31 and 32 show the graphs of these means for ABC total and for severe behaviours also at the three time points for the ITT dataset.

Table 8.23: ABC Descriptive Statistics T1-T3

ABC (n)	T1	T2	T3
	Mean (SD)	Mean (SD)	Mean (SD)
Total			
ITT, outliers in (50)	57.44 (22.46)	33.38 (24.31)	35.55 (19.10)
PP, outliers in (22)	62.05 (23.93)	28.18 (17.79)	35.55 (19.99)
Severe Behaviours			
ITT, outliers in (50)	6.04 (5.73)	1.78 (3.22)	2.00 (3.13)
ITT, outliers out (38)	4.87 (5.17)	0.13 (0.34)	0.97 (1.92)
PP, outliers in (22)	7.55 (6.50)	0.91 (1.80)	2.00 (2.83)
PP, outliers out (19)	6.00 (5.18)	0.11 (0.32)	1.39 (2.33)
Irritability			
ITT, outliers in (50)	19.22 (7.88)	11.72 (8.63)	12.18 (7.69)
ITT, outliers out (46)	18.63 (7.72)	9.96 (6.43)	11.15 (6.65)
PP, outliers in (22)	20.00 (6.81)	8.64 (5.38)	11.14 (7.19)
PP, outliers out (21)	19.33 (6.20)	8.29 (5.24)	10.81 (7.20)
Lethargy			
ITT, outliers in (50)	14.34 (8.55)	8.78 (7.69)	8.72 (6.57)
ITT, outliers out (49)	13.86 (7.92)	8.55 (7.59)	8.47 (6.39)
PP, outliers in (22)	16.36 (9.51)	8.36 (6.56)	9.95 (6.68)
Stereotypical Behaviour			
ITT, outliers in (50)	4.18 (3.85)	2.98 (4.61)	2.90 (2.98)
ITT, outliers out (45)	3.51 (2.82)	1.53 (1.50)	2.36 (2.27)
PP, outliers in (22)	4.55 (4.59)	2.50 (3.38)	3.36 (3.44)
PP, outliers out (21)	3.81 (3.09)	1.86 (1.56)	2.81 (2.32)
Hyperactivity			
ITT, outliers in (50)	15.52 (8.40)	8.48 (7.04)	9.50 (5.95)
ITT, outliers out (46)	15.61 (8.28)	7.04 (5.25)	8.93 (5.84)
PP, outliers in (22)	16.91 (9.69)	6.86 (6.85)	8.82 (6.12)
Inappropriate Speech			
ITT, outliers in (50)	4.04 (3.43)	2.5 (2.94)	2.26 (2.16)
ITT, outliers out (45)	3.96 (3.49)	1.71 (1.79)	2.36 (2.05)
PP, outliers in (22)	4.00 (3.37)	1.82 (2.11)	2.73 (2.35)
PP, outliers out (21)	3.86 (3.38)	1.71 (2.10)	2.43 (1.94)

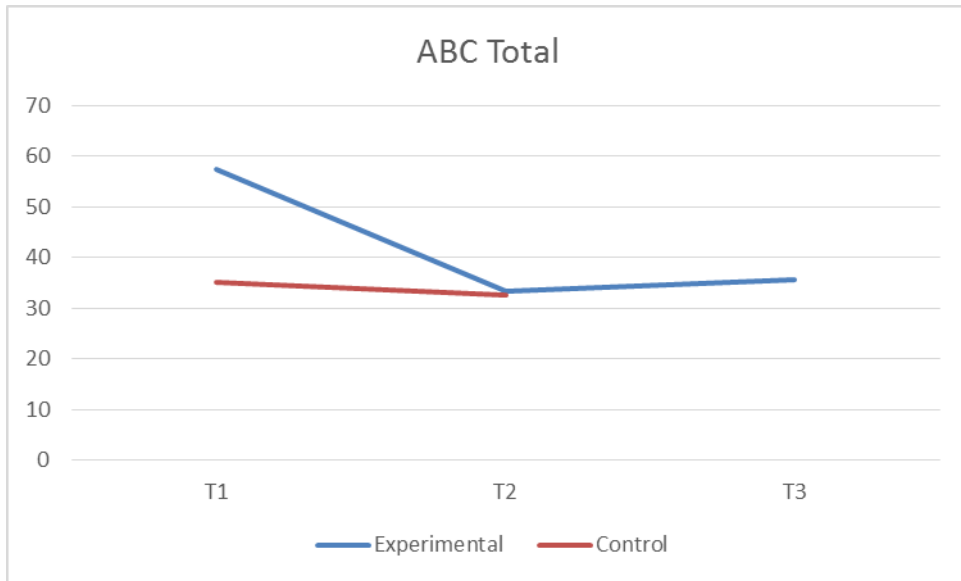


Figure 31. Mean ABC Total T1-T3

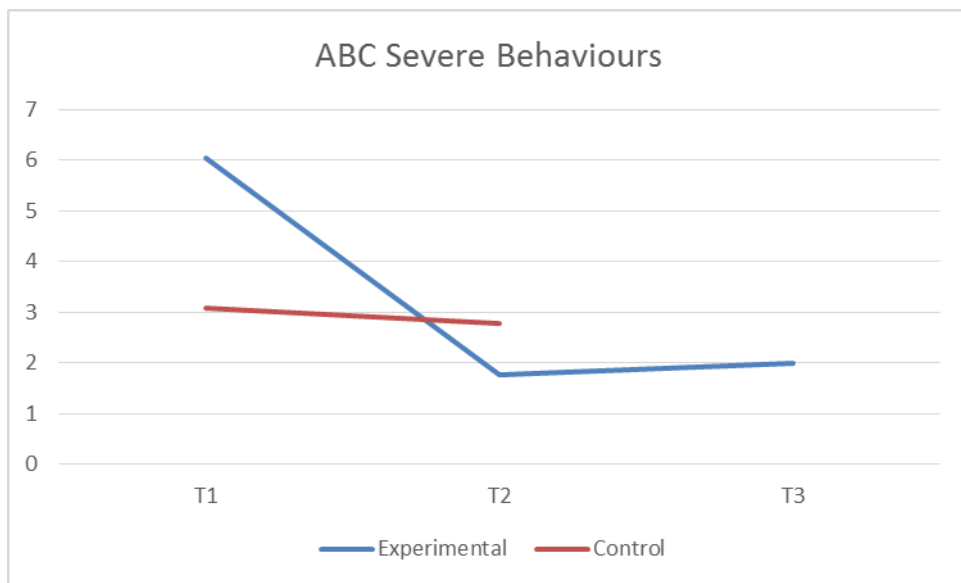


Figure 32. Mean ABC Severity T1-T3

8.4.3.2 Significance Testing

In order to test difference in ABC total, a related t-test was used. This showed a significant difference from T1-T3 ($t=6.635$; $p<0.001$; $df=49$; $d=1.9$, a large effect size), and no significant difference for T2-T3 ($t=0.723$; $p=0.473$; $df=49$; $d=0.21$, small effect size).

For severe behaviours, a Wilcoxon Signed Ranks non-parametric test was carried out as data were not normally distributed. This demonstrated a significant difference from T1-T3 ($Z=4.212$;

$p < 0.001$; $r = 0.59$, a large effect size), and no significant difference for T2-T3 ($Z = 0.663$; $p = 0.524$; $r = 0.09$, a negligible effect size).

In order to explore potential links between ABC and management, tests were carried out to check for any difference in ABC between the group whose manager had changed by T3, and the group whose manager remained the same. This showed that there was no significant difference between the groups ($t = 0.001$; $p = 0.999$; $df = 48$; $d = 0.00$, a negligible effect size), with the mean ABC total scores being the same for both groups, a mean of 35.55 for the unchanged group ($n = 22$) and a mean of 35.55 for the group whose manager had changed ($n = 28$).

8.4.3.3 Sensitivity Analysis

Two additional analyses were carried out in addition to the main analysis for ABC total at T1-T3, and T2-T3; the ITT dataset had no outliers. These showed that for ABC total the results did not change regardless of method of analysis; scores from T1-T3 changed significantly, and scores from T2-T3 did not, whether analysed on an ITT basis, or on a PP basis. The effect size in T1-T2 is similar either with ITT or PP analysis. Three additional analyses were also carried out in addition to the main analyses for severe behaviours at T1-T3, and two at T2-T3 (this had too many outliers for the analysis to be meaningful once these were removed). These showed that the results did not change regardless of method of analysis; the difference from T1-T3 is significant and the difference from T2—T3 is not significant, regardless of how the analysis is done. This information is summarised in Table 8.24 and descriptive statistics for the sensitivity analysis are above in Table 8.23.

Table 8.24: ABC Total and Severe Behaviours Sensitivity Analysis T1-T3 and T2-T3

ABC Total T1-T3	t	p	df	ES
PP (no outliers)	4.842	$p < 0.001$	21	2.11***
ABC Total T2-T3				
PP (no outliers)	1.786	0.089	21	0.78**
Severe Behaviours T1-T3	Z	p	n	ES
ITT, outliers out	4.463	$p < 0.001$	49	0.64***
PP outliers out	2.615	0.009†	18	0.67***
PP, outliers in	3.196	0.001†	22	0.68***
Severe Behaviours T2-T3				
PP outliers out	2.032	0.042	18	0.48**
PP, outliers in	1.854	0.064	22	0.39**

* Small effect size; ** Medium effect size; *** Large effect size; † Significant $p \leq 0.025$

8.4.3.4 Individual Change

In order to look at individual change, that is, not just change within the group as a whole, the change in ABC total score from T1-T3 was calculated for each participant in the experimental group. This is shown in Figure 33. This shows that overall positive change between T1-T3 in the experimental group is borne out for the majority of participants (82%).

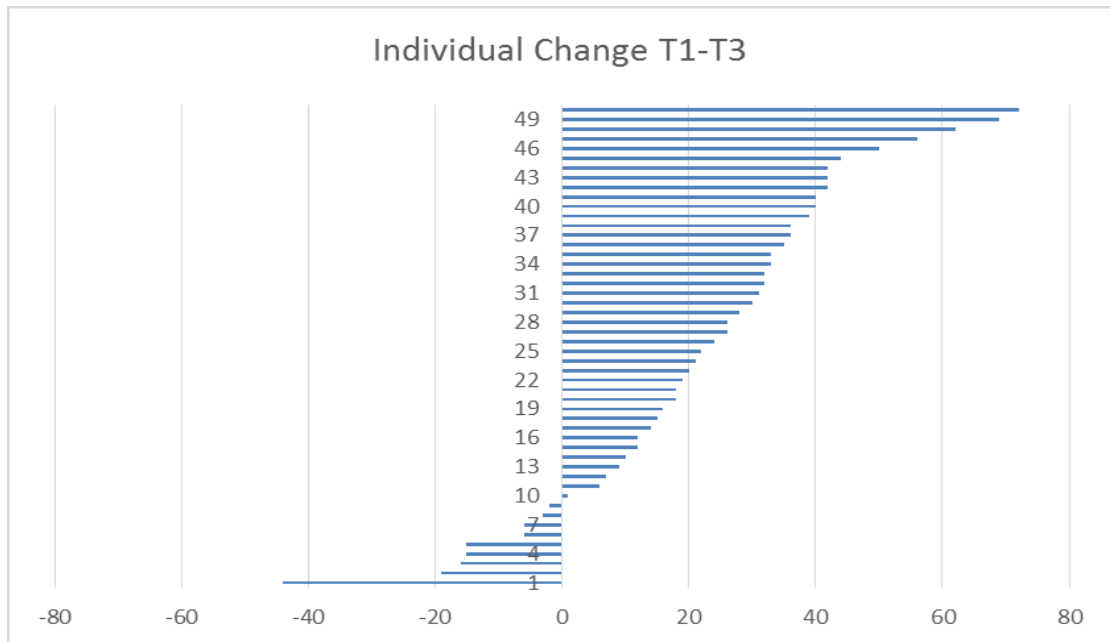


Figure 33. Individual Change in ABC Total Scores T1-T3

8.4.3.5 Reliable Change

Using the same approach to reliable change as earlier, the difference between T1 and T3 scores had to be at least 27.36 in order for this to be judged a reliable change. The individual change graphs were therefore redrawn, with a red line showing the score of 27.36 which would indicate a reliable change. Results are shown in Figure 34. This shows that 22 participants (44%) in the experimental group achieved a positive reliable change in ABC total score between T1-T3 and one (2%) had a negative reliable change.

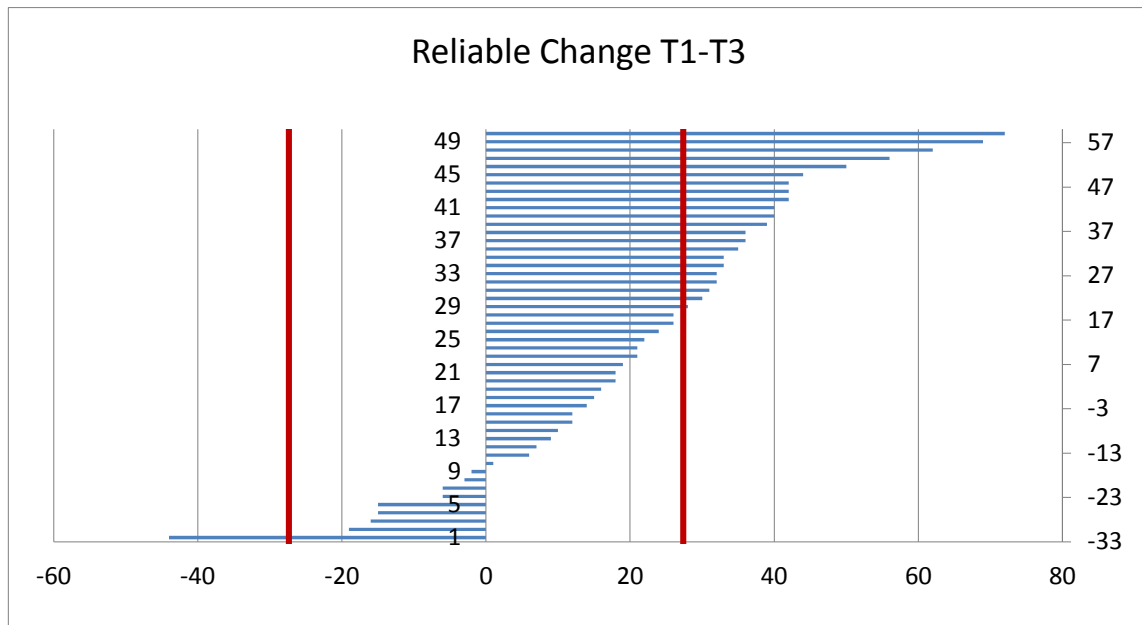


Figure 34. Reliable Change in ABC Total Scores T1-T3

8.4.3.6 Reduction to the Baseline

Another way of considering change in ABC total is to look at reduction to the baseline. This can be considered in three categories, reduction to 70% or more of the baseline, reduction to 30-69% of the baseline, and reduction to less than 30% of the baseline. Percentages in each category from T1-T3 are shown in Table 8.25 for the experimental group. The average reduction for ABC total was to 69% of the baseline and the average reduction for severe was to 33% of the baseline.

Table 8.25: Reduction to the Baseline in ABC Total and Severe Behaviours T1-T3

ABC	Experimental Percent (<i>n</i>)	
	ABC Total	Severe Behaviours
Reduction to 70% or more of the baseline	32 (16)	14 (7)
Reduction to 30-69%	50 (25)	28 (14)
Reduction to less than 30% of the baseline	18 (9)	58 (29)

8.4.4 Behaviour Recording Forms

8.4.4.1 Descriptive Statistics

Table 8.26 shows the descriptive statistics for the experimental group at T1, T2 and T3 for both frequency and severity, taken from BRF for the ITT and PP data sets, with and without outliers. Figures 35 and 36 show the graphs of the mean for frequency and severity for the ITT dataset.

Table 8.26: BRF Descriptive Statistics T1-T3

BRF (n)	T1	T2	T3
	Mean (SD)	Mean (SD)	Mean (SD)
Frequency			
ITT, outliers in (50)	10.40 (8.92)	5.20 (6.62)	5.30 (6.99)
ITT, outliers out (43)	9.33 (7.55)	3.23 (4.31)	2.93 (3.94)
PP, outliers in (22)	11.32 (10.66)	2.95 (4.59)	3.18 (5.39)
PP, outliers out (18)	7.39 (6.39)	1.50 (1.46)	1.11 (1.45)
Severity			
ITT, outliers in (50)	6.79 (1.94)	4.44 (3.46)	4.02 (3.41)
ITT, outliers out (49)	6.67 (1.77)	4.53 (3.43)	4.10 (3.39)
PP, outliers in (22)	6.61 (2.15)	3.80 (2.97)	3.23 (3.20)
PP, outliers out (20)	6.09 (1.41)	3.81 (2.89)	3.12 (3.03)

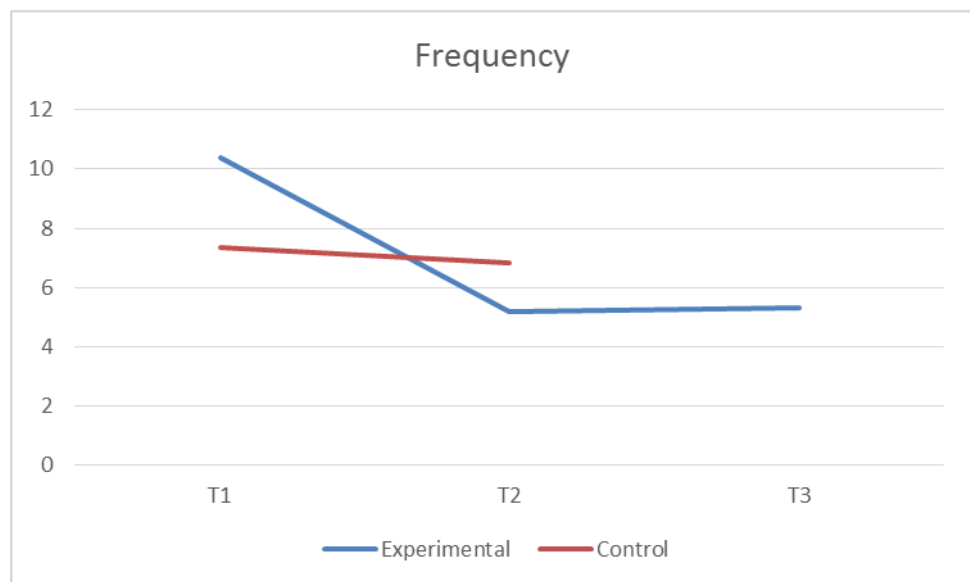


Figure 35. Mean Frequency T1-T3



Figure 36. Mean Severity T1-T3

8.4.4.2 Significance Testing

For frequency, as data was not normally distributed, a Wilcoxon Signed Ranks test was carried out on the difference in frequency from T1-T3 and this demonstrated a significant difference from T1-T3 ($Z=3.853$; $p=0.00$; $r=0.54$, a large effect size), and no significant difference from T2-T3 ($Z=0.206$; $p=0.843$; $r=0.03$, a negligible effect size).

For severity, as data was not normally distributed, a Wilcoxon Signed Ranks was carried out on the difference in severity from T1-T3 and this demonstrated a significant difference ($Z=5.098$; $p=0.00$; $r=0.72$, a large effect size).

8.4.4.3 Sensitivity Analysis

Three additional analyses were carried out in addition to the main analysis for frequency at T1-T3 and T2-T3. These show that the method of analysis does not change the results: T1-T3 continues to show significant change in frequency of challenging behaviour, and there is no significant change in frequency from T2-T3, regardless of intention to treat or per protocol analysis, and whether outliers are included or not. Effect sizes are also similar regardless of type of analysis.

Three additional analyses were carried out in addition to the main analysis for severity at T1-T3. These show that the method of analysis does not change the results and effect sizes are also similar regardless of type of analysis. The sensitivity analysis results are summarised in Table 8.27 and descriptive statistics for the sensitivity analysis are above in Table 8.26.

Table 8.27: Frequency Sensitivity Analysis T1-T3 and T2-T3 and Severity Sensitivity Analysis T1-T3

Frequency T1-T3	Z	p	n	ES
ITT, outliers out	4.870	p<0.001	43	0.74***
PP, outliers out	3.531	p<0.001	18	0.83***
PP, outliers in	3.947	p<0.001	22	0.84***
Frequency T2-T3	Z	p	n	ES
ITT, outliers out	0.541	0.603	43	0.08
PP, outliers out	1.031	0.302	18	0.24*
PP, outliers in	0.379	0.704	22	0.08
Severity T1-T3	Z	p	n	ES
ITT, outliers out	5.006	p<0.001	49	0.72***
PP, outliers out	3.825	p<0.001	20	0.86***
PP, outliers in	3.947	p<0.001	22	0.84***

* Small effect size; *** Large effect size

8.4.4.4 Reduction to the Baseline

Another way of considering change in frequency and severity is to look at reduction to the baseline. This can be considered in three categories, reduction to 70% or more of the baseline, reduction to 30-69% of the baseline, and reduction to less than 30% of the baseline. Percentages in each category from T1-T3 are shown in Table 8.28 for the experimental and control groups for both frequency and severity. The average reduction for frequency was to 67% of the baseline and the average reduction for severity was to 58% of the baseline.

Table 8.28: Reduction to the Baseline in BRF T1-T3

BRF	Percent (n)	
	Frequency	Severity
Reduction to 70% or more	28 (14)	54 (27)
Reduction to 30-69%	22 (11)	10 (5)
Reduction to less than 30%	50 (25)	36 (18)

8.4.5 Momentary Time Sampling

8.4.5.1 Descriptive Statistics

Table 8.29 shows the mean and standard deviation for the experimental group at T1, T2 and T3 for MTS behaviours for the ITT and PP data sets, with and without outliers. Figures 37-39 show the graphs of the means for total engagement, disengagement and challenging behaviour for the ITT dataset.

Table 8.29: Descriptive Statistics for MTS at T1-T3

MTS (n)	T1 Mean (SD)	T2 Mean (SD)	T3 Mean (SD)
Total Engagement			
ITT, outliers in (50)	70.52 (24.93)	76.02 (26.561)	77.08 (23.16)
ITT, outliers out (44)	72.76 (25.21)	83.36 (18.12)	82.77 (15.64)
PP, outliers in (21)	68.40 (31.23)	79.88 (26.41)	81.37 (23.37)
PP, outliers out (19)	70.56 (32.14)	84.86 (22.22)	86.05 (17.06)
Social Engagement			
ITT, outliers in (50)	26.76 (20.72)	27.46 (19.08)	16.97 (14.68)
PP, outliers in (21)	19.42 (18.31)	23.31 (14.50)	16.24 (16.86)
PP, outliers out (20)	17.11 (15.32)	22.64 (14.54)	14.91 (16.12)
Domestic Engagement			
ITT, outliers in (50)	13.69 (17.07)	11.09 (12.81)	16.44 (19.37)
ITT, outliers out (43)	8.97 (8.89)	10.43 (11.18)	10.92 (10.16)
PP, outliers in (21)	15.77 (17.92)	11.64 (15.70)	17.57 (20.42)
PP, outliers out (19)	11.48 (12.39)	8.92 (11.66)	15.37 (15.90)
Personal Engagement			
ITT, outliers in (50)	11.65 (11.62)	14.56 (14.34)	16.75 (13.11)
ITT, outliers out (42)	10.44 (9.88)	10.37 (7.15)	13.62 (9.45)
PP, outliers in (21)	11.89 (12.67)	18.20 (17.66)	12.57 (11.29)
PP, outliers out (16)	7.54 (8.04)	9.87 (6.97)	9.34 (9.76)
Other Engagement			
ITT, outliers in (50)	18.43 (22.96)	22.92 (21.27)	26.93 (23.76)
ITT, outliers out (44)	15.15 (19.87)	20.95 (20.06)	20.95 (16.20)
PP, outliers in (21)	21.33 (24.60)	26.73 (24.11)	34.33 (27.67)
Disengaged			
ITT, outliers in (50)	26.32 (24.81)	21.11 (24.73)	21.14 (23.29)
ITT, outliers out (43)	20.59 (19.26)	15.70 (18.20)	16.86 (18.15)
PP, outliers in (21)	29.43 (31.41)	19.45 (25.09)	18.82 (23.01)
PP, outliers out (18)	24.66 (31.12)	10.98 (13.22)	11.36 (13.61)
Challenging			
ITT, outliers in (50)	3.16 (6.60)	2.86 (6.36)	1.77 (4.81)
ITT, outliers out (37)	1.22 (2.05)	0.67 (2.44)	0.65 (0.99)
PP, outliers in (21)	2.16 (7.11)	0.67 (3.08)	0.48 (1.36)
PP, outliers out (16)	0.18 (0.52)	0	0

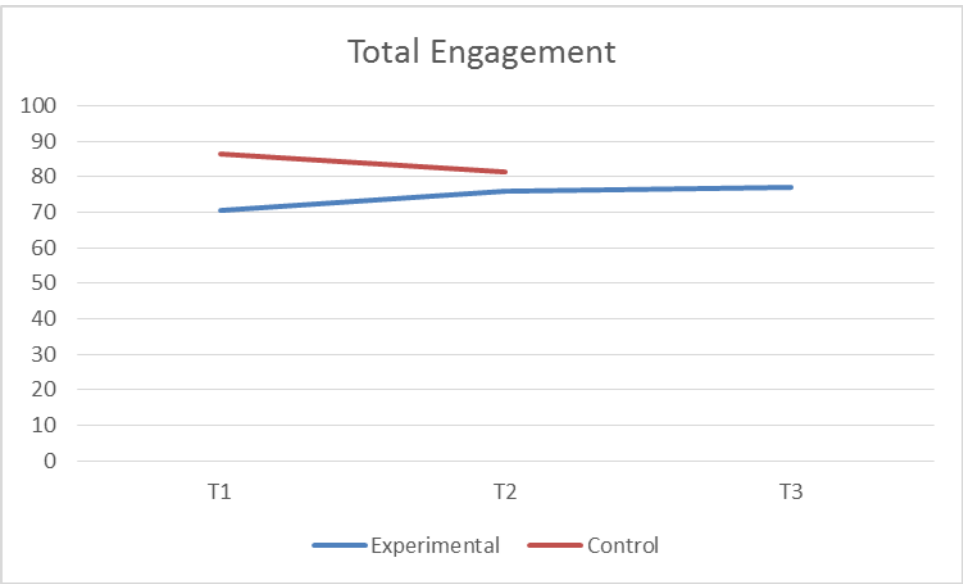


Figure 37. Mean Percentage Total Engagement T1-T3

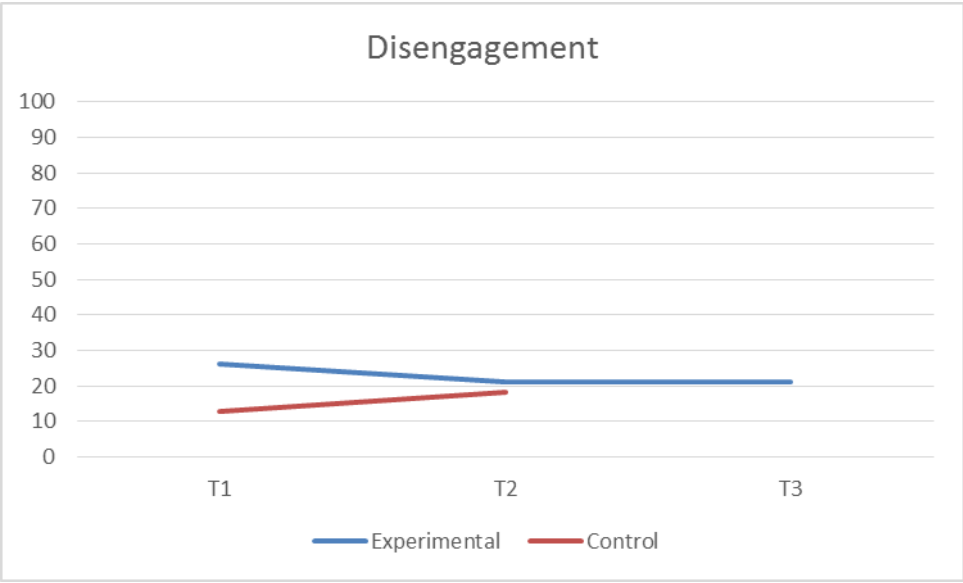


Figure 38. Mean Percentage Disengagement T1-T3

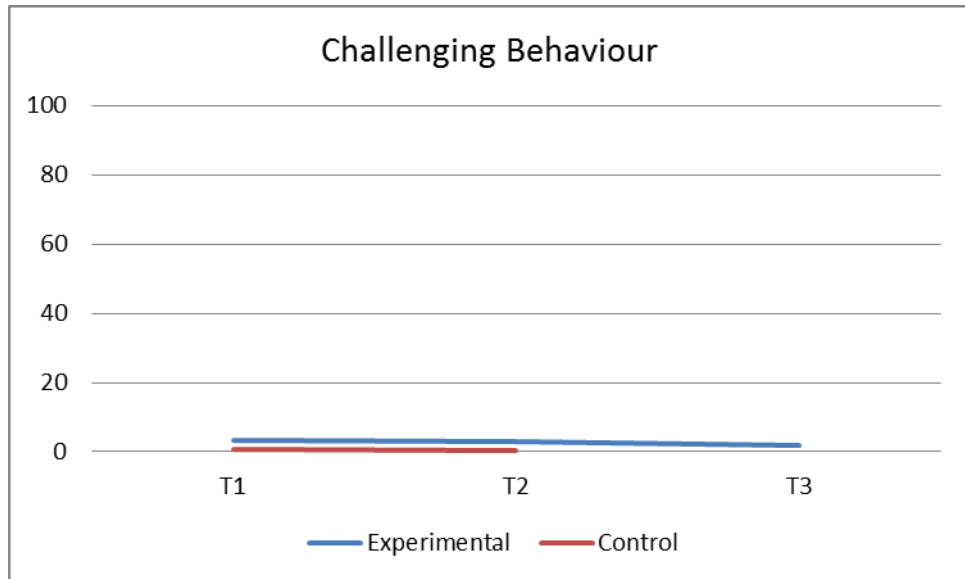


Figure 39. Mean Percentage Challenging Behaviour T1-T3

8.4.5.2 Significance Testing

Non-parametric tests were carried out on the MTS scores from T1-T3. These demonstrated no significant difference for total engagement ($Z=1.456$; $p=0.147$; $r=0.21$, a small effect size) or for disengagement ($Z=1.271$; $p=0.204$; $r=0.18$, a small effect size) or for challenging behaviour ($Z=1.882$; $p=0.060$; $r=0.27$, a small effect size).

In order to explore potential change in different types of engagement, tests were also done to establish whether either social or non-social engagement had changed significantly between T1-T3. This found that non-social engagement (non-social engagement was a combination of the three non-social engagements: domestic, personal, and other), increased significantly ($t=3.916$; $p<0.001$; $df=49$; $d=1.12$, a large effect size) from a mean at T1 of 43.76% to a mean of 60.12% at T3. Social engagement reduced significantly ($t=3.879$ $p<0.001$; $df=49$; $d=1.11$, a large effect size) from a mean of 26.76% at T1 to 16.97% at T3. As these were not major behaviours, no further analysis was done on these behaviours.

In addition, in order to explore potential links between engagement and management, tests were carried out to check for any difference in engagement between the group whose manager had changed by T3, and the group whose manager remained the same. This showed that there was no significant difference between the groups $U=2.43$; $p=0.203$; $n=50$; $r=0.18$, a small effect size), with a mean engagement for the unchanged group ($n=22$) of 80.76 and a mean engagement of 74.19 for the group whose manager had changed ($n=28$).

8.4.5.3 Sensitivity Analysis

Three additional analyses were carried out in addition to the main analysis for total engagement, disengagement, and challenging behaviour T1-T3. These analyses show that there is a significant difference for total engagement if analysed on an ITT basis with outliers removed; there is no significant difference for any of the other sensitivity analyses. The results are summarised in Table 8.30 and descriptive statistics for the sensitivity analysis are above in Table 8.29. As noted above, the PP analysis has much reduced numbers; in addition to the reductions noted earlier, there is one further service user who refused to be videoed/observed and therefore only 21 are PP for the observational data.

Table 8.30: MTS Sensitivity Analysis T1-T3

MTS Engagement T1-T3	Z/t	p	n	ES
ITT, outliers out	2.486	0.012†	46	0.37**
PP, outliers in	t=1.808	0.086	22	0.81***
PP, outliers out	t=2.190	0.041	21	1.00***
MTS Disengagement T1-T3	Z	p	n	ES
ITT, outliers out	Z=1.924	0.054	44	0.29*
PP, outliers in	Z=1.489	0.136	21	0.33**
PP, outliers out	Z=1.938	0.053	20	0.43**
Challenging Behaviour T1-T3	Z	p	n	ES
ITT, outliers out	2.062	0.038	42	0.32**
PP, outliers in	1.014	0.375	21	0.22*
PP, outliers out	1.342	0.500	16	0.34**

*Small effect size; ** Medium effect size; ***Large effect size; † Significant $p \leq 0.017$

8.4.6 Guernsey Community Participation and Leisure Assessment (GCPLA)

8.4.6.1 Descriptive Statistics

Table 8.31 shows the mean and standard deviation for the GCPLA range score and the other four sub-scales for the experimental group at T1, T2 and T3 for the ITT and PP data sets, with and without outliers. Figures 40 and 41 show the graphs of the means for GCPLA range and busy at the same time points for the ITT dataset.

Table 8.31: GCPLA Descriptive Statistics T1-T3

GCPLA (<i>n</i>)	T1	T2	T3
	Mean (SD)	Mean (SD)	Mean (SD)
Range			
ITT, outliers in (50)	20.00 (7.79)	20.08 (8.47)	20.59 (8.01)
ITT, outliers out (49)	20.06 (7.86)	20.24 (8.48)	20.79 (7.98)
PP, (no outliers) (22)	20.23 (7.54)	21.45 (7.38)	21.32 (8.05)
Busy			
ITT, outliers in (50)	12.34 (5.32)	12.36 (5.04)	13.61 (5.09)
PP, outliers in (22)	12.18 (5.31)	12.82 (4.33)	14.14 (4.49)
PP, outliers out (20)	10.90 (3.48)	12.40 (4.00)	13.65 (4.31)
Home-based			
ITT, outliers in (50)	4.52 (1.86)	4.18 (1.51)	4.52 (1.54)
ITT, outliers out (49)	4.61 (1.78)	4.27 (1.40)	4.61 (1.41)
PP (no outliers) (22)	4.73 (1.39)	4.59 (1.26)	4.86 (1.32)
Community			
ITT, outliers in (50)	15.50 (6.94)	15.92 (7.57)	16.07 (6.97)
ITT, outliers out (49)	15.12 (6.47)	15.65 (7.40)	15.93 (6.97)
PP (no outliers) (22)	15.45 (6.95)	16.86 (6.79)	16.45 (7.15)
Leisure			
ITT, outliers in (50)	11.90 (6.21)	11.92 (5.78)	12.35 (6.26)
ITT, outliers out (49)	11.51 (5.62)	11.76 (5.72)	12.13 (6.13)
PP (no outliers) (22)	12.59 (6.00)	12.55 (5.12)	13.00 (5.89)

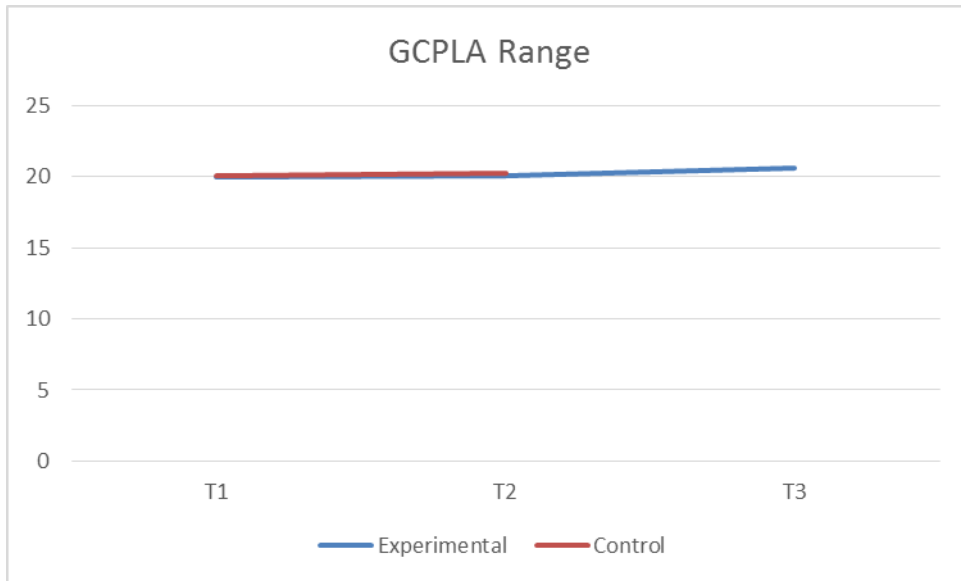


Figure 40. Mean GCPLA Range T1-T3

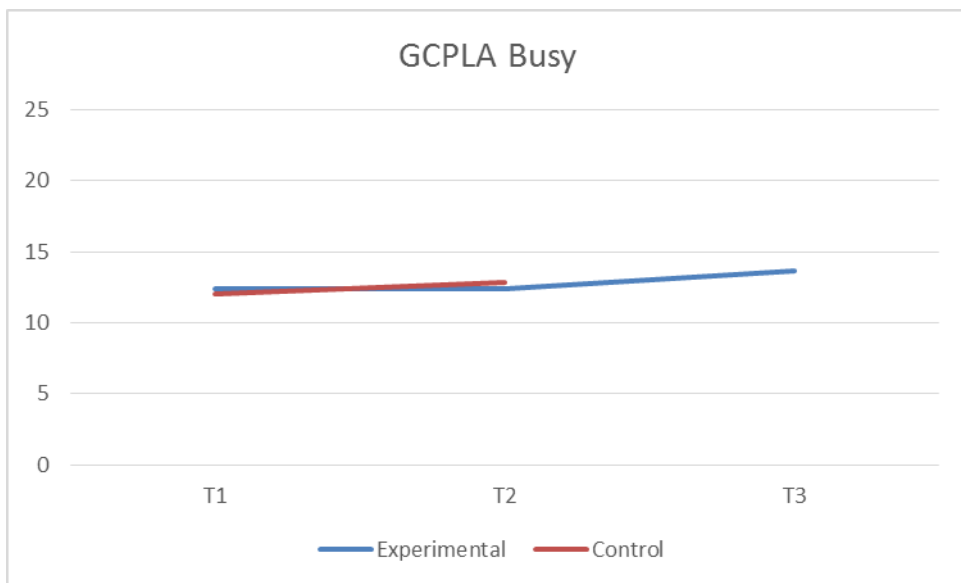


Figure 41. Mean GCPLA Busy T1-T3

8.4.6.2 Significance Testing

Tests were carried out on the GCPLA range and busy scores from T1-T3. These demonstrated no significant difference for range ($t=1.145$; $p=0.258$; $df=48$; $d=0.33$, a small effect size) or for busy ($t=1.868$; $p=0.068$; $df=49$; $d=0.53$, a medium effect size).

In order to explore potential links between GCPLA range and management, tests were carried out to check for any difference in range scores between the group whose manager had changed by T3, and the group whose manager remained the same. This showed that there was

no significant difference between the groups ($t=0.565$; $p=0.575$; $df=48$; $d=0.16$, a negligible effect size), with a mean range score for the unchanged group ($n=22$) of 21.32 and a mean range score of 20.02 for the group whose manager had changed ($n=28$).

8.4.6.3 Sensitivity Analysis

Two additional analyses were carried out in addition to the main analysis for both range and busy at T1-T3. These show that the results do not change dependent on method of analysis for range. Busy is significant using PP analysis with outliers out; however, this is based on only 22 participants out of the group of 50, so must be treated with caution. This information is summarised in Table 8.32 and descriptive statistics for the sensitivity analysis are above in Table 8.31.

Table 8.32: GCPLA Range and Busy Sensitivity Analysis T1-T3

GCPLA Range T1-T3	t	p	df	ES
ITT, outliers in	0.665	0.509	49	0.19
PP, (no outliers)	0.953	0.352	21	0.42*
GCPLA Busy T1-T3	t	p	df	ES
PP, outliers out	2.926	0.009†	19	1.34***
PP, outliers in	1.892	0.072	21	0.83***

* Small effect size; ** Medium effect size; ***Large effect size; † Significant $p \leq 0.025$

8.4.7 Correlations

A number of correlations were carried out between the main outcome measure, ABC total, and a range of other measures. This included managers, staff and other service user measures. None of these showed significant correlation with ABC total at T3. Assistance was significantly positively correlated with a range of measures: engagement, ASM, PSR, practice leadership and range. There was also significant positive correlation between practice leadership and a range of measures: managers' knowledge, PSR, GCPLA range, engagement and ASM. The results are summarised in Table 8.33.

Table 8.33: Correlations T3 for Managers', Staff, and Service User Measures

T3 Correlations	r	p
ABC/PSR	0.229	0.110
ABC/Practice Leadership	-0.202	0.160
ABC/Managers' Knowledge	-0.315	0.026
ABC/Total Engagement	Rho=0.261	0.067
ABC/Assistance	0.318	0.025
ABC/ASM	-0.32	0.023
ABC/GCPLA Range	0.273	0.055
ABC/GCPLA Busy	0.256	0.072
Assistance/Engagement	0.532	p<0.001
Assistance/ASM	0.615	p<0.001
Assistance/Staff Knowledge	0.204	0.155
Assistance/PSR	0.548	p<0.001
Assistance/Practice Leadership	0.595	p<0.001
ASM/Engagement	0.446	0.001†
ASM/Practice Leadership	0.507	p<0.001
ASM/PSR	0.476	p<0.001
Engagement/ Practice Leadership	0.626	p<0.001
GCPLA Range/Assistance	0.394	0.005†
GCPLA Range/PSR	0.384	0.006†
GCPLA Range/ Practice Leadership	0.377	0.007†
Managers' Knowledge/Practice Leadership	0.418	0.003†
PSR/Engagement	0.155	0.283
PSR/Managers' Knowledge	0.501	p<0.001
PSR/Practice Leadership	0.521	p<0.001

† Significant p≤0.01

8.5 Chapter Summary

8.5.1.1 Challenging Behaviour

Results from the ABC total and severe behaviours, are both significant at T1-T2 and at T1-T3, regardless of the method of analysis; there was no significant change from T2-T3 for either, except for severe behaviours using the per protocol analysis. Individual change calculations show that 82% of service users experienced reduced ABC total scores between T1-T3. However, using the reliable change calculations, only 44% of these can be considered a reliable change. One service user had a reliable negative change.

Results from the BRF show that frequency decreased significantly between T1-T2 and from T1-T3; severity did not decrease significantly from T1-T2, but from T1-T3 the decrease was significant.

Challenging behaviour was also analysed via MTS; however this occurred at a very low level, and although there was a small reduction from T1-T3, there was no significant difference. Due the low level of occurrence during video observation and the high number of outliers and extreme scores, this is not judged to be a reliable assessment of challenging behaviours.

In summary, it is judged that there can be a confident conclusion that the significant reduction in frequency and severity of service users' challenging behaviour following training, have been maintained at follow-up, as reflected by the ABC, the main outcome measure. The fact that this is evidenced in the intention to treat analysis gives confidence in the results.

8.5.1.2 Quality of Life

Results from the MTS analysis show there was no significant difference in engagement from T1-T2, nor from T1-T3; there were some significant changes within the sensitivity analysis, but as this is impacted by the different methods of analysis, it cannot be regarded as robust. It can therefore be concluded that there were no significant changes in total engagement from T1-T3. There were some changes within elements of engagement: non-social engagement increased significantly and social engagement decreased significantly.

Neither GCPLA range nor busy scores were significant at T1-2 or at T1-T3, using any method of analysis, demonstrating that there was no significant improvement in service users' overall use of the community or leisure activities.

In summary, it appears that the training made no significant immediate or long-term difference to service users' engagement or participation in community and/or leisure activities.

8.5.1.3 Correlations

ABC total was significantly negatively correlated with a number of measures at T2 for the experimental group: practice leadership, PSR, ASM, knowledge for both staff and managers, engagement and GCPLA range and busy. None of these was correlated with ABC total by T3.

The ASM was significantly positively correlated with both engagement and practice leadership for the experimental group only at T2 and also at T3. Practice leadership was significantly positively correlated with GCPLA range, with the PSR scores and with managers' knowledge, for the experimental group only at T2, and also at T3. The PSR was significantly positively correlated with managers' knowledge and with GCPLA range for the experimental group only at T2, and also at T3. Assistance and engagement were significantly positively correlated at T3.

9 Discussion of Managers' and Staff Results

9.1 Chapter Outline

This chapter will reflect on the impact of the Positive Behaviour Support (PBS) training on managers and their staff. As there is considerable overlap in the measures for managers and for staff, it was judged appropriate to combine the discussion of their results. Results for service users are considered separately in chapter 10, and then an overall discussion of results and the links between findings for managers, staff, and service users will be addressed in chapter 11.

The chapter will first consider the limitations of the study and then the rest of the chapter discusses the results for managers and staff in the context of the relevant literature. This chapter will not address any implications for PBS training or for the use of PBS generally; these areas will be considered in the overall discussion in chapter 11.

This discussion will be divided into two main sections: initial impact of the training and then maintenance of training effects. Within each of these main sections, there will be four further sections considering impact of the training: on support to staff, on support to service users, on PBS knowledge (for both managers and staff), and on attributions (for both managers and staff).

The sections considering support to staff will primarily focus on practice leadership, with some additional consideration of the Periodic Service Review (PSR) in the section on maintenance of training effects. The discussion on practice leadership will firstly outline how practice leadership is defined in the literature, and then will go on to consider the findings from this study in relation to practice leadership. These findings will then be compared to other studies reporting on practice leadership.

The sections considering support to service users will focus on the Active Support Measure (ASM) and on assistance. For the ASM, some individual item comparisons will be made; for assistance, comparisons will be made with a range of relevant literature, and there will be particular focus on the high levels of assistance found in this study. Consideration will be given to potential explanatory factors in relation to the high assistance, such as types of service, size of settings, staff ratios, and length of observations.

PBS knowledge and attributions about challenging behaviour will be discussed together for both managers and staff with comparison made with related literature. All of these sections will then be repeated in the second main part of the chapter, the maintenance of training effects.

9.2 Study Limitations

There were a number of limitations to this study. Firstly, although it is a control group study, it was not randomised. This was due to the fact that operational necessity meant that the managers of the services with the most challenging service users had to be prioritised for the PBS training.

This resulted in a number of significant differences between the experimental and control groups at T1; the experimental group had managers who were more knowledgeable about PBS, and they also had significant differences in a range of the staff Momentary Time Sampling (MTS) behaviours, for example, the experimental group provided less positive contact than the control group. These differences between the experimental and control groups at T1 were to some extent compensated for by carrying out additional analysis based on the change in scores from T1-T2; the sensitivity analysis always included a change score analysis, where this was not the primary analysis.

Secondly, the study was an internal evaluation, carried out by the writer who is an employee of the organisation where the study was carried out, and who is the manager of the team who provided the training and collected the evaluation data. There was therefore clearly a potential for bias; a number of measures were put in place to address the potential for bias, for example, blind inter-reliability checks for observation data, and double-scoring of 5% of all evaluation questionnaires.

Thirdly, the services in this study scored very high at T1 on a number of measures, for example in engagement and in staff assistance; it may therefore have been difficult to produce significant improvements due to already high scores.

Fourthly, the poor internal reliability found in the CHABA at T1 is a limitation as four of the subscales (Learned Negative, Biomedical, Emotional, and Physical Environment) had a Cronbach's alpha less than 0.6, which is generally regarded as the acceptable standard, and one of those (Learned Negative) scored less than 0.4.

The number of participants dropping out from the study can also be regarded as a limitation. Both managers and service users left the study for a range of reasons and therefore their data was not included in the per protocol (PP) analysis which had only 22 participants by T3 (44%). Significance of some results was changed by the approach to analysis and to some extent this undermines confidence in the PP results. The sensitivity analysis approach has compensated for this to a degree: in this study only results that are significant regardless of method of analysis are accepted.

Finally, the lack of reliability data in terms of adhering to the training which was delivered is also a limitation, as without it, it is not possible to evidence procedural integrity. This means that there was no way of assessing the consistency and accuracy of implementation of the training by the range of trainers involved; this therefore undermines the results to some extent.

Impact of Training (T1-T2)

9.2.1 Support to Staff – Practice Leadership

This study evaluated outcomes from training frontline managers in PBS. Outcomes were evaluated for managers, staff and service users; however it was the managers who received the

training, so it is they who are the starting point for any impact on the two other groups, that is, for any changes in staff practice and any changes in service users' support. Their role is therefore pivotal in terms of the theory of process of change within this study; for that reason it is worth exploring further what research tells us about the role of the frontline manager within social care services for people with learning disabilities.

It is generally agreed that the frontline management role within social care settings is multifaceted and varied, for example, Hewitt et al (2004) carried out research into frontline managers in services for people with learning disabilities and identified it as a skilled and varied role with 142 competencies involved. Clement & Bigby (2012) used the same 142 competencies in their study about frontline managers. They carried out a series of semi-structured interviews with experienced managers about their views of the competencies needed, and 92% of the original competencies were retained. The authors noted the complexity and breadth of the frontline manager role, and the difficulty at times of prioritising and making sense of the whole role.

Cherniss (1980) explored staff burnout in human services and as part of this focused on the role of the supervisor, describing supervision as a 'mentor relationship' through which staff are supported to grow and develop. He suggested that supervisors need to provide technical assistance, emotional support, feedback and protection from external pressures. In a study based on self-report by managers about how much time different tasks took, Beadle-Brown et al (2006) found working with service users, ensuring service quality and helping staff with difficulties were rated as the most important tasks; the authors also noted how varied the job was. A time-budget study by Gifford et al (2006) found that administration was the activity which took up most of managers' time, whereas Ford & Honnor (2000), found that managers believed recruitment, retention and training of staff was their main challenge. This range of studies demonstrates the varied nature of the role.

In addition to the wide range of tasks required by frontline managers, it is also generally agreed that frontline managers have a substantial impact on the behaviour of staff which in turn impacts on the extent to which service users' outcomes are met (Beadle-Brown et al, 2008; Clement & Bigby, 2007; Clement & Bigby, 2012; Mansell et al, 2002). Hewitt et al (2004) for example identified the important role of the frontline manager, noting that it is he/she who 'defines the job, provides the training, mediates the stress, creates the culture...' (p.133). This view sees frontline managers in a key role in terms of how staff experience their job and also how they go about carrying out their work-related activity. Over the last 10 years there has been increasing emphasis on the importance of frontline managers adopting a practice leadership role in order to improve the quality of support to service users (Ashman et al, 2010; Beadle-Brown et al, 2008; Deveau & McGill, 2013; Mansell et al, 2005b; Mansell & Beadle-Brown, 2012).

Practice leadership was originally identified as an important factor in services for people with learning disabilities by Mansell et al (1994), where the authors note that the role of house managers needed to be redefined as practice leaders, rather than as administrators. Mansell et al (2005b) went on to define practice leadership as: ‘the development and maintenance of good staff support, for the people served, through:

- focusing in all aspects of their work as manager, on the quality of life of service users and how well staff support this
- allocating and organising staff to deliver support when and how service users need and want it
- coaching staff to deliver better support by spending time with them providing feedback and modelling good practice
- reviewing the quality of support provided by individual staff in regular one to one supervision and finding ways to help staff improve it
- reviewing how well the staff team is enabling people to engage in meaningful activity and relationships in regular team meetings and finding ways to improve it’ (taken from Mansell & Beadle-Brown, 2012, p.108).

Mansell et al (2005b) also identified a number of aspects to the way the practice leader helps the team shape the quality of support they provide; these were identifying areas to improve via encouraging reflective practice and observing team members providing support; coaching and modelling desired behaviours; and reviewing and confirming improvement, via team meetings and supervision. Clement & Bigby (2008) described practice leadership as involving ‘coaching, directing, role-modelling, supervising, and supporting’ (p.51). This definition was further refined and discussed in Mansell and Beadle-Brown (2012, chapter 5).

The current study found no significant change in practice leadership scores from pre to post training, although there were slight increases for both the experimental and control groups; this would indicate that the PBS training did not impact on how frontline managers carried out their practice leadership role. The study also found that there was a significant positive correlation between practice leadership scores and Periodic Service Review (PSR) scores at T2; since the PSR is a measure of how fully the PBS plans are being implemented, and practice leadership measures how much managers are providing direction, support and mentoring it is not unexpected that this was positively correlated with the practice leadership score; well implemented PBS plans require good practice leadership to be in place.

In the current study part of the Staff Experiences and Satisfaction Questionnaire (SESQ) (Beadle-Brown et al, 2005) was used to assess practice leadership; although research using this is limited, there are a few studies which have used part of it for a practice leadership questionnaire (Beadle-Brown et al, 2012; Beadle-Brown et al, 2014; Deveau & McGill, 2013; McGill et

al, 2010). This research is summarised in Table 9.1, following which, each of the relevant studies is discussed in turn.

McGill et al (2010) carried out a comparative analysis of the quality of provision in the Hesley Group's homes in England. Ninety-eight staff returned the SESQ from a total of 447. Comparison was made with a set of unpublished data from a study in progress by Beadle-Brown et al (unpublished, cited in McGill et al, 2010). The comparison data came from a staff group of 505, working in a UK social care service provider.

Beadle-Brown et al (2012) evaluated the impact of active support training on a range of variables including service user engagement, staff assistance, and practice leadership. The authors found that practice leadership improved following active support training, with the percentage of staff reporting that their manager usually modelled good practice increasing from 42% to 78% and those reporting that managers usually gave feedback moving from 48% to 81%.

Deveau & McGill (2013) reported regarding the impact of practice leadership on staff experiences of working in services for people with challenging behaviour. They carried out a survey of 149 staff based on the SESQ with 19 items covering 3 main areas: team meetings, supervision and being observed by manager. Their method of scoring was slightly different from the Beadle-Brown et al studies (2012; 2014) in that most items received a score on a scale (always=4; usually=3; sometimes=2; never=0), giving a total possible score of 60. They found a mean of 41/60, which for comparison has been calculated as a practice leadership index of 69%.

Beadle-Brown et al (2014) explored the role of practice leadership in explaining variation in active support using data from a large organisation which implemented active support at two different time points. They reported two sets of data collection, one from 2005/06 and the other from 2009/10, and found that average levels of practice leadership were low, although this improved from a practice leadership index of 42% at T1, to 55% by T2, a 31% increase. The reasons for this were that managers were more likely to observe staff at least monthly, to model good support, and to give useful feedback. However, as this change was not following a specific training or other intervention, but just over a period of time, it is perhaps not helpful to compare the change in results from T1-T2 with results from the current study. Perhaps a more useful comparison is the Beadle-Brown et al (2012) study which reported a 12% increase in practice leadership index, following implementation of active support; this increase is very similar to that of the current study in relation to PBS training, both types of intervention having a similar impact on levels of practice leadership.

Table 9.1 Comparisons with Other Studies for Practice Leadership Index and for Team Meetings, Supervision and Observation

Study Details	Current Study Experimental			McGill et al, 2010	Comparison from McGill et al, 2010	Beadle-Brown et al, 2012			Deveau & McGill, 2013	Beadle-Brown et al, 2014
	T1 (n=50)	T2 (n=50)	% Change	(n=98)	(n unknown)	T1 (n=36)	T2 (n=32)	% Change	(n=140)	T2 (n=187)
Practice Leadership Index (range)	63 (31-89)	70 (25-94)	11	50	74	74 (37-94)	83 (65-95)	12	69	55 (0-100)
Team meetings Main focus is engagement	30	24	-20	30	32	-	-	-	51	45
Supervision Constructive feedback	60	68	13	50	58	-	-	-	-	71
Main focus is engagement	22	36	64	20	28	-	-	-	29	-

Study Details	Current Study Experimental			McGill et al, 2010	Comparison	Beadle-Brown et al, 2012			Deveau & McGill, 2013	Beadle-Brown et al, 2014
	T1	T2	% Change			T1	T2	% Change		
Observation by Manager										
Observation at least monthly	82	82	0	62	54	-	-	-	-	59
Main focus is engagement	26	48	85	29	48	-	-	-	56	55
Feedback usually	72	86	19	75	82	48	81	69	83	-
Models good support usually	40	64	60	22	40	42	78	86	-	44
Correction to support better usually	56	62	11	27	44	-	-	-	-	-
Can approach manager for support usually	78	80	3	55	79	-	-	-	-	-

Comparisons in the three main areas of team meetings, supervision and manager's observations can be made for the studies that publish this data. This indicated that practice leadership index scores ranged from 50% to 74%, putting the current study round about the middle of this at 63%. Between 30-50% of people said that engagement was the main focus of team meetings; the current study was at the bottom end of this at 30%, and in fact this fell to 24% by T2 (it was noted that the main focus of team meetings moved to health and safety issues by T2, perhaps a result of a major new initiative from the organisation's Health and Safety Advisor around that time). The percentage of staff who indicated that engagement was the main focus of supervision ranged from 20-29% in the comparison studies, with the current study scoring 22%; however this rose to 36% by T2, becoming the highest score in the group. It is noticeable that in the current study 82% of staff indicated that they were observed by their manager at least monthly; this is substantially higher than any other study, the next highest being 62%; it may be that this very regular observation by managers impacts on other aspects of support, for example, level of assistance to service users, although correlation between observation by manager and assistance was not shown to be significant.

9.2.2 Support to Service Users

9.2.2.1 Active Support

Although this study did not provide training in active support, it used the Active Support Measure (ASM) to evaluate outcomes from PBS training. The ASM was used as it is a helpful measure for assessing the quality of support from staff and in particular focuses on how staff facilitate participation for service users; it therefore was very suitable for evaluating outcomes from PBS, which has a considerable emphasis on supporting individuals to participate in activity. This study found that there was no significant difference in ASM scores from pre to post training, although there was a slight increase for the experimental group in this period.

Extensive evaluation has taken place of active support over the past 20 years, using both the ASM and Momentary Time Sampling (MTS) (this research is summarised in Chapter 1 and Chapter 3 of Mansell & Beadle-Brown, 2012). As the PBS literature does not refer specifically to use of the ASM, it is helpful to consider some of the active support research in relation to the current study. Table 9.2 compares each item on the ASM from the experimental group at T1 with Beadle-Brown et al (2011); it is noticeable that in all but item one, the means in the current study are substantially higher at T1, and also the current study had a higher percentage of participants scoring 'good' at T1 in most of the items. Of particular note are item seven 'graded assistance to ensure success' and item 13 'staff work as a coordinated team to support service users'; the means are markedly higher in the current study. These two statements are key elements of active support (e.g. Mansell et al, 2001, p.253), as graded assistance is an essential

element to facilitate participation regardless of ability or disability, and staff working together as a team is essential in order to achieve high levels of individual assistance to each service user. These scores being so comparatively high would seem to indicate a good level of active support and staff assistance to service users was already in place.

Active support scores have been categorised as follows: 67% or more equates to good active support; 33%-67% is mixed active support; and less than 33% is regarded as very weak active support (Mansell & Beadle-Brown, 2012, chapter 3); comparisons with the current study can be seen in Table 9.3. It is clear from the comparison figures that service users in the current study were receiving relatively high levels of active support pre training: 59% ASM for the experimental group at T1 in the current study compared to a range of 33% to 65%). However, although this is a relatively high pre training score, it is still in the mixed active support category and therefore there was still room for increase; it must however be noted that this study was not providing training in active support and therefore perhaps significant increase in an active support specific measure would not necessarily be likely.

The current study found significant positive correlations between ASM scores at T2 and PSR scores at T2, between ASM scores and practice leadership at T2, both for the experimental group only, and between ASM scores and MTS assistance at T2 for both experimental and control groups; the ASM and assistance correlation is to be expected, as they are both measures of staff's interactions with service users, and in this study they are both based on the same observation period. This association has been found in a number of other studies also (e.g. Beadle-Brown et al, 2012).

Beadle-Brown et al (2011) found that practice leadership scores were slightly higher in teams providing good active support (practice leadership mean 54%), compared to teams providing weak active support (practice leadership mean 49%). In the current study services providing good active support had a higher practice leadership mean (68% at T1, 77% at T2) than teams providing weak active support (practice leadership mean 48% at T1, 40% at T2).

Beadle-Brown et al (2014) used a questionnaire along with observational measures and explored relationships between practice leadership and active support. They found that practice leadership had a significant impact on active support but that more important was the quality of management; when management quality was high, then good practice leadership produced a difference in active support, but not otherwise. Quality of management was assessed by the Quality of Management Scale (Freeman, 1992, in Carpenter et al, 2000). They concluded that combining good management with practice leadership was most effective in achieving higher levels of active support. Beadle-Brown et al (2015a) also report positive correlations between the ASM and practice leadership, and also between practice leadership and contact from staff,

although not of assistance. They also found that in services where practice leadership was better, so too was active support.

9.2.2.2 Momentary Time Sampling (MTS) Assistance

Although MTS measured a number of different staff behaviours, the one of primary interest is assistance, as this has been associated with higher rates of service user engagement (Mansell & Beadle-Brown, 2012). This study found that there was no significant difference between pre and post training in staff assistance to service users as measured by MTS.

Table 9.3 shows percentage assistance from the MTS, along with comparisons from other studies reporting the same data. It is noticeable that assistance levels in the current study are markedly higher than in comparison studies which range from 2-23%. These include an early study evaluating active support by Jones et al (1999), an evaluation of active support in Australia by Stancliffe et al (2007), and a PBS study by McGill et al (in press) who reported on a PBS intervention designed to prevent challenging behaviour by taking a proactive prevention approach. This high level of assistance is consistent throughout the current study, high scores were noted from T1 onwards; as described in the method chapter (section 5.11.2) this was one of the reasons for additional coding being carried out in order to check the accuracy of this data. These checks indicated that the coding was accurate, and therefore this high level of assistance is an area worth further exploration.

In considering why the assistance levels in the current study are so much higher than in other studies, there are a number of possible factors. Firstly, as noted earlier, the levels of active support are high in comparison to most other studies, although those studies with equally high ASM scores, do not demonstrate equally high assistance levels (e.g. Beadle-Brown et al, 2014 has an ASM of 61%, but assistance is only 9%). It may be that it is not the overall score of the ASM that is specifically relevant to assistance; rather it may be specific questions within it which are more central to the definition of active support, such as graded assistance and the staff working as a coordinated team. As noted above, these are areas where the current study scored significantly better than the comparison study, Beadle-Brown et al (2011); however data on the detailed scores of the ASM are limited, and therefore it is difficult to make any firm conclusions in this respect.

Table 9.2: Item Comparison on the ASM for Mean Scores at T1 and Percentage Scoring 'Good' at T1

	T1 Mean		T1 % Scoring 'Good'	
	Current Study Experimental	Beadle-Brown et al, 2011	Current Study Experimental	Beadle-Brown et al, 2011
1. Age appropriateness of activities	1.95	2.33	34	48.5
2. Real rather than pretend activities	1.83	1.73	22	0
3. Choice of activities	1.57	0.94	12	0
4. Demands presented carefully	1.57	1	12	0
5. Tasks appropriately analysed to facilitate involvement	1.67	1.03	18	0
6. Sufficient staff contact for service users	2.05	0.82	30	0
7. Graded assistance to ensure success	1.63	0.39	22	0
8. Speech level matches developmental level of service user	2.10	1.24	18	9
9. Interpersonal warmth	2.38	2.09	28	30
10. Differential reinforcement of adaptive behaviour	2.05	1.88	14	0
11. Staff notice and respond to service user communication	2.31	1.24	28	9
12. Staff manage challenges well	1.62	1	12	0
13. Staff work as a coordinated team to support service users	1.69	0.52	6	0
14. Teaching is embedded in everyday activities	1.10	0	0	0
15. Written plans are in routine use	0.62	0	2	0

Table 9.3 ASM and Assistance Comparisons

Study (<i>n</i>)	ASM		Assistance		
	Mean (Range)		Mean % (Range)		% Increase
	T1	T2	T1	T2	T1-T2
Current Study Experimental (50)	59 (2-91)	64 (24-100)	35.53 (0-90)	35.83 (3-100)	0.8
Current Study Control (22)	69 (18-91)	68 (24-91)	53 (1-98)	43 (0-98)	20
McGill et al, in press Experimental (35)	42	55	3.34	6.23	87
Control (38)	47	42	2.53	3.00	19
Beadle-Brown et al, 2015a (171)	50 (2-97)	-	3 (0-28)	-	-
Beadle-Brown et al, 2015b (107)	59 (5-95)	-	6.3 (0-69)	-	-
Beadle-Brown et al, 2014 (125)	61 (18-97)	69 (15-100)	9 (0-48)	11 (0-80)	22
Mansell et al, 2013 (147)	39	-	3	-	-
Beadle-Brown et al, 2012 (33)	33 (17-54)	64 (25-93)	2 (2-24)	10 (0-47)	400
Totsika et al, 2010 (21)	52	58	7.97	6.06	-24
Mansell et al, 2008 Ex- perimental (169)	65	-	12	-	-
Mansell et al, 2008 Control (190)	54	-	8	-	-
Stancliffe et al, 2007 (31)	-	-	7.27	11.42	57
Mansell et al, 2002 (23)	50	66	-	-	-
Jones et al, 2001b (188)	-	-	8.1	17.7	119
Jones et al, 1999 (19)	-	-	5.9	23.3	250

Another factor in understanding the high assistance may be a link with practice leadership. The current study found a significant positive correlation between practice leadership and ASM scores, similar to Beadle-Brown et al (2015a); in particular the very regular observation by managers in the current study which is substantially higher than in a range of comparison studies,

may be a factor in understanding the high levels of assistance. However, it must also be noted that although observation is regular, the number of staff reporting the focus of this as being service user engagement is comparatively low, 26% compared to other studies reporting 29-56%. In addition, correlation between observation by manager and assistance was not shown to be significant; it may therefore be that manager's observation does not have an impact on levels of assistance provided to service users.

The size of the settings in which service users in this study lived may be a factor in the high levels of assistance; there is some research to show that smaller sized living units are associated with better outcomes for service users (Stancliffe, 1997), particularly when the units are very small sized (Tossebro, 1995). Felce et al (1991) observed engagement in a range of different settings and found that in every case, larger staff and service user groups were associated with lower service user activity. There was a noticeable difference in the size of the settings in the current study to those in comparison studies. Over 70% of the service users in this study lived alone, and on average there were 1.1 service users per setting in the experimental group. This is very different from the settings in the comparison studies, which are group homes with a mean number of service users per setting ranging from 3-6. It is also higher than the percentage of people with learning disabilities living alone in Scotland; the most recent national statistics indicated 54.7% of all adults with learning disabilities are the only person with learning disabilities in their accommodation (SCLD, 2015). Table 9.4 summarises this information. In addition, it is clear from some studies that the setting described is not an ordinary house, but more of a semi-institutional setting, for example, Perry et al (2011) describe the settings as being on the outskirts of towns or villages and having four bedrooms and two bathrooms, with offices, meeting rooms and a smoking room; the description does not appear to be of an ordinary house in the community.

Staffing ratios may also be an aspect worth exploring as it is another area where this study differs substantially from comparison studies. Felce et al (2002) found that a larger size of service user group predicted lower staff to service user ratios, and Felce et al (2003) found that settings with fewer service users were associated with higher staffing input; in this study the average whole-time equivalent (WTE) staff to service user ratio is 3:1 for the experimental group. Comparison studies have markedly smaller ratios, ranging from 1.22:1-2.84:1. Table 9.4 also summarises this information. Felce et al (2002) found evidence that higher staff ratios (along with more experienced staff) predicted higher levels of assistance to service users, so the considerably higher staff ratio in the current study may be a factor in the high levels of assistance. There were similar findings in Felce et al (2003) where higher staff ratio was shown to have an

association with higher staff attention to service users. However other studies have found no association between staff ratios and assistance (Mansell et al, 2003; Felce et al, 1998), and Felce et al (2003) also note that higher staff to service user ratio was only one of three variables that collectively explained 32% of the variance in staff attention.

Table 9.4: Setting Comparisons

Study	Average Service Users per Setting (Range)	WTE Staff:Service User Ratio (Range)
Current study (T1) Experimental	1.1 (1-4)	3.00:1 (0.5:1-6:1)
Current study (T1) Control	1.36 (1-4)	2.19:1 (0.79:1-4:1)
McGill et al, in press Experimental	3.45	-
Beadle-Brown et al, 2015b	3 (1-8)	-
Mansell et al, 2013	4.44* (2-6)	1.25:1 (0.44:1-3.42:1)
Beadle-Brown et al, 2012	5.5* (3-8)	-
Perry et al, 2011	3.5* (3-5)	-
Mansell et al, 2008 Experimental Control	4.99*	1.37:1 1.22:1
Stancliffe et al, 2007	4.4* (4-5)	1.49:1**
Bradshaw et al, 2004	3.67	-
Felce et al, 2003	3.2	1.8:1**
Felce et al, 2002	3.3 (2-6)	1.5:1
Mansell et al, 2002 Experimental Control	6 (4-7) 5 (3-10)	1.5:1 1.6:1

*calculated from number of houses and number of service users; **calculated from number of hours, based on 35 hour staff week

More information regarding the types of services in the current study is summarised in Table 9.5. ‘Shared living with 1:1 staff support’ means that the service is funded for 1:1 staffing during waking hours, thus covering the period of observation, 4-6pm. ‘Shared living with shared support’ means that the service user was not funded for 1:1 support, although they may have been receiving that at the time of the observation. This indicates that 93% of the service users in this study, both from experimental and control groups, were supported on at least a 1:1 staff basis. This appears to be a notable difference from the comparison studies and along with the low average number of service users per setting, may be an additional factor in explaining the high levels of assistance. An additional factor may be that carrying out observations in an individual setting with only one service user, and often only one staff member, is both more intrusive and more intense for those being observed since there is no perceived ‘break’ from observation while others are observed; it may be that these circumstances impact the behaviour of staff, and in particular the level of contact to service users.

Table 9.5: Type of Service in Current Study

	Living alone	Shared living but 1:1 staff support	Shared living and shared support
Experimental (50)	35 (70%)	13 (26%)	2 (4%)
Control (22)	16 (73%)	3 (13.5%)	3 (13.5%)
Total (72)	51 (71%)	16 (22%)	5 (7%)

Other factors in relation to the high assistance in the current study which have been considered are: length of time sample interval, time of observation, average length of observation per service user, and staff to service user ratio during the observation. The comparisons for these factors are shown in Table 9.6; the time of observation was fairly consistent and the sample interval of 20 seconds in the current study was not unusual, although there is some evidence that sample intervals of more than 30 seconds are less accurate (Brulle & Repp, 1984). There was a difference in length of observation per service user; in the current study average observation in minutes for the experimental group was 92, whereas for the comparison studies this ranged from 20-35. However, it is not clear that length of observation per service user would have any link to levels of observed assistance, although there is some evidence of increased probability of inaccuracy where contact is of short duration (Harrop et al, 1994). The main relevant difference between the current study and the wider literature which may impact on level of staff assistance is in relation to staff to service user ratios during observation. In the current study there was on

average 1.08 staff for every service user during observation in the experimental group; in the comparison studies the average staff per service user ranged from 0.36-0.83 during observation.

Table 9.6: Observation Comparisons

Study	Time Sample Interval	Time of Observation	Average Observation per SU in Minutes (Range)	Staff:Service User Ratio During Observation (Range)
Current study (T1) Experimental	20s	4-6pm	91.66 (30-120)	1.08:1 (0.5:1-2:1)
Current study (T1) Control			81.66 (31-120)	0.94:1 (0.5:1-1:1)
Beadle-Brown et al, 2015b	1m	4-6pm	35 (5-105)	0.88***
Mansell et al, 2013	1m	4-6pm	22 (3-50)	-
Beadle-Brown et al, 2012	1m	4-6pm	24 (5-40)	0.75:1 (0.67:1-1:1)
Perry et al, 2011	20s	evening meal	-	-
Mansell et al, 2008	1m	meal time	-	-
Stancliffe et al, 2007	NA*	evening meal	20	0.36:1**
Felce et al, 2003	-	4.30-7.30pm	-	0.56:1*** (1.11:1-0.25:1)
Felce et al, 2002	20s	4-7pm	-	0.53:1*** (0.2:1-1.42:1)
Felce et al, 1998	20s	evening meal	-	0.83:1***

*not MTS, continuous recordings of duration of each behaviour; **calculated from mean number of staff on duty and mean number of service users; ***calculated from ratios of number of service users to one staff member

Aside from the issue of the high assistance in this study, it is also noticeable that assistance did not change at all over the course of the study. Given that this training programme was a year-long, practice-based, person-focused PBS training course, it is perhaps surprising that there was no impact on assistance. Considering the content of the course, there was a substantial focus on developing detailed support plans in order to assist service users to participate more fully in activity. There was a specific workshop on developing activity-based support plans, and this

workshop contained some elements of active support training, in that the four key principles were taught (Mansell & Beadle-Brown, 2012, p57-65). However, despite all of this, there was no change in levels of staff assistance. One potential explanation for this is that assistance started high and therefore had little opportunity to increase further; however given it was only a mean of 35%, it appears at least theoretically possible that it could have increased further, particularly considering that 96% of the experimental group were supported on at least a 1:1 basis during observation, and 70% had no other service users living with them. It should be noted however that very high levels of assistance could be aversive for some individuals and may in fact increase escape-related behavioural challenges. As there are no studies reporting such a high assistance at T1, it is difficult to make any useful conclusions in this respect. Another potential reason for the lack of change in assistance however, may be the lack of hands-on training within the training programme. The active support literature has evidenced that lack of the hands-on training or coaching has a substantial impact on the training's effectiveness (Jones et al, 2001b). Other studies have also shown that a combination of classroom based training and on the job coaching is the most effective format of staff training (van Oorsouw et al, 2009), and it may be that this is the element which this particular training programme lacked in order to demonstrate an impact on staff assistance to service users.

9.2.3 PBS Knowledge

This study found a significant increase in managers' knowledge regarding PBS following the training and no significant change in staff's knowledge of PBS following the training done by their managers. Following training, managers' knowledge of PBS was also significantly positively correlated with PSR scores at T2. Staff knowledge has been shown to be an important aspect of managing challenging behaviour (Hastings & Remington, 1994a; Hastings & Remington, 1994b; Hastings et al, 1995), as an understanding of the reasons for behaviour and knowledge of behavioural principles are likely to impact on staff's responses. Lack of knowledge has been found to have an impact on staff anxiety and on burnout (Allen et al, 1990; Bromley & Emerson, 1995). Lack of knowledge about PBS has also been found to be a barrier to its implementation, and several studies have indicated a lack of behavioural knowledge in staff (Rae et al, 2011; Wishart et al, 2013).

However, in their regression analysis study of factors predicting helpful staff responses to challenging behaviour, Wishart et al (2013) note that there was a large percentage of variance unaccounted for by knowledge, therefore indicating there are other factors which impact on the relationship between knowledge and practice. Some suggestions for failing to implement best

practice are outlined in Emerson et al (2000b) and include resources and organisational inefficiency and also in McKenzie et al (2005) where staff report lack of consistency and difficulty in holding staff meetings as barriers. The PSR is relevant here as a method of measuring the level of consistent implementation of PBS plans, as well as the regularity of team meetings.

There are a number of studies that report on impact on knowledge of PBS following training and a range of different measures are used, most of them developed specifically for the training being carried out. There does not appear at this stage to be a widely used knowledge test for PBS, and this would be a useful addition to the field; it may be that this could be developed based on the new PBS competency framework (Denne et al, 2013). For ease of comparison of the difference measures, scores from each of the studies have been converted into percentages; difference between T1 and T2 percentage scores have then been calculated as a percentage change from T1. This information is summarised in Table 9.7.

Wardale et al (2014a) developed a knowledge acquisition test based on O'Neill et al (1997) which they also used in their study of PBS training in forensic settings (Wardale et al, 2014b). They report that the test covered the following elements: objectives of PBS, behaviour chains, functions of behaviour, data collection methods, topography and measurement, and the Competing Behaviour Model, including teaching alternative and functionally equivalent skills. The maximum score was nine.

Wills et al (2013) reported on their PBS training to staff in local services delivered by a health challenging needs service. Thirty-eight staff completed the 10-point multiple choice questionnaire pre and post training and there was a significant increase in knowledge following the training. Lowe et al (2007b) evaluated the impact of PBS training on staff knowledge following a 10-day training course. Staff knowledge was assessed via questionnaires of 15 sets of questions before, after and at one-year follow-up to PBS training. Significant increases in knowledge were found immediately after the training with further increases evident over time for some staff. McGill et al (2007) reported on a two-year university diploma course with outcomes for three cohorts of students. Knowledge was assessed via the Self Injury Questionnaire (SIBUQ) (Oliver et al, 1996) which is a 27 item multiple choice questionnaire, with a knowledge subscale of 11 questions about behavioural techniques. The study found a significant increase for knowledge between T1 and T2 on this scale and the authors conclude that an extended programme of training can have a beneficial impact on knowledge. Allen & Tynan (2000) carried out an evaluation of training focused on reactive strategies which was delivered by a health intensive support service which works within a PBS model. The reactive strategies training delivered to 109 staff was evaluated via a questionnaire, designed for this study, consisting of 20 multiple choice questions. The study first compared a trained and non-trained

group, and then the non-trained group were trained and their scores before and after training were compared; it is these latter scores that are reported here for comparison.

Table 9.7: PBS Knowledge Mean Percentage Comparisons

Study (n)	Number of Training Days	T1 % Score	T2 % Score	% Increase T1-T2	Significance Level
Current Study Managers Experimental (50) Control (22)	10 NA	71 60	83 67	17 12	0.033 NS
Current Study Staff Experimental Control	NA NA	56 54	61 61	9 13	NS NS
Wardale et al, 2014a (234)	4	24	66	175	<0.001
Wardale et al, 2014b (6)	3	33	63	91	-
Wills et al, 2013 (38)	2.5-3.5	51	68	33	<0.001
Lowe et al, 2007b (205)	10	55*	73*	33*	<0.001
McGill et al, 2007 (79)	57	64	71	11	-
Allen & Tynan, 2000 (58)	2-3	69	75	9	<0.005

* Average of registered and non-registered staff scores

9.2.4 Attributions about Challenging Behaviour

The role of attributions in responding to challenging behaviour has been explored in the literature over many years (Hastings & Remington, 1994a and 1994b; Hastings et al, 1997; Weiner, 1980). Attributions are staff's beliefs about challenging behaviour and their 'making sense' of the reasons why it happens; positive attributions are considered to be those that see the behaviour as external, uncontrollable, unstable, and not personal (Weiner, 1980). Helpful attributions about behaviour (for example, that it is uncontrollable) have been found to be associated with helpful staff responses; studies have also shown that staff are less likely to offer help if they believe behaviour to be under the service user's control (Dagnan et al, 1998; Wanless & Jahoda, 2002). If staff incorrectly attribute causes of challenging behaviour to factors within the service user, then this may result in a reluctance to implement interventions which do not fit with this belief. For example, Hastings (1995) reported that 74% of staff believed service users' challenging behaviour to be intentional, so addressing these unhelpful attributions has been regarded as a relevant target for staff training in challenging behaviour.

This study found no significant change in attributions regarding challenging behaviour following PBS training as measured by the Challenging Behaviour Attributions Scale (CHABA) (Hastings, 1997), either for staff or for managers. Hastings (1997) hypothesised that Learned Positive (LP) attributions would decrease and Learned Negative (LN) attributions would increase as a result of training. In this study both LP and LN attributions decrease following training, for both staff and managers, although this change is not significant. Apart from Physical Environment (PE) and Stimulation (ST), all the subscales reduced following training, for both staff and managers, but without any significant change. This lack of any significant change is somewhat surprising, especially given the number of other studies that have reported significant changes in the CHABA subscales following PBS training. Findings from some comparison studies are detailed in Table 9.8 and a short summary of each is below.

Wardale et al (2014b) reported on the CHABA following a four-day PBS training course and found significant increases from pre to post training for all subscales except Emotional (EM). Gore & Umizawa (2011) used the CHABA to evaluate their PBS training to staff and carers of people with challenging behaviour; they report 33 family carers completed the measure and they found a significant decrease on the LN scale, and a significant increase on the Biomedical (BM) scale. In addition, when taken as a whole group, including staff and carers, there was a significant decrease on the EM subscale following training. Mansell et al (2008) used the CHABA as part of a wider evaluation of the impact of active support training and this is an interesting study as other studies all report on changes following behavioural training. The authors reported that the experimental group scored significantly higher on the ST subscale than the control group; and the control group scored higher on LN, which is perhaps not what would have been expected.

Lowe et al (2007b) reported that attributions changed following training for both registered and non-registered staff. Registered staff showed significant increases on all domains; non-registered staff showed the same with the exception of LP which did not change. Grey et al (2002) is one of the first studies to report on the use of the CHABA; they used it to evaluate the impact of a longitudinal training in multi-element behaviour support for 34 staff. Their use of the CHABA differed from some others in a couple of ways; firstly they asked staff to complete it with regard to a service users with whom they worked, rather than based on a written vignette; secondly they presented their scoring in a different way. While most studies give the mean score for each subscale, Grey et al (2002) calculated the percentage of staff who scored each subscale at more than zero, thus indicating that that particular causal factor was regarded as applicable and then carried out an analysis on the change in percentages, although these are only reported in terms of significant change at the third data collection point. They also noted what they felt to

be a number of deficiencies in the CHABA which they felt may cast doubts on conclusions; this included that the subscales appear to lack content validity and they suggested that accurate inferences could not be drawn from the CHABA in relation to how staff might respond to challenging behaviour.

There are a number of other studies which, like the current study, found no difference in attributions as measured by the CHABA, following staff training. Tierney et al (2007) reported on attributions from 48 staff following a three-day workshop which included PBS plan and functional assessment approaches; they found no significant change on any CHABA subscale pre and post training. McGill et al (2007) also reported on the use of the CHABA to evaluate their university diploma course and they found no significant change on any subscale except EM, although LN subscales were not reported due to poor reliability, which the authors noted meant that some of the most potentially relevant data could not be analysed. Gore & Umizawa (2011) reported no significant differences on any of the subscales for the 27 staff that completed the measure pre and post training. On the whole it appears that there are inconsistent results in the use of the CHABA, and therefore it is perhaps more difficult to draw any clear conclusions about the lack of significant change in the current study.

There are a number of studies exploring attributions using measures other than the CHABA and it is also worth considering these briefly. A systematic review of changes in attribution following staff training in challenging behaviour was carried out by Williams et al in 2012. This found 11 studies which measured attributions, five of which used the CHABA and eight of which reported changes post training. The authors noted that core characteristics of training did not distinguish the studies which reported attribution changes. The first published study to explore attribution change as a result of training was Berryman et al (1994), using the Causal Attributions for Challenging Behaviour Scale (Berryman, 1991). The study reported on two groups receiving either a non-aversive or a traditional behavioural one-day workshop. Pre training very few staff listed negative reinforcement or tangible reinforcement as causal variables; following the non-aversive training more staff attributed challenging behaviour to negative reinforcement, and were less likely to attribute emotional reasons. Kalsy et al (2007) used The Controllability of Beliefs Scale (Dagnan et al, 2004) to evaluate their four-hour workshop for staff working with people with learning disabilities and dementia, and reported significant reduction in controllability from T1-T2. McDonnell et al (2008a) used this same scale following a three-day training course and compared a training group with a control group; they reported no significant difference between the groups for thoughts about challenging behaviour following training. Rose et al (2014) also used this following a one-day PBS workshop for 65 staff, with data collection at four time points - a week pre training, immediately pre training, post training

and at two months follow-up. They found a significant reduction in controllability attributions between pre and post training.

However, few studies have found a direct relationship from attribution to helping behaviour through regression analysis (Jones & Hastings, 2003) and some studies have indicated that knowledge is more important than attributions. Wishart et al (2013) carried out a study to explore the extent to which knowledge and attributions predicted staff helping behaviour in relation to challenging behaviour. They used a self-report knowledge questionnaire developed and used in previous studies (McKenzie et al, 1999; Rae, Murray & McKenzie, 2011) to measure and code staff knowledge about challenging behaviour and its management; this was reported to have good interrater reliability. To measure attributions they used an adapted version of the Leeds Attributional Coding System (Munton et al, 1999); this codes attributions such as whether the service user was in control of their behaviour, whether the origin of the behaviour was with the person or not, and whether the cause was permanent or not. They then scored staff reports of their response to a challenging situation, based on an adaptation of the method in Hastings (1996). They found that knowledge and helpful attributions were significantly correlated with reported helping behaviour; but in a regression analysis, only knowledge significantly contributed to the variance. The authors noted that this points towards staff training to increase knowledge, rather than to change attributions.

Table 9.8: CHABA Comparisons

	Current Study Managers E		Current Study Staff E		Wardale et al, 2014		Mansell et al, 2008		Tierney et al, 2007		McGill et al, 2007		Lowe et al, 2007b		Grey et al, 2002**		Hastings 1997
	T1	T2	T1	T2	T1	T2	E	C	T1	T2	T1	T2	T1	T2	T1	T2	-
Learned Positive	1.01	0.94	1.04	0.93	1.03	1.23	-	-	-	-	1.23	1.09	1.29	1.44	74	85	1.15
Learned Negative	0.91	0.83	0.92	0.70	0.80	1.13	0.52	0.62	-	-	*	*	0.97	1.43	6	15	0.47
Bio-medical	0.45	0.43	0.39	0.25	0.57	0.74	-	-	0.44	0.61	0.28	0.22	0.73	1.10	9	3	0.19
Stimulation	0.60	0.64	0.32	0.33	0.46	0.67	0.32	0.23	0.48	0.51	0.81	0.78	0.71	1.07	15	24	0.44

emo- tional	0.98	0.88	1.03	0.93	0.99	0.98	-	-	0.91	0.89	0.91	0.66	1.18	1.39	41	50	0.80
Physi- cal Envi- ronme nt	0.46	0.53	0.27	0.30	0.49	0.79	-	-	0.39	0.60	0.57	0.51	0.84	1.25	3	9	0.38

Hastings aggression; Lowe non-registered staff; E = experimental group

* Not reported because unreliable

** Reported percentage of staff who scored each factor more than 0, indicating the causal factor is applicable

9.3 Maintenance of Training Effects (T1-T3)

9.3.1 Support to Staff – Practice Leadership

The current study found no significant change in practice leadership scores from pre training to follow-up, although there were slight increases from T1-T2, these had reduced by T3 although not to the original level; however none of these changes was significant. This indicates that the PBS training did not impact on how frontline managers carried out their practice leadership role, either immediately following training or over time. However, there was an association between practice leadership and staff's support to service users; ASM and assistance were both significantly positively correlated with practice leadership at T3. The study also found that there was a significant positive correlation between practice leadership scores and Periodic Service Review (PSR) scores at T3, and also between practice leadership scores and managers' PBS knowledge at T3.

One of the factors that may have impacted on the maintenance of training effects is the high dropout by T3; by the time of follow-up to training, only 44% (22) of the managers in the experimental group had completed the training and were still working in the same service; the remaining managers had either left the organisation (three), or the course (four) or had moved within the organisation, from the service where they were working at the time of the PBS training (15), or the service user had left the organisation or died and therefore had no data available (six). Comparing the group whose manager had changed with the group whose manager was the same showed that this did have an impact on practice leadership, with a significant difference in practice leadership scores. It would appear likely that this would impact the long-term maintenance of the PBS interventions, and this potential link between practice leadership, change of management, and impact on service users will be explored further in chapter 11.

9.3.2 Periodic Service Review

The Periodic Service Review (PSR) (La Vigna et al, 1994) was used in this study as a means of assessing the level of implementation of PBS following the training. La Vigna et al (1994) developed the PSR after identifying that poor management was the main cause of implementation problems in PBS interventions; the PSR was developed to provide a way of improving and maintaining quality of services with regards to PBS. There is very little research available on the use of the PSR; literature searches found only six studies reporting on its use, although there are a number of studies that refer to using it but without presenting any data (e.g. Grey & McClean, 2007; McClean & Grey, 2012). It is also difficult to be clear on what percentage

score on the PSR is likely to impact on outcomes for service users. La Vigna et al (1994) describe a PSR score of 85% as being that likely to indicate consistent implementation of behavioural recommendations (p. 9); however, there is no research available to confirm this statement or to consider the impact of lower scores, such as the 66% achieved in the current study. Studies with PSR data are summarised in Table 9.9.

Table 9.9: PSR Comparisons

Study (n)	PSR Percentage Score Mean (Range)
Current Study Experimental T2 (50 individuals) T3 (50 individuals)	66 (33-95) 61 (14-95)
McGill et al, in press T2 (11 services)	80 (30-92)
Lowe et al, 2010 T1 (13 services) T2 (13 services)	69 (54-83) 79 (50-93)
McClean et al, 2007 (5 individuals)	48-95*
McKenzie et al, 2006 With behavioural guidelines (11 services) Without behavioural guidelines (8 services)	21 (5-63) 7 (5-11)
Jones, 1998 (33 individuals)	78 (34-95)

* Maximum scores for each individual (this study reported PSR scores at different timescales for each individual, so it is not possible to include a meaningful mean score)

Jones (1998) is one of the first studies reporting on the use of the PSR; he used it to monitor outcomes and processes while introducing changes in a vocational service in order to be more focused on meeting services users' goals. The study reported that the PSR helped to create a positive working environment and improved staff performance, and that the organisation achieved a continual upward trend. McKenzie et al (2006) describe the use of the PSR to measure staff practice in relation to behaviour guidance; they found that PSR scores were

significantly higher for services which had formal behaviour guidelines than for services that did not, a mean of 21% (range 5-63%) versus a mean of 7% (range 5-11%).

McClellan et al (2007) reported data for five individuals using the PSR over 24 months and reported maximum PSR scores of between 48%-95% for the five individuals; however, behaviours still reduced to near zero for all the individuals, despite some lower PSR scores indicating that the PBS plans were not fully implemented. Lowe et al (2010) reported on the evaluation of PSR in 13 specialist health services for people with challenging behaviour; data were collected after the PSR had been in place for around 12 months in each of the settings. Average PSR scores increased from 69% (range 54-83%) to 79% (range 50-93%). Staff were also asked their views regarding PSR and these were generally positive. The study found no relationship between knowledge or attitude scores and change in PSR scores; however positive attitudes towards the PSR did appear to be related to knowledge. The authors link the PSR to practice leadership in that direct observation of staff practice, coaching and modelling, were addressed in the PSR standard on positive monitoring. Commenting on the Lowe et al (2010) article, Baker & Shephard (2010) reported on use of the PSR in a challenging needs service over a 15 year period where the service used PSR to monitor themselves against its aims and objectives; in addition PSR had been used to monitor PBS interventions for around 36 service users. Detailed data were not presented, but the authors noted that PSR is most useful for complex PBS plans with many different elements implemented over a prolonged period of time, with a number of different agencies and individuals involved in the support. McGill et al (in press) report on the use of the PSR in their study to explore prevention of challenging behaviour. Of the 11 services in the study, all of the services reported achieving the standards on the PSR, with an average of 80% (range 30-92%) over a period of 11 months.

In the current study PSR scores decreased from T2-T3 although this was not a significant decrease; scores are fairly low in comparison to most of the available literature, with an average score of 66% post training (range 33-95%), dropping to an average of 61% at T3 (range 14-95%). The PSR was positively associated with the ASM at T2 and T3, and with assistance at T3; this may indicate a relationship between PSR scores and outcomes for services users, and this potential relationship will be explored further in chapter 11, which considers the overall results. The PSR was also positively associated with practice leadership, and PBS knowledge for managers; this is different from the Lowe et al (2010) study which did not find a relationship between change in PSR scores and knowledge, but as no other studies were found which reported on these correlations, it is not possible to put these results into further context.

9.3.3 Support to Service Users – Active Support

This study reported a significant increase in ASM scores from T1-T3, although there had been no significant difference from T1-T2. The percentage of people receiving good active support also rose from 38% at T1 to 58% at T3, and services providing good active support had a higher practice leadership mean (72% at T3) than teams providing weak active support (practice leadership mean 43% at T3).

As at T2, the ASM score at T3 was found to be positively correlated with practice leadership at T3, thus indicating that in services where managers provided better support and coaching to their staff, then these staff provided better levels of active support to service users. This association was also found in Beadle-Brown et al (2015a). However, Mansell et al (2008) noted that demonstrating the reasons why active support is well implemented or not has proved difficult within the current literature. Their study demonstrated that most organisational and management factors were only weakly associated with good staff performance. Other research has indicated that the role of the first-line manager is important in determining the quality of support provided by staff (Larson & Hewitt, 2005), although this was not specifically about levels of active support.

Comparing the group whose manager had changed with the group whose manager stayed the same showed that this change did have an impact on ASM, with the stable management group having a staff team with significantly higher ASM scores. This indicates that where there is stability of management, staff provide higher levels of active support. This link between change of management, active support, and the potential corresponding impact on service users will be explored further in Chapter 11.

There are limited studies available with three data collections for the ASM. Totsika et al (2010) reported on the outcomes following the interactive training element of active support; they found that the ASM rose between T1-T2, but then at T3 decreased to less than the original score. The continued rise in ASM scores in the current study may be due to staff becoming more confident using PBS approaches over time, and thus their level of active support to service users increases over time, and therefore we see the gradual increase which eventually becomes significant; the increase in ASM over time may also be linked to service user factors, such as levels of challenging behaviour, and this will be explored further in Chapter 11.

9.3.4 Support to Service Users – Assistance

This study found no significant difference in staff assistance to service users over time, although this rose slightly from T1-T3 (7% increase). There are few studies available with three data collection points for MTS; one of these is Jones et al (1999) with their study of the introduction of active support to 19 services users across five supported living settings. The study found significant increase in assistance from T1-T2 and then a decrease by the third data collection point which was 8-12 months after the introduction of active support; however this still amounted to an increase of 171% between baseline and follow-up. Totsika et al (2010) also reported on assistance at three time points; they found a small increase (14%) between T1 and T3. Stancliffe et al (2007) also found a substantial increase (87%) between baseline and follow-up. These three comparison studies are reporting on active support interventions, so perhaps it is not surprising that they report higher increases in assistance than the current study, since increased assistance is a key element of active support. Also, the current study as previously noted had higher average assistance than comparison studies, and this may also be a reason for the lack of significant change over time. It is noteworthy however, that the high assistance level in the current study is maintained at all three time points, with remarkably little change over time. This information is summarised in Table 9.10.

Assistance was found to be significantly positively correlated with PSR scores at T3, indicating that fuller implementation of the PBS plans was associated with increased assistance from staff to service users. Assistance was also significantly positively correlated with practice leadership scores at T3, indicating that higher practice leadership scores were associated with increased assistance from staff to service users; this is not unexpected as practice leadership has a focus on staff's assistance to service users. Change of manager impacted on assistance also; there was a significantly lower rate of assistance at T3 from the group whose manager had changed in comparison with the stable management group. This may indicate that stability of management supports higher levels of assistance from staff to service users.

Table 9.10: Assistance Comparisons T1-T3

Study (n)	% ASM (range)				% Assistance (range)			
	T1	T2	T3	T1-T3% Increase	T1	T2	T3	T1-T3% Increase
Current Study Experimental (50)	59 (2-91)	64 (24-100)	68 (22-100)	15	35.53 (0-90.51)	35.83 (3-100)	37.94 (6-70)	7
Totsika et al, 2010	52	58	50	-4	7.97	6.06	9.08	14
Stancliffe et al, 2007	-	-	-	-	7.27	11.42	13.56	87
Jones et al, 1999 (19)	-	-	-	-	5.9	23.3	16	171

9.3.5 PBS Knowledge

This study found that there was a significant increase in managers' knowledge regarding PBS following the training, and this was maintained at follow-up. Managers' knowledge was positively correlated with PSR scores at T3, and also with practice leadership at T3, indicating that those managers who knew most about PBS provided the highest levels of practice leadership, and had staff teams who achieved higher implementation of PBS interventions.

There was no impact overall on staff knowledge of PBS following their managers' training. There was no correlation between staff PBS knowledge and assistance at follow-up, indicating that their level of knowledge about PBS was not a factor for staff in terms of the amount of assistance they provided to service users.

There are a limited number of studies which report on knowledge following PBS training at follow-up, that is, with more than two data collections; the findings from those that do are summarised in Table 9.11. Lowe et al (2007b) reported on follow-up scores one year on from training; however in this time, participants continued to learn through working through course material to complete assessment portfolios, and this may be part of the reason for the very substantial ongoing increases. In addition, the questionnaires were not administered under test conditions as at T1 and T2, but instead data were abstracted as evidence from submitted portfolio; it is likely that this change in administration method had an impact on results.

Table 9.11: PBS Knowledge Comparisons at T3

	T1 % Score	T2 % Score	T3 % Score	% Increase T1-T3	Significance Level
Current Study Managers					
Experimental (50)	71	83	84	18	0.00
Current Study Staff					
Experimental (50)	56	61	59	5	NS
Lowe et al, 2007b (65)	44	67	92	109	<0.001
McGill et al, 2007	64	71	80	25	<0.001

9.3.6 Attributions about Challenging Behaviour

This study found no significant change in attributions regarding challenging behaviour between pre training and follow-up as measured by the CHABA, either for staff or for managers. Although there were some slight increases in the subscales between pre training and follow-up, none of the changes were significant for any subscale, for either staff or managers.

There are a limited number of studies which report on attributions using the CHABA with more than two data collections; the findings from those that do are summarised in Table 9.12. Lowe et al (2007b) reported that attributions changed following training but that these reverted to baseline levels over time. At T3 some differences were noted between registered and non-registered staff with the latter attributing challenging behaviour to the factors of LP and BM to a significantly greater extent than registered staff. McGill et al (2007) reported the EM subscale to decrease significantly over time, but no other changes are noted at T3. Grey et al (2002) carried out analysis on the change in percentages from T1-T3 and found significant increases for LN and ST, and a significant decrease for LP over time.

9.4 Chapter Summary

This study has shown that following PBS training there has been a significant increase in managers' knowledge which was maintained over time, and there was also a significant increase in levels of active support provided by staff over time. There was no significant change in staff knowledge, attributions made about challenging behaviour by staff or managers, practice leadership, or staff assistance to service users. Assistance is positively correlated with the ASM, although the ASM increases significantly over time and assistance does not. While lack of significant change is perhaps not surprising in terms of staff knowledge (since they did not attend

the training), or in terms of attributions (other studies have also found little change), it is difficult to account for the lack of change in assistance, particularly for a PBS training course with such an emphasis on support to service users, except perhaps in relation to a possible ceiling effect. There is also evidence of a link between stability of management and support provided to service users via both active support and levels of assistance, and between stability of management and levels of practice leadership; these associations merit further research investigation.

Table 9.12: CHABA Comparisons at T3

	Current Study Managers*			Current Study Staff*			McGill et al, 2007			Lowe et al, 2007b			Grey et al, 2002***		
	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3
Learned Positive	1.01	0.94	0.88	1.04	0.93	0.88	1.23	1.09	1.05	1.29	1.44	1.33	74	85	59
Learned Negative	0.91	0.83	0.73	0.92	0.70	0.89	**	**	**	0.97	1.43	1.24	6	15	24
Biomedical	0.45	0.43	0.37	0.39	0.25	0.48	0.28	0.22	0.30	0.73	1.10	0.80	9	3	21
Stimulation	0.60	0.64	0.43	0.32	0.33	0.37	0.81	0.78	0.66	0.71	1.07	0.74	15	24	23
Emotional	0.98	0.88	0.86	1.03	0.93	0.97	0.91	0.66	0.59	1.18	1.39	1.22	41	50	35
Physical Environment	0.46	0.53	0.42	0.27	0.30	0.43	0.57	0.51	0.51	0.84	1.25	0.82	3	9	9

* Experimental group ** Data not reported as not reliable *** Reported percentage of staff who scored each factor more than 0, indicating the causal factor is applicable

10 Discussion of Service User Results

10.1 Chapter Outline

This chapter will reflect on the impact of the Positive Behaviour Support (PBS) training on service users. Firstly, the chapter will first consider the limitations of the study and then go on to discuss the results for service users in the context of the relevant literature. The overall discussion of results and the links between findings for managers, staff, and service users, along with any implications for PBS training or research, will be considered in the overall discussion in chapter 11.

The discussion in this chapter will be divided into two main sections: initial impact of the training and then maintenance of training effects. Within each of these two main sections, there will be two further sections, considering impact of the training on challenging behaviour and on quality of life. This will also include some examination of how challenging behaviour and quality of life relate to each other.

The section considering challenging behaviour will consider the three measures used to evaluate challenging behaviour: the Aberrant Behaviour Checklist (ABC) (Aman & Singh, 1986), Behaviour Recording Forms (BRF), and the use of Momentary Time Sampling (MTS). As well as comparison with relevant literature, there will also be consideration of how these measures relate to each other, particularly the ABC and the BRF.

The section considering quality of life will first briefly describe different areas of quality of life and then will go on to focus on engagement in activity, measured by MTS, and use of the community, measured by the Guernsey Community Participation and Leisure Assessment (GCPLA) (Baker, 2000). A range of studies have provided data on engagement and on the GCPLA so consideration will be given to these and to how the current study fits within them. Finally, the chapter will look at maintenance of training effects, and the service user results from T1-T3.

10.2 Study Limitations

As previously noted, this study was not randomised. This resulted in a number of significant differences between the experimental and control groups at T1. For service users there were significant differences between the experimental and control groups in the ABC and in MTS for total engagement and disengagement at T1. However, in relation to the non-randomised nature of the study, advice from the statistician at the time of the study design was that where there was a large change in score from T1-T2, then lack of randomisation is regarded as less of a concern;

so to some extent this is true in the current study since ABC scores changed substantially for the experimental group from T1-T2. However, it must also be noted that the ABC was completed by the manager and therefore was not assessed blind, since the manager knew whether they were in the experimental group or not.

Another limitation already mentioned is the high level of managers' dropout from the course, either by leaving the course or leaving the organisation, and therefore the number of participants who were not per protocol (PP) and had to be excluded from the PP analysis; for service users this was even greater as in addition to the other reasons for exclusion, three service users who refused to be observed, also had to be excluded from these analyses.

In addition to limitations already noted in the previous chapter, there were a number of limitations specific to service users. One limitation was the difficulty of finding suitable evaluation measures to assess quality of life in individuals with severe learning disabilities, many of whom lacked verbal communication. The observation data was used as a method for assessing quality of life, that is, for those service users who are unable to specifically say if they are happier, more involved, receiving better support, then the observation data may be able to establish these changes more clearly. However, this is not an ideal evaluation method for assessing quality of life, as it does not give service users an opportunity to comment directly, but rather to have their behaviour interpreted by others as an indication of their views; it would have been more person-centred to have had an evaluation measure to which service users could directly contribute, rather than the more externally-located method of observation by others.

Secondly, the observations were only done on one occasion at each time point, so they were very much a snapshot of the service at that point. It is clear that since the two-hour observation was a one-off at that time point, it may not have been truly representative of a person's support or activities, for example, if the service user was sick that day, they may have been less involved in activity than usual; or if a staff member was new, they may have been less confident in supporting the service user to participate in activity. However, the aim of the study was that these potential anomalies would resolve themselves across the whole participant range.

In terms of assessing changes in challenging behaviour, there were a number of limitations. The main outcome measure, the ABC, was completed by managers who were not blind to whether they were in the experimental group or not, and this knowledge could have impacted on their scoring, particularly for those managers who had undertaken the training. There was also no separate reliability check for the ABC which may have gone some way to compensate for this. The BRF were also not completed blind as these were recorded by staff who knew whether their manager had received the training or not, although the BRF did have a number of reliability checks in place. The third measure of challenging behaviour was scored blind, but was

limited in terms of the approach to coding behaviour via MTS. The codes used were from Jones et al (1999), which only code challenging behaviour as having occurred, not whether it was major or minor challenging behaviour. It appears in retrospect that it would have been more useful to have had different codes for different levels of behaviour. As it is, it is not possible to distinguish between different levels of challenges, although it appears, from informal feedback from the videos, as if most forms of challenge were minor and mainly stereotypical behaviours.

Finally, the lack of internal reliability found in the ABC is a limitation; this was tested at T1 and found a Cronbach's alpha of 0.49, which indicates poor reliability, so this also undermines confidence in results to some extent.

10.3 Impact of Training (T1-T2)

10.3.1 Challenging Behaviour

In this study challenging behaviour was measured in three different ways: ABC, BRF, and MTS. The ABC is the main outcome measure and will be considered in most detail, with some comparisons made with BRF. Challenging behaviour was also measured via MTS; however this occurred at a very low rate for both experimental and control groups, so perhaps conclusions from this may be regarded as less robust.

This study has demonstrated that for the services where managers were trained in PBS there has been a significant reduction in challenging behaviour from T1-T2 as indicated both by reduction in ABC total scores, and reduction in frequency on the BRF. Both these measures demonstrated significant reduction for the experimental group while no significant change was found for the control group. This study has also shown a significant reduction in severity of challenging behaviour following PBS training, as shown by the reduction in ABC severe behaviour scores. These reduced significantly for the experimental group while no significant change was found for the control group. However, due to the limitations noted above, both in study design and measures used, some caution must be applied to these results.

10.3.1.1 Aberrant Behaviour Checklist (ABC)

In order to put the results from this study into context and to see how the participants relate to participants in other studies, it is useful to consider other studies which have reported on use of the ABC, and in particular to see how the mean ABC total for this study (at T1, 57 for experimental and 35 for the control) compares with other studies.

In the active support literature Mansell et al (2013), Felce et al (2002), Jones et al (1999) reported means of 30, 39, and 36 respectively, while the average score in the ABC manual (Aman

& Singh, 1986) is 24 (taken from the average of the US male and female scores for 154 residents of institutions, in the age group 31-40, as the group closest in age to participants in this study). However, in studies focusing more specifically on challenging behaviour, means are higher. For example Lowe et al (1995) reported on the characteristics of those referred to specialist teams and the mean for this group was much higher at 63; for Perry et al (2011), considering resettlement outcomes for people with severe challenging behaviour, the reported mean was 47, for those described as having severe challenging behaviour; for Allen et al (2011) reporting on outcomes from specialist PBS teams the mean was 59.8; and Daynes & Baker (2014), reporting on family training for people with challenging behaviour had a pre-training mean of 47. Hassiotis et al (2009) in their study exploring the impact of specialist health behaviour teams using PBS, reported median ABC scores of 36 and 47 in the two groups; in the current study, median scores for the experimental group are 55 and 26 for the control group. Overall this shows that the participants in the experimental group in this study sit at the high end of scores on the ABC; Table 10.1 summarises this information.

Some studies have tried to more specifically set criteria for severe challenging behaviour via the ABC. Robertson et al (2004) defined severe behaviour as either those for whom the score of two sub-scales (Irritability and Hyperactivity) added together is more than 30, or who scored five or more items on these sub-scales as severe. Using the first option as a standard, 64% of the experimental group in the current study had severe challenging behaviour at T1 ($n=32$), with a mean ABC score of 69, while only 18% ($n=4$) of the control group had severe behaviour, as measured by this standard. Given all of the above, it seems reasonable to refer to the experimental group in this study as having severe challenging behaviour.

In this study the experimental group had a reduction in ABC total score to 60% of the baseline, which is around average in terms of comparisons in the literature, and severe behaviours reduced to 29% of the baseline at T2; this is at the lower end of reductions for severe behaviours, though there are few studies published with specific data on ABC severe behaviours. These range from a reduction to 16% of the baseline for Daynes & Baker (2014) to a reduction to 54% of the baseline in Beadle-Brown et al (2012). In relation to ABC total McGill et al (in press) reported a reduction to 35% of the baseline in their study working to prevent challenging behaviour using PBS in a preventative way through proactively implementing a range of positive interventions into community houses, and then monitoring these via a Periodic Service Review approach; Perry et al (2011) considering outcomes from resettlement in the community also reported a good reduction in ABC scores, to 48% of the baseline. Studies reporting impact on ABC following training family carers in PBS appear to report smaller reductions in this, for

example, MacDonald et al (2013) reported outcomes from PBS training to family carers of people with autism, and Daynes & Baker (2014) reported on PBS training to family carers with reductions to the baseline of 82% and 70% respectively.

An interesting study for comparison was carried out by Tyrer et al (2008) where the authors compared the use of two antipsychotic drugs against a placebo as treatment for challenging behaviour. Amongst other measures the study used the ABC and reported on a reduction to 43% of the baseline in median ABC total for placebo at T2 (four weeks from baseline); this compared with a reduction to 54% of the baseline for Risperidone and a reduction to 70% of the baseline for Haloperidol. The current study reported a reduction to 49% of the baseline in median ABC total from T1-T2; as with the two antipsychotics, this was a lower reduction than achieved by the placebo. The authors concluded that the psychological effect of a formal external intervention or spontaneous resolution may have been responsible for the substantial improvement in reported challenging behaviour. However this effect is not seen in the current study control group, who showed no improvement in ABC scores over the period. Table 10.1 also summarises this information.

Table 10.1: Comparison of ABC Scores at T1 and T2 for Total and Severe Behaviours, and Mean Reduction to Percentage of Baseline

Study (<i>n</i>)	ABC Total			ABC Severe Behaviours		
	Mean (Range)			Mean (Range)		
	T1	T2	Reduction to % of Base- line	T1	T2	Reduction to % of Base- line
Current Study Experimental (50)	57.44 (13-99)	33.38 (0-94)	60	6.04 (0-21)	1.78 (0-10)	29
Current Study Experimental (50)*	55*	27*	49*	-	-	-
Current Study Control (22)	35.23 (5-104)	32.64 (0-86)	93	3.09 (0-25)	2.77 (0-19)	90
McGill et al, in press (proactive PBS interven- tion)						
Experimental (35)	42.65	14.75	35	-	-	-
Control (38)	49	41.33	84	-	-	-
Beadle-Brown et al, 2015b (110)	40 (0-33)	-	-	-	-	-
Daynes & Baker, 2014 (PBS training to carers) (6)	47.33 (23-80)	33 (9-58)	70	3.17 (1-5)	0.5 (0-2)	16

Study (<i>n</i>)	ABC Total				ABC Severe Behaviours	
	Mean (Range)				Mean (Range)	
Mansell et al 2013 (follow-up to active support) (147)	30 (0-104)	-	-	-	-	-
MacDonald et al, 2013 (PBS training to carers) (37)	55 (9-124)	45 (8-115)	82	5.3 (0.24)	2.6 (0-17)	49
Beadle-Brown et al, 2012 (active support training) (30)	23 (0-76)	18 (1-50)	78	3.12	1.67	54
Allen et al, 2011 (PBS team interventions) (26)	59.8	28.1	47	9.7	1.9	20
Perry et al, 2011 (resettlement outcomes) (19)	47.1	22.5	48	-	-	-
Hassiotis et al 2009 (PBS team interventions)						
Experimental (32)	36*	25.5*	71*	-	-	-
Control (31)	47*	40*	85*	-	-	-
Mansell et al, 2008 (active support training)						
Experimental (169)	22.06	-	-	-	-	-
Control (190)	24.04	-	-	-	-	-
Tyrer et al 2008 (antipsychotic and placebo study) (86)	51*	21.5*	42*	-	-	-

Study (<i>n</i>)	ABC Total			ABC Severe Behaviours		
	Mean (Range)			Mean (Range)		
Felce et al, 2003 (association of support variables) (163)	33 (0-109)	-	-	-	-	-
Felce et al, 2002 (comparisons of supported housing) (97)	39.2 (0-120)	-	-	-	-	-
Jones et al, 1999 (active support training) (19)	36	-	-	-	-	-
Felce et al, 1998 (service comparison for severe CB) (41)	94 (59-147)	-	-	18 (5-42)	-	-
Lowe et al, 1995 (referrals to specialist teams)						
Experimental (30)	63 (13-128)	-	-	12.5	-	-
Control (21)	41 (9-86)	-	-	4.6	-	-
Singh et al, 1986 (ABC Manual) (154)	24	-	-	-	-	-

*median scores; Mansell et al (2008) data taken from control group; Felce et al (1998) data taken from specialist community house; Perry et al (2011) taken from full pre and full post

10.3.1.2 Behaviour Recording Forms (BRF)

The findings in the current study reflect other studies where there has been reduction in frequency of challenging behaviour, as measured by real-time behaviour recordings, following PBS training. One of the most common ways of reporting this is in terms of percentage reduction from baseline of reported incidents of challenging behaviour. For example, Dench (2005) reported reduction to less than 30% of the baseline for 56% of the group; Grey & McClean (2007) report a reduction to below 30% of the baseline for 70% of the target group; Allen et al (2011) reported a reduction to 22% of the baseline at discharge, timescale not specified; McClean & Grey (2012) reported an average decrease to 39% of the baseline at 3 months; Crates & Spicer (2012) reported an average reduction to 50% of baseline at 3 months. In the current study, using the information from BRF, 52% (n=26) of the experimental group reduced to 30% or less of the baseline in frequency, although 27% (n=6) of the control group also did. This appears to demonstrate that results from the current study are similar to that reported generally in the PBS training literature. McClean et al (2005) defined substantial change as a reduction to 0-29% of the baseline, to 30-69% of the baseline as moderate. The experimental group in the current study showed on average a reduction to 59% of the baseline in frequency using BRF, thus showing a moderate change, using this definition and 52% achieved a substantial change using this definition.

Using the episodic severity scoring method based on the BRF however showed no significant difference between the experimental and control groups in terms of change in score from T1-T2. When analysed separately a significant change occurred for both the experimental and control groups; however this sensitivity to different methods of analysis clearly casts doubt on these results. For studies not using the ABC there are a number of different methods for measuring severity. Crates & Spicer (2012) reported a reduction to 31% of baseline at 3 months using an episodic severity approach similar to this study. McClean & Grey (2012) reported a reduction to 70% of the baseline at 6 months using the Challenging Behaviour Rating Scales (CBRS). A number of studies use the Checklist of Challenging Behaviour (Harris et al, 1994), but do not report specific details of any changes; for example, Grey & McClean (2007) referred to significant reduction in severity but without specifying a percentage decrease, and Gore & Umizawa, (2011) reported no significant decrease in severity, but without specific data. In the current study, episodic severity measures based on BRF showed an average reduction to 66% of the baseline at T2, with 32% of the experimental group achieving a reduction to less than 30% of the baseline. This information is summarised in Table 10.2.

Table 10.2: Comparative Reductions to Percentage of Baseline from T1-T2 Using BRF

Study (n)	Reduction to % of Baseline T1-T2 for Frequency	Reduction to % of Baseline T1-T2 for Severity
Current Study Experimental (50)	<30% (for 52% of the group) 59% (average reduction)	<30% (for 32% of the group) 66% (average reduction)
Current Study Control (22)	<30% (for 27% of the group) 107% (average reduction)	<30% (for 27% of the group) 71% (average reduction)
Crates & Spicer, 2012 (32)	50% (average reduction)	31% (average reduction)
McClean & Grey, 2012 (31)	39% (average reduction)	70% (average reduction)*
Allen et al, 2011 (26)	22% (average reduction)	-
Grey & McClean, 2007 (30)	<30% (for 70% of the group)	-
McClean et al, 2005 (65)	<30% (for 32% of the group)	
Dench, 2005 (25)	<30% (for 56% of the group)	-
Baker, 1998 (5)	<30% (for 100% of the group) 11% (average reduction)	-

*based on CBRS

As there were two main methods of measuring challenging behaviour within this study, it is perhaps useful to compare the outcomes from them both. For example, using the ABC rather than BRF, only 18% of the experimental group reduced to less than 30% of the baseline compared with 52% of the group achieving this when measured by BRF. There is therefore a less positive result using the ABC rather than using BRF. However, in terms of average reduction to the baseline, both measures were very similar with an average reduction to 59% of the baseline using BRF, compared to reduction to 60% of the baseline using the ABC. Correlations carried out between ABC and BRF showed a positive correlation at T1 and T2 for the experimental group, but this was not significant. One clear difference between the two ways of measuring challenging behaviour is that the ABC was completed by managers at each data collection point as a one-off reflection of their views of behaviour, whereas the BRF were completed by staff for each incident of challenging behaviour over a month at each data collection point. It may be that the different methods of recording and the fact they are completed by staff in different roles and with different attitudes, has resulted in the differences between them; however further work is required to explore this in more detail. To the writer's knowledge, no other study has reported on differences in results linked to alternative ways of measuring changes in challenging behaviour following PBS training. Aman & Singh (1986) in the ABC Manual carried out one-hour

behaviour observations in order to confirm ABC ratings for 36 service users; they reported that the behaviour observations tended to support the ABC ratings. Grey & McClean (2007) reported outcomes using both the Checklist of Challenging Behaviour (CCB) (Harris et al, 1994) and behaviour recording forms, but did not make any direct comparisons between the two ways of recording behaviour. Allen et al (2011) use both the ABC and a record of frequency of target behaviour; they reported significant decreases in both, but did not make any comparisons. McClean & Grey (2012) similarly using the Challenging Behaviour Rating Scale (based on Harris, 1994) reported significant decreases in both measures, but did not detail any comparison of reductions using the different measures.

10.3.1.3 Momentary Time Sampling (MTS)

Challenging behaviour was also measured in this study via MTS; however levels of observed challenging behaviour were very low and changes were not significant. It is notable that using the ABC, the experimental group in this study can be regarded as having severe challenging behaviour, however the levels of observed challenging behaviour are lower than in other studies. Using MTS to measure observed challenging behaviour is more common in active support literature than in studies reporting PBS interventions, therefore it is these that are most useful for comparison although active support has a focus on increasing participation rather than on challenging behaviour specifically. Most of the challenging behaviour noted in other studies using MTS is low-level stereotypical behaviour (Beadle-Brown et al 2012; Perry et al, 2011; Robertson et al, 2004); however, this type of stereotypical behaviour is less likely where there are high levels of engagement (Mansell et al 2008; Mansell et al, 2001) and the high levels of engagement for service users in the current study may be a potential explanation for the low levels of challenging behaviour recorded via MTS.

Reduction to the baseline can also be calculated for observed challenging behaviour; in this study there was a reduction to 87% of the baseline for the experimental group and to 26% of the baseline for the control group. Other studies have reported greater decreases, for example, Mansell et al (2001) in their study of people re-settled into community houses where active support was used reported a decrease in major challenging behaviour to 44% of the baseline, although this change was not significant. There are also studies reporting little change or even increase, for example, Bradshaw et al (2004) reporting on evaluation carried out after introduction of active support, found a significant increase and Jones et al (2001a) in their study introducing active support to 38 supported living settings, found little change in observed challenging behaviour. Of the various studies presenting data from MTS about challenging behaviour, McGill et al (in press) report the greatest reduction, to 40% of the baseline for the experimental group;

apart from the current study, this is also the only study reporting this type of data following a PBS intervention, rather than following active support. It appears reasonable to suppose that interventions based on PBS would have more impact on observed challenging behaviour than those focused on increasing participation; however using MTS as a measure, the current study does not reflect that. Table 10.3 summarises this information.

Table 10.3: MTS Challenging Behaviour Comparisons T1-T2

Study (<i>n</i>)	Percentage Challenging Behaviour (range)		Reduction to % of Baseline
	T1	T2	
Current study (stereotypy and other CB) Experimental (50) Control (22)	3.16 (0-32.59) 0.68 (0-6.11)	2.76 (0-32.59) 0.18 (0-2.00)	87 26
McGill et al (in press) Experimental (35) Control (38)	25 26	10 19	40 73
Beadle-Brown et al, 2015b (107) Self-stimulation Other CB	14.5 (0-89) 5.4 (0-94)	- -	- -
Beadle-Brown et al, 2012 (30) Stereotypy Self-injury	23 (0-100) 0.37 (0-8)	19 (0-61) 0.23 (0-7)	83 62
Perry et al, 2011* (19) (stereotypy and other CB)	19.8	15.2	78
Totsika et al, 2010	5.89	2.90	49
Bradshaw et al, 2004 (11) Experimental Control	8.5 21.6	20.6 17.3	242 80
Mansell et al, 2001 (13) Major CB (self-injury, aggression, destruction) Minor CB (mainly stereotypy)	9 29	4 16	44 55
Jones et al, 2001a (106) (stereotypy)	18	17	94

*data taken from full pre and full post

10.3.2 Quality of Life

10.3.2.1 Quality of Life

Quality of life is a key outcome in services for people with learning disabilities. The eight domains of quality of life have been identified in a series of studies as: emotional well-being; interpersonal relationships; material well-being; personal development; physical well-being; self-determination; social inclusion; rights (Beadle-Brown, 2006; Beadle-Brown et al, 2009b; Schalock et al, 2002; Wang et al, 2010). It is felt that the most relevant of these to the current study are interpersonal relationships and social inclusion, as these are most likely to be indicated by the measures used; MTS assessing levels of engagement, and the GCPLA assessing use of community and participation in activity. As there is growing acceptance of the need for quality of life to be measured and used to help inform evaluation of services for people with learning disabilities (Brown et al, 2009), then quality of life changes were a key feature of the current study, particularly since lifestyle enhancement is a stated aim of PBS (see chapter 2).

10.3.2.2 Engagement in Activity

Engagement in activity has long been recognised as an important aspect of quality of life (e.g. Emerson & Hatton, 1996; Felce & Emerson, 2004) and many of the domains of quality of life (Schalock et al, 2002) require participation in activity and social engagement in order to achieve them. Mansell et al (2008) described engagement as ‘the vehicle by which many aspects of quality of life are realized’ (p398), and Felce & Emerson (2004) discussed the importance of engagement in activity in terms of maintaining physical and mental health, as well as helping to develop a positive role and status. Engagement in activity is an important end in itself and is also a stepping stone to many other positive outcomes.

This study has demonstrated an increase in total engagement for service users in the experimental group; however this change is only significant at T2 when analysed using a per protocol (PP) approach. When the whole group of participants are included in an intention to treat (ITT) approach, then the change is not significant. In the ITT approach, data from 12 participants who did not meet PP requirements were part of the analysis. This included data from six participants whose manager left the training or the service, and therefore although their data is real, it is not PP; it also included imputed data for six participants, all in the experimental group. Three of these were imputed because they refused to be observed, so they were imputed with the group’s average score. The other three participants were imputed with the worst occurring score in the group, because the service users’ service had broken down due to the level of their challenging

behaviour and the service no longer being able to cope. It appears that the intention to treat approach to missing data, made the difference between engagement being significant or not.

There are a number of factors to explore in terms of these results. The initial levels of total engagement were very high; 70.52% at baseline for the experimental group. Most of the comparison studies in relation to observed engagement come from the active support literature over the past 20 years. Engagement is measured in various ways, with different definitions; some studies differentiate between social and non-social engagement and some have a range of codes for different types of engagement, as in this study. However it is clear that in general, most total engagement levels are substantially lower, for example, Hatton & Emerson's review in 1996 of 47 studies providing data on service user activity found an average engagement of 48%; Mansell & Beadle-Brown (2012) reviewed the literature with regard to level of engagement and found average engagement of 39% (range 17%-70%) across 24 studies (p.48). Other studies report engagement of 20%, 42% and 17% at baseline (Beadle-Brown et al, 2012; Stancliffe et al, 2007; Bradshaw et al, 2004, respectively). All of these studies saw increases in engagement following active support training (to 41%, 50%, and 26% respectively), although still not to the level of the current study.

However, although the mean engagement is high, it is not unknown in terms of other studies; Mansell et al (2008) found engagement scores of 77% in their experimental group and 70% in the control group. They noted that assistance scores were low and therefore concluded that the high engagement may be explained by higher ABS than in some comparative studies (an average of 175 across the whole sample). That is unlikely to be the explanation for the high baseline engagement in the current study as ABS of 133 is not higher than average in the comparative studies (e.g. Mansell et al, 2013 evaluating AS implementation in Australia reported an ABS mean of 154, and Jones et al, 2001a considering the impact active support training reported an ABS mean of 144). This information is summarised in Table 10.4.

The mean for assistance of 35% for the experimental group at T1 was also much higher than most other studies (see Section **Error! Reference source not found.**), so this may account for the high levels of engagement in the current study as higher engagement in activity is generally associated with higher levels of staff assistance (Beadle-Brown et al, 2012; Jones et al, 2001a; Jones et al, 1999).

Table 10.4: Comparisons for Engagement T1-T2 and ABS

Study (<i>n</i>)	% Engagement (Range)		ABS (Range)
	T1	T2	T1
Current Study (T1)			
Experimental (50)	70.52 (3.53-100)	76.02 (8.00-100)	133 (35-262)
Control (22)	86.38 (0-100)	81.50 (10.07-100)	148 (52-248)
McGill et al, 2015			
Experimental (35)	42	62	-
Control (38)	-	-	-
Beadle-Brown et al, 2015b (107)	44.1 (0-95)	-	113 (27-248)
Beadle-Brown et al, 2015a (171)	47 (0-100)	-	-
Beadle-Brown et al, 2014 (125)	46 (0-100)	46 (0-100)	-
Mansell et al, 2013 (147)	51	-	154 (39-253)
Beadle-Brown et al, 2012 (33)	20 (0-82)	41 (0-91)	77 (27-154)*
Totsika et al. 2010 (21)	41.13	37.49	-
Mansell et al, 2008			
Experimental (169)	77	-	-
Control (190)	70	-	-
Stancliffe et al, 2007 (31)	42.46	49.54	-
Bradshaw et al, 2004	16.6	26.0	91 (29-178)
Mansell et al, 2001 (13)	14 (4-26)	36	-
Jones et al, 2001a	46.7	54.6	144 (20-293)
Felce et al, 1998**	35		90 (21-178)

*estimated based on Short ABS;**data from specialist community houses

People with more severe learning disabilities tend to experience lower levels of engagement than more able people (Emerson et al, 2000b; Felce et al 1998), and previous studies, mainly from the active support literature, have demonstrated positive correlations between adaptive behaviour and engagement (Mansell et al, 2013; Beadle-Brown et al, 2012; Mansell et al, 2008;

Jones et al, 2001a), demonstrating that the more able people, that is, those with higher adaptive behaviour scores, are most likely to be engaged in activity. The current study found a positive correlation (at 0.05) between ABS and engagement at T1 for the experimental group only. This appears to suggest that prior to the training, the least able in the group were the least engaged; after the training, this association was no longer present; this differs from some other studies which found the association both pre and post intervention (Jones et al, 2001a).

In the current study the experimental group were the significantly more challenging group at T1; however no correlation was found between engagement and challenging behaviour (either via ABC or BRF) for the experimental group at T1. Following the training, there was an association between challenging behaviour and engagement in that difference in ABC total from T1-T2 was significantly negatively correlated with difference in engagement from T1-T2 for the experimental group. This shows that those individuals that had the greatest reduction in their level of challenging behaviour also had the greatest increase in their engagement between T1-T2. This is reflected in some studies demonstrating an inverse association between engagement and levels of challenging behaviour, for example, Felce et al (2002), and Thompson et al (1996) who found that level of challenging behaviour predicted inactivity. However, other studies found that there was no association between challenging behaviour and engagement in activity (Emerson et al, 2000b; Felce et al, 2000).

As most of the comparison studies come from the active support literature, it is useful to consider McGill et al (in press), which reports on a PBS-based intervention. Although they report significant decrease in challenging behaviour for the experimental group, they note that quality of life as measured by observed engagement did not improve significantly more for the experimental group than for the control group. Similar to the current study, it appeared that the impact of PBS interventions is more easily demonstrated on levels of challenging behaviour than it is on quality of life. The authors of that study note that the limited time given to the observations may have been insufficient to account for variation and may explain the lack of significant change; it is possible that this was also a factor in the current study and the length of observation has already been noted as a limitation of the current study.

10.3.2.3 Participation in Community and Leisure

In addition to service user engagement, another method of considering quality of life was to measure service users' involvement in the community and in leisure activities, as participation in the community has long been regarded as an indicator of quality of life (e.g. O'Brien, 1987; Wolfensberger, 1972). The GCPLA was used and the results from the current study showed no significant change from T1-T2, in fact practically no change at all in the range of busy scores

for either the experimental or control group, thus indicating that there was no change in service users' use of community or leisure activities following the PBS training.

There are a number of factors to consider in relation to these results. There are few published studies available for GCPLA comparison in order to put the results of the current study in context. Jung's unpublished master's thesis (2013) reviewed the use of the GCPLA and found six published studies which used it, one of which had used a modified version; some other studies have used questionnaires based on the GCPLA but these are too different for useful comparison (e.g. Leyin, 2008). Jung's thesis however, did include data on range and busy scores from three unpublished datasets; data from the NHS Learning Disability Team in Hastings was on 26 service users and had a range score of 20.93, and a busy score of 11.41; data from 35 service users from a PhD student in Scotland had a range score of 21.89 and a busy score of 12.51; and data on 29 people with learning disabilities from a staff member in the Tizard Centre where Jung's study was based, had a range score of 26, but did not include busy scores.

For published studies that did use the GCPLA there is sometimes limited opportunity for comparison, for example, Pilling et al (2007) used a postal questionnaire to explore the experiences of pupils at residential special schools; questions on access to services and use of the community were based on the GCPLA. No data are reported about range or busy scores, so these cannot be used for comparison; however the study found that participation in leisure and community was greater than in Baker (2000) but the use of public transport was lower. Watson et al (2008) used the GCPLA in their evaluation of a project supporting young adults with learning disabilities; however they primarily used the questions around being accompanied or not as they reported individuals found it difficult to answer regarding frequency of activity. There are therefore no useful data in their study to compare with the current study.

Baker (2000) included some normative data when he first published the measure; this gave a service user mean range score of 18 for 38 service users with a range of learning disabilities, and a staff score of 24 as a comparison, for 41 staff members. Service users' data was either collected via self-report or completed on their behalf. The study also demonstrated that the range of service users' activities was smaller than that of staff's and that these were mostly undertaken with staff, rather than alone or with friends. Service users were also less busy, though this was not a significant difference.

Abraham et al (2002) investigated the relationship between self-esteem and community participation, using a modified version of the GCPLA. They used interviews for self-report and modified the GCPLA by the addition of pictorial symbols to represent activities. They reported a range score of 20.3, and a busy score of 11.2. This study also reported figures from an unpublished MSc dissertation (Holmes, 1994), with a range score of 16.7 and a busy score of 8.3.

Baker's follow-up study in 2007 evaluated impact of moving from hospital to community-based support for 60 people with learning disabilities; regression was also used in order to identify factors potentially influencing use of community. The GCPLA was completed by either the key worker or manager, and the study reported the resettlement group had a significantly increased range score and community score at T2, compared to the control group; however there was no significant change in home-based activity. The mean range score post-discharge was 16, but no other detailed figures are included in the reported data; this was between the 25th and 50th percentile for the learning disability data in his 2000 study. The current study reported a T2 mean of 20 for both experimental and control groups, which was between the 50th and 75th percentile of the learning disability data in Baker (2000).

One additional study was found in addition to these; Allen et al (2011) reported on their PBS Clinical Practice and Outcomes Project, which provided outcomes from specialist behavioural teams using PBS. They found a significant increase in range score from 18.8 at T1 at point of referral, to 21.1 at T2, at point of discharge from the behavioural team's caseload.

From the limited data available, the pre training range scores in the current study appear to be slightly higher in comparison to other studies; this may be a contributory factor to the lack of change post-training. This information is summarised in Table 10.5.

Table 10.5: GCPLA Comparisons at T1 and T2 for Range and Busy

Study (n)	GCPLA Range Mean (Range)		GCPLA Busy Mean (Range)	
	T1	T2	T1	T2
Current Study Experimental (50)	20.00 (3-43)	20.08 (3-38)	12.34 (2-25)	12.36 (1-25)
Current Study Control (22)	20.09 (3-41)	20.23 (6-36)	12.00 (1-22)	12.82 (4-23)
Allen et al, 2011 (38)	18.8	21.1	-	-
Hastings NHS (26)*	20.93	-	11.41	-
Scotland PhD student (35)*	21.89	-	12.51	-
Tizard Centre (29) *	26	-	-	-
Baker, 2007 Experimental (26)	-	16	-	-
Abraham et al, 2002 (50)**	20.3		11.2	
Baker, 2000 Service Users (38)	18	-	11.3	-
Staff (41)	24	-	13.5	-
Holmes (1994) *	16.7	-	8.3	-

* Unpublished data; **modified GCPLA

Both Baker studies (2000 & 2007) considered association between range score and adaptive behaviour and challenging behaviour. In both studies an association with ABS scores is reported, but not with challenging behaviour as measured via the Behaviour Problems Inventory (BPI) (Rojahn et al, 1989). In the 2000 study there was a significant positive correlation between GCPLA range and ABS scores, and a non-significant negative correlation between GCPLA range and BPI scores. In the Baker 2007 study these results were repeated with GCPLA range positively correlated with ABS, but not significantly correlated with BPI, thus indicating that level of ability has an association with use of the community and leisure activity, but level of challenging behaviour does not.

In the current study these results were different; GCPLA range was not correlated with ABS; however there was a significant negative correlation between GCPLA range and ABC total at T2, and also between the busy subscale and ABC total at T2 for the experimental group only, indicating that service users with the highest levels of challenging behaviour, used the community less often and participated in activities less regularly.

This relationship between challenging behaviour and use of the community is not surprising since that is one of the factors in the definition of challenging behaviour used earlier (section 1.3.1); ‘behaviour of such an intensity, frequency or duration as to threaten the quality of life and/or the physical safety of the individual or others and is likely to lead to responses that are restrictive, aversive or *result in exclusion* (Royal College of Psychiatrists et al, 2007, p.10) [emphasis added]. The lack of association between adaptive behaviour and use of the community is unexpected, since this is reflected in a range of previous research (Baker, 2000; Baker, 2007; Emerson et al, 2000b; Stancliffe & Lakin, 1998).

10.4 Maintenance of Training Effects (T1-T3)

10.4.1 Challenging Behaviour

10.4.1.1 Aberrant Behaviour Checklist

This study found an average reduction to 69% of the baseline from T1-T3 in mean ABC total scores, less of a reduction than from T1-T2, but still a significant change overall. 44% also had a positive reliable change, although 2% also had a negative reliable change.

As other relevant comparison studies use ABC median, it is helpful to consider this also; the current study found a reduction to 64% of the baseline in median ABC total scores. Tyrer et al (2008) reported reduction to 59% of the baseline from T1-T3 with placebo, reduction to 54% of the baseline with Risperidone, and to 48% of the baseline with Haloperidol. However their T3

data was only 12 weeks after baseline, so markedly shorter than in the current study. Hassiotis et al (2009) reported data at baseline, 3 months and 6 months in their randomised control trial into the impact of PBS interventions carried out by a specialist health team, and they report a reduction to 57% of the baseline at follow-up. Hassiotis et al (2011) also carried out a two-year follow-up to this study; they reported significantly lower ABC scores for the experimental group, but do not include raw data so cannot be included here for comparison. Most trials for interventions into challenging behaviour have data for less than 6 months (Hassiotis & Hall, 2008), so the current study is unusual in that there is around 20 months between baseline and follow-up data collection. This information is summarised in Table 10.6.

Table 10.6: ABC Comparisons T1- T3

Study	ABC Total (Median)			
	T1	T2	T3	Reduction to % of baseline
Current Study	55	27	35.25	64
Tyrer et al, 2008	51	21.5	29.5	59
Hassiotis et al, 2009				
Experimental	36	25.5	20.5	57
Control	47	40	41	87

10.4.1.2 Behaviour Recording Forms

Using BRF this study reported an average reduction to 67% of the baseline and there was a reduction to less than 30% of the baseline for 50% of the group. Although there was a slight increase in frequency from T2-T3, the decrease from pre training to follow-up was still a significant change. McClean et al (2005) reported on outcomes from PBS training and found a reduction in frequency to less than 30% of the baseline for 77% for the 138 individuals at follow-up an average of 22.5 months after implementation. Grey & McClean (2007) also reported on outcomes from PBS training and at six months follow-up they found a reduction in frequency to 11% of the baseline for 19 of the participants, and McClean & Grey (2012) reported a reduction in frequency to 18% of the baseline at follow-up up.

For severity as measured by the BRF there was a reduction to 58% of the baseline from T1-T3, and 36% of the group achieved a reduction to 30% of the baseline. Unlike frequency, severity decreased from T2-T3, but this was very slight and there was no significant change in severity as measured using BRF. McClean & Grey (2012) reported a reduction in severity to 60% of the baseline at follow-up using the Challenging Behaviour Rating Scales (CBRS), but neither of the two other comparison studies include severity data at T3. This information is summarised in Table 10.7.

Table 10.7: Comparative Reductions to Percentage of Baseline from T1-T3 Using BRF

Study	Reduction to % of Baseline T1-T3 for Frequency	Reduction to % of Baseline T1-T3 for Severity
Current Study Experimental	<30% (for 50% of the group) 67% (average reduction)	<30% (for 36% of the group) 58% (average reduction)
McClean & Grey, 2012 (31)	18% (average reduction)	60% (average reduction)*
Grey & McClean, 2007 (19)	11% (average reduction)	-
McClean et al, 2005 (138)	<30% (for 77% of the group)	-

* using CBRS

10.4.1.3 Momentary Time Sampling

With an MTS approach to recording challenging behaviour, this study reported a reduction in challenging behaviour from a mean of 3.16 to 1.77 at T3, a reduction to 56 % of the baseline; however this change was not significant. Although Jones et al (1999) and Stancliffe et al (2007) both reported three time points for MTS data, neither of them included detailed data on challenging behaviour, therefore no comparisons can be made, although Stancliffe et al (2007) note that there was no significant difference in challenging behaviour. Totsika et al (2010) reported observed challenging behaviour increasing by T3, following a reduction at T2. This information is summarised in Table 10.8.

Table 10.8: MTS Challenging Behaviour Comparisons T1-T3

Study (n)	T1	T2	T3	Reduction to % of Baseline T1-T3
Current study (50)	3.16	2.86	1.77	56
Totsika et al. 2010 (21)	5.89	2.90	7.52	128

10.4.2 Quality of Life

10.4.2.1 Engagement in Activity

This study found no significant difference in engagement from baseline to follow-up for service user engagement, although there was a slight increase (9%). Other studies reporting on engagement at three time points are Totsika et al (2010) who also reported no significant change, and Stancliffe et al (2007) saw a 27% increase from T1-T3. Jones et al (1999)'s study on implementation of active support saw a substantial increase of 73% from T1-T3. This information is summarised in Table 10.9. As the levels of assistance from staff to service users in the current study do not change significantly from T1-T3, perhaps it is not surprising that engagement does not change significantly either, as we know from the literature that the two are often associated (Felce & Emerson, 2004, p.361-2).

In terms of the different types of engagement, there was a significant increase in non-social engagement from a mean at T1 of 43.76% to a mean of 60.12% at T3. Social engagement also changed significantly and reduced from a mean of 26.76% at T1 to 16.97% at T3. This reduction in social engagement may be linked to the increase in active support, as service users become more engaged in activity rather than socially; this potential link between active support provided by staff and service users' engagement will be explored further in chapter 11.

Table 10.9: Comparisons for Engagement T1-T3

Study (n)	T1	T2	T3	% Increase T1-T3
Current study (50)	70.52	76.02	77.08	9
Totsika et al. 2010 (21)	41.13	37.49	41.81	2
Stancliffe et al, 2007 (22)	42.46	49.54	53.81	27
Jones et al, 1999 (19)	33.1	53.4	57.2	73

10.4.2.2 Participation in Community and Leisure

This study found no significant change in range or busy scores from T1-T3, although scores do increase slightly at T2 and increase again at T3. Only one study was found that reported using the GCPLA at more than two time points. Hassiotis et al (2012), in their study linked to Hassiotis et al (2009) which reported on the impact of PBS input from a specialist health team, report that both the experimental and control groups increased their activities by follow-up of 24 months with median range score of 19 for the experimental group and 18 for the comparison group; follow-up median range score in the current study was for 20 the experimental group but this was unchanged from T1. As the Hassiotis et al (2012) study does not report detailed data at all three time points, it is difficult to make any further comparisons.

There is very little change in any other of the GCPLA subscales from T1-T3 in the current study; given the emphasis within the training on developing support plans based on participation in activity, it is perhaps surprising to see no change at all in the mean score for GCPLA home-based; however this is similar to the lack of change in domestic engagement as measured via MTS, despite having an active support emphasis within the training.

10.5 Chapter Summary

This study has demonstrated significant reductions in challenging behaviour following PBS training, which are largely maintained at follow-up; these are shown both by the main outcome measure the ABC, and also by the BRF. However, due to study design and limitations in the measures used, some caution must be applied to these results. The lack of significant change in

quality of life measures is in contrast to these apparent significant improvements in challenging behaviour. Although it is positive to note that the improvement in challenging behaviour has been achieved without any detrimental impact on engagement or community participation, it is perhaps surprising that this significant improvement in behaviour has not resulted in an associated robust and significant improvement in quality of life measures. Consideration was given to the option that the lack of change in the GCPLA was due to the measure not being sensitive enough; however given that the same lack of change was also noted within service user engagement as measured by MTS, then this seems an unlikely explanation. Perhaps it is more likely that for both engagement and for participation in community and leisure activities, there is a ceiling effect as both these measures indicate high scoring prior to the training.

The modest (non-significant) increases in engagement and in participation over time may indicate that quality of life factors are beginning to change. It may be that the decrease in challenging behaviour took time to effect any change in service users' level of activity; it may be that as time passed and the reduction in challenging behaviours became more of an ongoing fact, this impacted on staff behaviour in some way (perhaps confidence) and increased the likelihood that they would support service users to participate in activities more often. The potential links between staff and service user outcomes will be explored in the next chapter.

11 Overall Discussion

11.1 Chapter Outline

This chapter will reflect on the whole study and its results. Firstly the chapter will present a brief summary of results along with some of the limitations of these. Following this, the chapter will return to the original theory of process of change as presented in chapter 4, and will reflect on how far this theory came true in reality for both challenging behaviour and quality of life. Updated models of change will be presented, demonstrating the *actual* process of change, and links will be explored between results from managers and staff, and results for service users. The chapter will finish with concluding comments, including critical evaluation of measures used in the study, implications for: our understanding of Positive Behaviour Support (PBS), for PBS training, for a framework to implement PBS, and for future research.

11.2 Summary of Results

The main findings from this study are potentially ambiguous in that it appears that PBS training for managers resulted in significant reductions in service users' challenging behaviour, and that these significant reductions were maintained over time; however,, there are a number of limitations to the study in terms of design and measures used, and therefore it cannot be concluded with confidence that there have been reductions in behaviour. These limitations have been outlined elsewhere but include: lack of procedural validity in terms of the actual training delivered; potential of bias given this was an internal evaluation; lack of a randomised control group; limitations in the measures used, particularly in relation to the measurement of challenging behaviour, and lack of internal reliability in some measures. Having acknowledged these limitations however, the data do show a reduction in challenging behaviour and therefore this chapter will consider that and possible implications arising from any reduction.

Regardless of ambiguity in relation to challenging behaviour, it is clear that the training did not result in any significant improvement to service users' quality of life, either in terms of engagement or in use of the community and participation in leisure activities. Subsidiary findings are that although managers' knowledge of PBS increased, their attributions about behaviour did not change, nor did the levels of practice leadership they provided to their staff. For staff, there were no changes in any of their measures, apart from on the Active Support Measure (ASM) which increased significantly by follow-up; there were no changes to staff knowledge, attributions or support provided to service users (assistance or positive contact). The study also saw

considerable dropout by T3, when only 44% of managers had completed the training and were still managing support for the same service users.

11.3 Theory of Process of Change

As outlined in chapter 4, this study was based on a theory of change which expected that PBS training for frontline managers of social care services would have an impact on not only them and their performance in their role, but also on their staff teams, and on the service users that they support, impacting on both their challenging behaviour and their quality of life. Specifically, it was expected that both managers and staff would increase their knowledge of PBS and that this would influence both groups' understanding of challenging behaviour as evidenced by a change in their attributions. Both of these factors were expected to lead to better person-centred PBS plans being developed, which were based on an understanding of the functions of behaviour and which used positive and helpful methods to address these. The increasing implementation of these improved support plans would be indicated by the Periodic Service Review (PSR).

During the training, it was also expected that managers would increase their understanding of practice leadership and that evidence of increased use of practice leadership would be demonstrated by managers observing their staff more regularly. This observation would have a focus of supporting service users to engage in activity, with managers giving feedback to their staff based on these observations. This increased practice leadership was expected to lead to better support to service users, particularly in terms of assistance to participate in activity

This improved support to service users was expected to have a number of positive outcomes; firstly, it was expected that there would be increased engagement in activity and increased participation in leisure activities and use of the community. In addition to these quality of life outcomes, it was expected that challenging behaviours would reduce as a result of the improved staff support and the well-implemented PBS plans.

However, the actual process of change was somewhat different from what was anticipated; this is outlined in Figure 42 in relation to challenging behaviour, and Figure 43 for quality of life, although it is acknowledged that there is overlap and interaction between the two figures. These figures show the changes that actually took place and provide an overview of how these fit with the expected theory of change in relation to the three categories of participants: managers, staff and service users. Further discussion of these figures and the actual change process is below. It would have been useful to have more systematically tested the theory of process of

change via structural equation modelling which would have investigated the relationships between the three variables, that is, the groups of managers, staff and service users. However, this was beyond the scope of this thesis.

11.4 Actual Process of Change for Challenging Behaviour

The process of change for challenging behaviour is presented in Figure 42; this shows that reduction did occur in service users' behaviour as expected, following PBS training for managers although it is acknowledged that these results must be treated with caution. The changes appear to be less linear than expected; the process of changes for managers, effecting changes for staff, and therefore bringing about changes for service users, does not appear to be as clear as may have been anticipated. Each of these groups will be considered in terms of their relationship with the model of change for challenging behaviour and with the results seen in this study.

For managers, all three areas did change as expected, that is, their knowledge increased, they developed new PBS plans based on function and they cascaded these plans to their staff to implement. The importance of implementation of PBS plans as measured by the PSR was as expected; Aberrant Behaviour Checklist (ABC) total was negatively correlated with PSR at T2, indicating that the greater the implementation of PBS, as indicated by higher PSR scores, the greater the decrease in challenging behaviour, as indicated by lower scores on the ABC. None of the published studies using PSR have considered this relationship, but McGill et al (in press) noted this association, in that the change in ABC score from T1-T2 was significantly correlated with PSR score.

However there were some issues which arose that were unanticipated; one of these was the high level of management turnover. Staff turnover in social care settings has been shown to be an issue (e.g. de Kock et al, 1987; Felce et al, 1993; Hatton et al, 2001; Jones et al, 1999) and this was one of the pragmatic reasons for choosing to focus training on managers rather than direct care staff, as it was presumed that turnover would be low. However this proved not to be the case and in the course of the 20 months of the data collection period of this study, the turnover in the managers was 56%. These managers had either left the organisation or the service where they were working at the time of the PBS training (44% $n=22$), or the service user had left the organisation or died and therefore the manager no longer supported them (12% $n=6$). In relation to challenging behaviour it seemed likely that the significant turnover in management would have an impact on the long-term maintenance of PBS, since the main person with the training had moved on; however there was no significant difference between the group whose managers had changed and the group with stable management in terms of ABC scores at T3; this is also somewhat surprising.



Manager	Staff	Service User
<p>1. YES – PBS knowledge increased significantly, from 71% to 83%.</p> <p>2. YES – five new PBS protocols were developed for each service user as part of the training, all based on function.</p> <p>3. YES – PBS plans cascaded to staff, evidenced by PSR score of 66%; also ASM Q15 (‘written plans are in routine use’) increased by 69%.</p>	<p>4. NO – no significant change in staff PBS knowledge.</p> <p>5. NO – no significant change in staff attributions about challenging behaviour.</p> <p>6. YES – PBS plans were followed to a limited extent, as evidenced by PSR score and by increase in Q15 on ASM.</p>	<p>7. NO – no significant change in positive contact or assistance from staff.</p> <p>8. YES – PBS plans are implemented, evidenced by PSR score and by Q15 on ASM.</p> <p>9. YES – evidenced by significant reduction in ABC (total and severe) and BRF (frequency) however limitations to the study potentially undermine these results.</p>

Figure 42. Actual Process of Change for Challenging Behaviour

Another element of the managers' part of the model which was perhaps not as expected, was the role that practice leadership appeared to play in reducing challenging behaviour; according to the original theory of change, practice leadership was expected to be a factor in quality of life changes, rather than specifically for challenging behaviour. However, practice leadership was significantly correlated with ABC scores at T2. In relation to the staff areas within the model, these do not change as expected and to some extent the expected linear process of change in managers, leading to changes in staff, breaks down. There was no change in staff knowledge of PBS and likewise no change in their attributions, both of which were expected to provide increased motivation to follow the PBS plans. The level of implementation of PBS plans is reflected by the PSR score of 66% at T2; due to the lack of research evidence on the PSR, it is difficult to be clear about the impact that a score of 66% would have on service user outcomes. In terms of the La Vigna standard this would not be regarded as a satisfactory score. However, the score does indicate that two thirds of the behavioural and operational advice was being followed, so it seems likely that even this limited implementation would have had some sort of positive impact on outcomes for service users. It therefore appears that changes in knowledge and attributions were not necessary to facilitate a limited implementation of PBS plans. The ASM also demonstrated that there was substantial increase in use of written support plans, an increase in score of 69% for question 15 on the ASM ('written plans are in routine use'); although this increase was not significant when analysed along with the control group. It therefore appears that the main staff characteristic that was relevant for reducing behaviour was their implementation of PBS plans; their knowledge and beliefs did not change and therefore appear not to have played any significant role.

Finally, the figure shows that as far as can be evaluated in terms of the limitations noted service users' challenging behaviour reduced in line with expectations; however this was not due to more positive support from staff as neither assistance nor positive contact increased significantly, so this was also not as expected. It was foreseeable that staff assistance could affect the rate of challenging behaviour and that this relationship could go in two different directions. Increased assistance may be experienced as increased demands, therefore triggering escape-related challenging behaviour; however assistance would be likely to reduce attention-related challenging behaviour, as service users would experience the assistance as positive attention which allows them to engage socially and in activity, thus reducing challenging behaviour related to an attention function (although perhaps this type of challenge would be less likely in the current study due to the individual services, with 1:1 staff support). If the functions of each service user's behaviour had been included as part of the analysis, this would have been a useful

addition and may have allowed a better understanding of the impact of assistance on the different functions of behaviour; this may be a useful area for future research. However, in the current study there was no significant relationship between staff assistance and levels of challenging behaviour, at any of the three time points, so, overall, assistance appears not to have been a contributory factor in the reduction in challenging behaviour. This is surprising, particularly since assistance and challenging behaviour are often correlated and research has found that staff give less attention to service users who are more challenging (e.g. Felce & Perry, 1995). Considering this from a different direction, it was also surprising that decreased challenging behaviour did not lead to a corresponding rise in staff assistance.

It is therefore difficult to fully explain the change in service users' behaviour, particularly in terms of the staff measures; it may be that staff knowledge and attributions are less important in leading to change in challenging behaviour than has been previously thought. It may be that all that is required is for staff to follow the PBS plans, and that as long as these are implemented to a limited degree, then behaviour change for service users will follow.

11.5 Actual Process of Change for Quality of Life

The process of change for quality of life is shown in Figure 43; this presents a somewhat different picture from the change in challenging behaviour in that there is no evidence of changes in quality of life (QoL), and therefore the process of managers' PBS training effecting service users' quality of life, did not occur as expected.

For managers, it was expected that there would be an increase in their practice leadership following the training; however this did not occur and therefore perhaps the model of expected change breaks down immediately. There is some evidence of increased focus on service user engagement, as during managers' observations of their staff, focus on engagement increased by 85% and during supervision, focus on engagement increased by 64%; however neither of these changes was significant. Practice leadership however does appear to be a factor in relation to quality of life outcomes in this study, as high practice leadership scores are associated with better active support and with higher levels of engagement. Practice leadership was also significantly positively correlated with busy scores on the Guernsey Community Participation and Leisure Assessment (GCPLA), indicating that the services with better practice leaders were likely to have service users who participated more in activity. The association does not indicate causation, however it seems unlikely that greater participation in activity would result in better practice leadership, or that staff providing good active support would cause their managers to demonstrate good practice leadership; intuitively it is more likely that the causation is in the opposite direction and that good practice leaders support their staff to provide better active

support, and to involve service users more in activities. This reflects other studies which have seen practice leadership associated with better service users' outcomes, for example, Beadle-Brown et al (2014) examined the effect of practice leadership on the implementation of active support and found it to be significantly associated with improved outcomes for service users due to the increased implementation of active support, and Beadle-Brown et al (2015a) also found correlation between practice leadership scores and service user engagement.

As noted earlier, management turnover was a major factor in this study. Comparisons between the group whose managers had changed and those with a stable management showed that there was a significant difference in their practice leadership scores, with the stable management group having a higher score. Given that practice leadership was associated with many other factors, then it would be expected that these factors may also be impacted by the change in management. This proved to be the case as there were also significant differences between the stable management group and those who had new managers by T3 in terms of staff assistance and active support to service users.

For the staff part of the model, assistance to service users does not change at all, so this is an important element of the model of change not proceeding as expected. Scores on the ASM do increase significantly by T3, and the fact that the ASM only changed over time and not immediately following training is worth noting. One potential explanation may be that as staff experienced less challenges from service users over a period of time, that this increased the likelihood of their providing active support, perhaps due to an increase in confidence as challenges were reduced. Or, it may be that staff felt more positive towards service users because they were experiencing less challenging behaviour, and therefore were more likely to support them to participate. Jones et al (2013) discussed the impact of challenging behaviour, even stereotypical challenging behaviour which may be seen as not particularly challenging, but can still impact on staff perception of service users, which in turn effects the degree to which they are likely to interact with them positively (Hastings & Remington, 1994b). In this sense, the increased active support may be seen as a *result* of reduced challenging behaviour, not a *cause* of reduced challenging behaviour.

The ASM was also significantly positively correlated with engagement. This association has been found previously in the literature (e.g. Mansell & Beadle-Brown, 2012, chapter 3) However in the current study, despite the ASM increasing significantly over time, there was no change in service users' engagement.



Manager	Staff	Service User
<p>1. NO – observation does not increase, but was already regular (82% at least monthly); 60% increase in modelling good support, but not significant.</p> <p>2. NO – during observation 85% increase; in supervision, 64% increase; however, changes are not significant and in team meetings 20% reduction.</p> <p>3. YES – 4 new proactive PBS protocols all with QoL emphasis (e.g. improve communication).</p>	<p>4. NO – 19% increase in feedback at least usually; 21% increase in ASM Q13 ('staff work as a coordinated team to support service users'), but neither significant.</p> <p>5. NO – no significant increase in staff assistance to service users, although this is already high.</p> <p>6. YES – PBS plans were followed to a limited extent, evidenced by PSR score and by increase in Q15 on ASM.</p>	<p>7. NO – no significant increase in assistance from staff.</p> <p>8. NO – no significant increase in engagement; although non-social engagement increases and social engagement decreases.</p> <p>9. NO – no change in GCPLA busy or range.</p>

Figure 43. Actual Process of Change for Quality of Life

The level to which PBS was implemented as measured by the PSR, was the only element of expected changes for staff which occurred as anticipated and this was not to the expected level, as the PSR score only indicated a limited implementation. The PSR was associated with a number of factors directly relating to outcomes for service users; it was significantly positively correlated with both engagement and GCPLA range scores, indicating that higher level of implementation of PBS plans is associated with better quality of life outcomes; however, despite this limited level of implementation, there was no associated improvement in service users' quality of life.

In the service users' part of the model, none of the anticipated changes took place; there was no significant change in assistance from staff, and no significant change in engagement. There were some changes in engagement in that non-social engagement increased significantly from T1-T3 and social engagement reduced significantly; this may be due to the increase in the levels of active support provided by staff so that service users were more engaged in participating in activity rather than only in social interaction.

Finally, there was no significant change in use of the community or participation in activity. It is surprising that despite the PSR being implemented at 66% and a 69% increase in routine use of written support plans, the impact these factors had on challenging behaviour is not replicated here for changes in quality of life. Given that in the new PBS plans written for each person, four out of the five protocols were proactive with an emphasis on quality of life, this is difficult to explain. It appears counterintuitive that the implementation of a number of proactive PBS protocols, with a focus on quality of life, would produce significant reduction in challenging behaviour but have no impact on quality of life. It appears that in terms of the findings from this study, there are some elements which are difficult to explain. It is clear however from the literature, particularly the studies discussed in chapters 2 and 3 of this thesis, that difficulty in evidencing changes in quality of life is a common theme in PBS studies.

11.6 Critical Evaluation of Measures Used in the Study

Following the conclusion of this study, there are a number of factors that are worth considering in relation to the measures. Measuring behavioural change presents a challenge as it is difficult to do this blind, and therefore there is potential for bias. Both the ABC and the BRF were completed by either managers or staff who knew that they were in the experimental group and therefore this casts some doubt on these data. Data in relation to challenging behaviour was also taken during the observation periods, but occurrence of this was at too low a level to be usefully analysed. However it is worth noting that data from both the ABC and BRF follow an expected pattern, that is, the behaviour reduces following training and then begins to increase again. In addition, the two measures were completed by different groups, and yet both follow a similar pattern. It may be expected that

managers in the experimental group would demonstrate bias when completing the ABC measure, as they may wish to show a positive impact from the training they had undertaken; intuitively it seems less likely that staff would be motivated in the same way when deciding whether or not to complete a behavioural recording form following an incident of challenging behaviour.

It is also worth noting the if bias was a factor in relation to the measures completed for challenging behaviour, then this same bias did not impact in terms of quality of life measures; the GCPLA was also completed by managers and showed no significant change throughout the study. It appears likely that had managers in the experimental group wished to demonstrate positive bias on the GCPLA, then this would have been easy to do, as the measure is straightforward to complete and it is clear how to score it positively.

Other than issues of bias, there are a number of other factors to reflect on in relation to the measures used in this study. The CHABA showed no significant change for either staff or managers in any sub-scale at any time point. It may therefore be judged not to be the most appropriate or useful measure for measuring staff attributions. The GCPLA although it has a level of fine detail not available in some comparative measures, does not have a section in relation to domestic activity within the house, and this may have been a factor in this study, as non-social engagement did increase significantly over time, perhaps as a result of the emphasis on domestic activity in relation to Active Support.

Finally, it is worth noting the low levels of internal reliability found at T1 of this study within the ABC and CHABA. The ABC which is the main outcome measure had a Cronbach's alpha of 0.49, regarded by some as unacceptably low; and values for the CHABA subscales Learned Negative, Biomedical, Emotional and Physical Environment fell below 0.6, and would therefore be regarded as poor reliability.

11.7 Concluding Comments

11.7.1 Implications for Our Understanding of Positive Behaviour Support

This study has demonstrated that PBS training may have resulted in significant changes in challenging behaviour which were maintained over time; however, due to study limitations, it is not possible to have complete confidence in these results. It is clear that overall impact on quality of life was disappointing; it is true that engagement started high and therefore perhaps had little opportunity for increase, and it is also true that there were modest increases in both engagement and participation, but despite these factors there was very little impact on quality of life. This is perhaps especially surprising given that PBS has such a focus on quality of life and indeed regards quality

of life changes as equally, if not more important than changes in behaviour; a better life for people with learning disabilities and challenging behaviour is the fundamental premise of PBS.

It is difficult to be clear whether the lack of change in quality of life is because of a ceiling effect, that is, high-scoring measures at baseline which effectively had little room for improvement, or whether the intervention, the PBS training, was not effective in achieving any quality of life changes. With hindsight there are a number of adaptations to the study that may have helped to better identify quality of life changes. The observation periods could have been longer; this would have allowed a more thorough assessment of how and if support had changed, and would also have allowed a more established baseline for observed behaviours. Methods other than observation could have been used; perhaps interview with family carers could have helped identify what people important to the service user thought had changed following the training. Longer term follow-up would also potentially have given some clearer indication of any changes; for individuals with learning disabilities and complex needs, it would perhaps not be expected to see lifestyle changes take place quickly. It may be that quality of life changes take longer to achieve than changes to challenging behaviour, and therefore perhaps that is why they are not evidenced in most of the literature, as most studies considering impact of PBS training do not have long-term data collection. In connection with this point, it is worth noting that in the current study there were interesting differences in how changes took place in the two areas, challenging behaviour and quality of life. Challenging behaviour appeared to improve following training and then over time began to increase slightly, presumably as the training effects began to wear off. However for some aspects of quality of life (total engagement and busy), there were small improvements at T2 which were followed by further small improvements at T3, so the training effects increased as time went on, similar to the pattern of the ASM. These changes are not significant, but it would have been an interesting addition to this study to have had a fourth data collection in a further six months.

And yet in many ways, this result in relation to quality of life is not surprising; both the review of PBS effectiveness in chapter 2 and the literature review in chapter 3 come to a similar conclusion: that the evidence for PBS impacting on quality of life is limited. It is challenging to understand why this is the case, and there really only appear to be two alternatives: either PBS is not able to impact quality of life in reality as it is expected to do in theory, or, there is an impact on quality of life but the measures being used are not capturing this change.

The latter possibility was partly the premise of this study; that measures were required which would be sufficiently fine to capture the level of changes that might be expected from a service user group with severe learning disabilities, limited verbal communication and a range of other complexities such as autism and challenging behaviour. Thus observational methods were chosen as

more likely to evidence changes in behaviour and changes in presentation than other types of evaluation such as questionnaires or interviews. Observation of people with learning disabilities has been used in research over the past 50 years, and there is a host of literature about how best to do this, what methods to use, when it should be done, how to ensure data is reliable, and so on. However, despite all of this guidance, perhaps observational research is still not addressing the fundamental question of how do we understand and evidence quality of life changes for this particular client group? With particular reference to the current study, it is also pertinent to note that only a very small area of quality of life was even addressed. If the eight domains of quality of life are taken as the current accepted definition (Beadle-Brown et al, 2009b; Schalock et al, 2002; Wang et al, 2010), then this study attempted to address only two of these domains: interpersonal relationships and social inclusion, as indicated by level of engagement (both in activity and social engagement) and by use of the community. The remaining domains of emotional well-being, material well-being, personal development, physical well-being, self-determination, and rights, were not assessed in this study, and yet they are arguably equally as important as the two that were evaluated.

The other possibility is that PBS does not have an impact on quality of life and in many ways that is a tempting conclusion, not just based on the findings in the current study, but also based on the findings from the literature review. However, there is an intuitive belief which is difficult to disregard; that supporting people with learning disabilities in a more person-centred way, providing better and more active support, developing PBS plans based on an understanding of their behaviour and what it means for them, *must* have an impact on how the individual experiences their life, and must result in an experience of support that is better, happier, and more ‘quality’. PBS is fundamentally based on a person-centred approach, where the service user will receive support more suited to their needs, personality and ambitions. It brings together important people in their life in order to best identify these, and then it systematically develops behavioural approaches to implement this in practice. At every stage it is tailored to the person, based on their presentation, and continually evaluated to check its suitability to meet their needs and address their outcomes. With all of this as a part of PBS, it appears impossible for PBS not to have an impact on service users’ quality of life, and arguably the reduction of challenging behaviours are some indication of an improved quality of life. So far though, it appears we have not found the right tools to evidence this to the expected extent.

11.7.2 Implications for Training in Positive Behaviour Support

The training evaluated in this study was based on a specific premise: that training managers in PBS could have sufficient impact on the staff who provide the direct support so that their practice would

change in a way that impacts on the challenging behaviour of service users. Although there are limitations in study design which warrant caution in conclusions, there are indications that this model of training has been successful; no direct support staff attended training and yet challenging behaviour appeared to reduce significantly and remained significantly reduced at follow-up. This finding may have implications for the approach to training in PBS, and could suggest that training efforts are better directed at frontline management rather than direct support staff. This appears to provide a cost effective and more achievable approach to implementing PBS on a large-scale, particularly within large organisations where training coverage for all direct support staff is difficult to achieve, for example, in social care settings where turnover of direct support staff is so high. Many organisations try to provide an introductory one or two-day training in PBS to all direct support staff working with people with challenging behaviour; perhaps it would be more achievable, and, based on the findings from this study, potentially more effective, to focus more resources on in-depth PBS training to managers.

There was also substantial turnover in the managers within this study, meaning that less than half the group had ongoing support from a manager who had completed the PBS training. Many of the measures would have changed significantly if the per protocol approach had been adopted and it is possible that the lack of change in quality of life is related to this factor. It is also worth noting that if this level of turnover were to be extrapolated, then within five years of starting the training with 50 managers, there would be less than seven trained managers left. This level of turnover creates a number of difficulties; it makes it very difficult to carry out research into long-term impact of any intervention involving staff within social care settings, as the confounding factor of turnover may impact long-term results. It also more importantly, impacts on the service provided to the individuals living in these settings, and ultimately to their quality of life; long-term consistency of support may be difficult to achieve and it is likely that in these circumstances, service users will have to cope with entirely new staff teams learning all about them (e.g. their ways of communicating, the best way to support participation, the factors likely to serve as antecedents for challenging behaviours) over and over again, many different times over the years.

However, high turnover is a likely to remain a factor in care settings, and therefore PBS training has to address this, and needs to be effective even where there is substantial management turnover. One option would perhaps be to provide shorter PBS training (the training in this study lasted a year), and to provide this on a continuous 'rolling' basis, so that although training would be less in-depth, it would be provided more often and could therefore keep up with changes in management. Perhaps for the current world of social care, two six-month training sessions in PBS are more effective than a more in-depth training programme lasting a year, which may see substantial dropout before the training is even complete.

The lack of impact that PBS training has been able to evidence on quality of life, both within this study and within the literature as a whole, suggests that there are additions or changes needed to PBS training in order to achieve this. If we hold onto the belief as noted above, that PBS interventions *can* effect quality of life changes for individuals with learning disabilities, it appears that PBS training may need to change in order to evidence this. Allen et al (2005) noted that within the PBS model, quality of life is both an intervention and an outcome; perhaps this needs to be more clearly delineated in PBS training. One option would be that firstly we use quality of life interventions (along with behavioural interventions) in order to achieve the outcome of changing behaviour, then over time, we move to interventions with a more deliberate focus on quality of life as the outcome. This would mean that we view PBS training as being in two phases, firstly with a focus on reducing behaviour, and then with a second phase to more specifically address improving quality of life. It is likely that this would also fit with the priorities and motivation of most participants attending PBS training; quality of life changes are probably less of a priority for participants attending training, most of whom are looking for solutions to make the challenging behaviour stop, rather than coming from a starting point of wanting to improve the individual's quality of life. Once there is evidence that behaviour has been reduced, and staff are more confident and effective in supporting the person successfully, then perhaps that is the time to move onto a more deliberate focus in training on improving quality of life. This may feel uncomfortable for those committed to PBS as having a dual focus of reducing challenging behaviour and improving quality of life and it may feel like a watering down of the PBS model, or even a reverting to more old-fashioned behaviour change approaches. However, if PBS training cannot evidence significant, long-term and large-scale changes in service users' quality of life, then it appears necessary to consider some changes to how that training is delivered.

11.7.3 Implications for a Framework to Implement Positive Behaviour Support

The findings from this study, combined with a range of relevant literature noted below, have a number of implications for how PBS might be successfully implemented within large organisations.

The following are proposed as a helpful framework for embedding PBS within large organisations:

- Senior management support – system-wide implementation is reinforced by strategic support at a senior level, for example developing a PBS policy which makes the use of PBS mandatory, and which prohibits the use of any punishment-based approaches (Allen et al, 2012; Allen, 2011)
- Organisational investment – commitment of organisation resources, via the setting up of a specific PBS team is also an important element in ensuring the use of PBS, as this

gives staff and managers access to internal PBS expertise, without having to wait for access to external supports; it also gives a strong internal message of support to PBS as the organisation's method for addressing challenging behaviour (Allen et al, 2012; Allen et al, 2011; Allen, 2009)

- Organisation-wide implementation – the importance of ensuring competence across the whole organisation, so that implementation is systematic and comprehensive, not done on an ad hoc basis, dependant on the inclination of individual managers (Allen et al, 2013b; Allen et al, 2012; DOH, 2007)
- Practice Leadership – strong practice leaders with a commitment to PBS are important at all levels to ensure ongoing implementation of PBS, and to help encourage reflective practice (Allen et al, 2013b; Mansell and Beadle-Brown, 2012)
- Training – as evidenced in this study, PBS training is one element which can assist staff teams in implementing PBS and provide a framework for successful organisation-wide implementation (Allen et al, 2012; Denne et al, 2013; McGill et al, in press)
- Quality assurance – management monitoring, using a tool such as the PSR, is also useful in ensuring integrity of PBS implementation at a service level (Allen, 2009; La Vigna et al, 1994)
- Person-centred values – having clear and shared organisational values based on a person-centred approach are a useful first step to the implementation of PBS (Carr et al, 2002; Gore et al, 2013)
- Involvement of users and carers – having feedback from carers and where possible users, regarding the use of PBS and what is most successful and useful from their point of view, is important to achieve maximum implementation of PBS (Albin et al, 1996; Allen et al, 2013b; Carr et al, 1999)

11.7.4 Implications for Future Research

This study and the framework noted above have indicated a number of areas for potential future research. PBS is an approach to addressing challenging behaviour which has at its heart, the impetus to improve quality of life; quality of life improvements are the life-blood of what makes PBS, PBS. However, so far research evidencing quality of life changes linked with PBS interventions is noticeably limited, particularly for large-scale interventions. This therefore appears to be a significant

area for future research, to enable the field of PBS to evidence the quality of life changes which it sees as so central to its reason for existence. It would be particularly positive to see the use of rigorous designs, such as randomised control trials with blind assessors.

The lack of change in service user engagement in the current study is in contrast to many of the active support studies which have reported substantial increases in service user engagement, usually combined with increases in staff assistance. It may be that the lack of change in the current study was due to the lack of a focus on hands-on training or on-the-job coaching, and therefore if the training programme in the current study were to be re-run, it would be useful to add in more focus on active support. This would include on-the-job coaching for the managers and then training to enable the managers to provide on-the-job coaching for their own staff. Not only would this potentially allow for more of an impact on assistance and engagement than seen in the current study, it would also potentially support longer-term maintenance, if this type of practical coaching could become part of the day-to-day, expected practice within social care settings. Research focusing on a blended intervention combining PBS and active support, would therefore be a useful addition to the field. The two approaches would combine together well as they are both based on similar origins, applied behaviour analysis and person-centred values.

The current study found a number of areas where practice leadership was associated with positive changes both in staff behaviour and in outcomes for service users; a study focusing more specifically on practice leadership intervention would therefore be a useful contribution to the field of research in learning disabilities and behaviour support. At the moment, most of the research reporting on outcomes in relation to practice leadership come from the active support literature, and it would perhaps be useful to widen this out and explore an intervention with practice leadership at its centre. If we are to see long-term positive changes being created in the lives of people with learning disabilities and behaviour support needs, then it seems likely that a focus on practice leaders and creating long-term sustainable practice leadership within organisations, will be vital to achieving this; research exploring this would therefore be a helpful contribution.

Long-term maintenance of improvements in challenging behaviour remain a challenge for the sector, and so it would be useful to see longer-term follow-up of PBS interventions. The current study followed up 20 months after baseline and at that point the improvements in behaviour were already beginning to reduce; it would have been useful to have a fourth data collection point another six months later, but time constraints did not allow for that. This would have allowed us to further study and explore the factors supporting and undermining long-term change in challenging behaviour, particularly in social care settings, where the turnover of direct care staff is such an issue, and also as this study has demonstrated, where even frontline management changes very regularly.

The vast majority of people with learning disabilities in the UK now live in the community, most of them supported by social care service providers, and all government policy and good practice guidance indicates that this move from health settings to social care settings is likely to continue. Research focusing particularly on this area would therefore be a useful contribution to the field. Social care providers are sometimes poorly funded, often employ staff with very little experience, and frequently are commissioned to support combinations of service users who are poorly matched to live together and whose support hours are not based on assessed need; in these circumstances it would be very useful to have clearer evidence about the type of input and interventions which are most useful in creating and sustaining positive changes in the lives of people being supported by these agencies, particularly where this can also demonstrate a cost-effective element to any intervention.

PBS is developing as a field and currently within the UK (particularly in England) it has a high profile, being recommended in the UK government's guidance and, following the Winterbourne abuse scandal, being recognised as the most ethical, effective and suitable approach to supporting people with learning disabilities and challenging behaviour. The human rights agenda is now being seen as more relevant for people with learning disabilities and the scandals of out-of-area placements with their cutting people off from family, their high costs and their link with abusive practices are increasingly regarded as an unacceptable way forward in the support of some of society's most vulnerable citizens. This has been evident in a range of good practice guidance and government publications, and there is clearly an increasing emphasis on decreasing the use of restrictive interventions particularly those that have high risk of injury or greater harm attached to them and/or cause pain. PBS is currently being proposed as the approach to fill this gap, to support the reduction of restrictive intervention, to allow local care providers to develop their services so that everyone can be supported in their local community, to equip ordinary social care providers and families with the skills to support those with even the most challenging behaviours. It therefore appears to be a very prescient time for further PBS research in order to build the body of evidence and to create further momentum in achieving national sustainable changes in the lives of people with learning disabilities and behaviour support needs.

12 References

- Aarons, G. A., et al. (2012). Expanding the domains of attitudes towards evidence-based practice: The Evidence Based Practice Attitude Scale-50. *Administration and Policy in Mental Health and Mental Health Services Research*, 39(5), 331-340.
- Abraham, C., et al. (2002). Self-esteem, stigma and community participation amongst people with learning difficulties living in the community. *Journal of Community and Applied Social Psychology*, 12(6), 430-443.
- Ager, A. (1990). *The Life Experiences Checklist: Manual*. NFER-Nelson: Windsor.
- Albin, R., et al. (1996). Contextual fit for behavioural support plans: a model of 'goodness of fit'. In: Koegel, L.K., Koegel, R.L. and Dunlap, G. eds. *Positive Behaviour Support: Including People with Difficult Behaviour in the Community*. Paul Brookes, p.81-98.
- Allen, D. (2009). Positive behavioural support as a service system for people with challenging behaviour. *Psychiatry*, 8(10), 408- 411.
- Allen, D. (2011). *Reducing the use of restrictive practices with people who have intellectual disabilities*. Kidderminster: British Institute of Learning Disabilities.
- Allen, D., et al. (2013a). Towards the Prevention of Behavioural and Psychiatric Disorders in People with Intellectual Disabilities. *Journal of Applied Research in Intellectual Disabilities*, 26(6), 501-514.
- Allen, D., et al. (2013b). Implementing positive behavioural support: changing social and organisational contexts. *International Journal of Positive Behavioural Support*, 3(2), 32-41.
- Allen, D., et al. (2012). The impact of a whole-organisation approach to positive behavioural support on the use of physical interventions. *International Journal of Positive Behavioural Support*, 2(1), 26-30.
- Allen, D., et al. (2011). Assessing the effectiveness of positive behavioural support: The P- CPO Project. *International Journal of Positive Behavioural Support*, 1(1), 14-23.
- Allen, D. G., et al. (2007). Predictors, costs and characteristics of out of area placement for people with intellectual disability and challenging behaviour. *Journal of Intellectual Disability Research*, 51(6), 409-416.
- Allen, D., et al. (2005). Positive Behavioural Support: Definition, Current Status and Future Directions. *Tizard Learning Disability Review*, 10(2), 4-11.
- Allen, D. and Tynan, H. (2000). Responding to aggressive behavior: Impact of training on staff members' knowledge and confidence. *Mental Retardation*, 38(2), 97-104.

- Allen, D., et al. (1997). Changing care staff approaches to the prevention and management of aggressive behaviour in a residential treatment unit for persons with mental retardation and challenging behaviour. *Research in Developmental Disabilities*, 18(2), 101-112.
- Allen, P., Pahl, J. and Quine, L. (1990). *Care Staff in Transition: The Impact on Staff of Changing Services for People with Mental Handicaps*. London: HMSO.
- Aman, M.G. and Singh, N.N. (1986) *Aberrant Behavior Checklist Manual* New York: Slosson Educational Publications.
- American Psychiatric Association (1994). *The diagnostic and statistical manual of mental disorders-IV (DSM-IV)*. Washington, DC: American Psychiatric Association.
- Anderson, C. M. and Freeman, K. A. (2000). Positive behavior support: Expanding the application of applied behavior analysis. *Behavior Analyst*, 23(1), 85-94.
- Anderson, J. L., et al. (1996). A team training model for building the capacity to provide PBS in inclusive settings. In: Koegel, L.K., Koegel, R.L. and Dunlap, G. eds. *Positive Behavioral Support: Including People with Difficult Behavior in the Community*. Baltimore: Paul Brooks, p.467-490.
- Anderson, J. L., et al. (1993). Issues in providing training to achieve comprehensive behavioral support. In: Reichle, J. and Wacker, D.P. eds. *Communicative Alternatives to Challenging Behavior: Integrating Functional Assessment and Intervention Strategies*. Baltimore: Paul Brooks, p.363-406.
- Association of Chief Executives of Voluntary Organisations (ACEVO) (2015). *Winterbourne View – Time is Running Out*. London: ACEVO.
- Ashman, B., Ockenden, J. and Beadle-Brow, J. (2010). *Person-Centred Active Support: A Handbook*. Pavilion Publishing.
- Baer, D. M. and Wolf, M. M. (1987). Some still-current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, 20(4), 313-327.
- Baer, D. M., Wolf, M. M. and Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, 1(1), 91-97.
- Baker, D.J. (1998). Outcomes of behaviour support training to an agency providing residential and vocational support to persons with developmental disabilities. *Journal of the Association for Persons with Severe Handicaps*, 23, 144-148.
- Baker, P. (2007). Individual and service factors affecting deinstitutionalization and community use of people with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities*, 20 105-109.

- Baker, P. (2000). Measurement of community participation and use of leisure by service users with intellectual disabilities: The Guernsey Community Participation and Leisure Assessment. *Journal of Applied Research in Intellectual Disabilities*, 13, 1–18.
- Baker, P. and Allen, D. (2011). Editorial. *International Journal of Positive Behavioural Support*, 1(1), 3-6.
- Baker, P. and Shephard, J. (2010). Reflections on the periodic service review as a practice leadership tool in services for people with intellectual disabilities and challenging behaviour. *Tizard Learning Disability Rev*, 15(3), 29-32.
- Baker, P and Shephard, J (2005). The rebranding of behavioural approaches for people with learning disabilities and challenging behaviour, *Learning Disability Review*, 10(2), 12–15.
- Ball, T., Bush, A. and Emerson, E. (2004). *Psychological interventions for severely challenging behaviours shown by people with learning disabilities: Clinical practice guidelines*. Leicester: British Psychological Society.
- Bambara, L. M., Mitchellkvaack, N. A. and Iacobelli, S. (1994). Positive Behavioral Support for Students with Severe Disabilities - an Emerging Multicomponent Approach for Addressing Challenging Behaviors. *School Psychology Review*, 23(2), 263-278.
- Beadle-Brown, J. (2006). Person-centred approaches and quality of life. *Tizard Learning Disability Review*, 11(3), 4-12.
- Beadle-Brown, J., Bigby, C. and Bould, E. (2015a). Observing practice leadership in intellectual and developmental disability services. *Journal of Intellectual Disability Research*, Article first published online: 8 June 2015. Available from: <http://onlinelibrary.wiley.com/chain.kent.ac.uk/doi/10.1111/jir.12208/epdf> [Accessed 17August 2015].
- Beadle-Brown, J., et al. (2015b). Quality of life and quality of support for people with severe intellectual disability and complex needs. *Journal of Applied Research in Intellectual Disabilities*.
- Beadle-Brown, J., et al. (2014). Practice leadership and active support in residential services for people with intellectual disabilities: an exploratory study. *Journal of Intellectual Disability Research*, 58(9), 838-850.
- Beadle-Brown, J. Hutchinson, A. Whelton, B. (2012). Person-centred active support – increasing choice, promoting independence and reducing challenging behaviour. *Journal of Applied Research in Intellectual Disabilities* 25, 291-307.
- Beadle-Brown, J., et al. (2011). *Evaluation of active support over time*. Canterbury: Tizard Centre.

- Beadle-Brown, J., et al. (2009a). People with learning disabilities in 'out-of-area' residential placements: Views of families, managers and specialists. *British Journal of Developmental Disabilities*, 55(108), 15-31.
- Beadle-Brown, J., Murphy, G. and Di Terlizzi, M. (2009b). Quality of life for the Camberwell cohort. *Journal of Applied Research in Intellectual Disabilities*, 22(4), 380-390.
- Beadle-Brown, J., et al. (2008). The role of practice leadership in improving and maintaining quality in intellectual disability services. *Journal of Intellectual Disability Research* 52(8-9), 755.
- Beadle-Brown, J., et al. (2006). People with learning disabilities in 'out-of-area' residential placements: Reasons for and effects of placement. *Journal of Intellectual Disability Research*, 50(11), 845–856.
- Beadle-Brown, J., et al. (2005). *The Staff Experiences and Satisfaction Questionnaire for use in Learning Disability Services*. Canterbury: Tizard Centre.
- Beadle-Brown (unpublished). Cited in McGill, P., Richardson, L., and Langthorne, P. (2010).
- Berryman, J., Evans, I. M. and Kalbag, A. (1994). The Effects of Training in Nonaversive Behavior Management on the Attitudes and Understanding of Direct-Care Staff. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(3), 241-250.
- Berryman, J. (1991). *The development of an instrument designed to assess direct care staffs causal attributions of behavior problems*. Unpublished Master's dissertation, State University of New York at Binghamton, Binghamton, NY.
- Bradshaw, J., et al. (2004). Implementation and evaluation of active support. *Journal of Applied Research in Intellectual Disabilities* 17 13-148.
- British Psychological Society and Royal College of Psychiatrists (2015). *Challenging Behaviour and Learning Disabilities: Prevention and interventions for people with learning disabilities whose behaviour challenges*. NICE guideline 11. [Online]. Available from: <http://www.nice.org.uk/guidance/ng11/evidence/full-guideline-60708493> [Accessed 6th September 2015].
- Broadhurst, S., and Mansell, J. (2007). Organisational and individual factors associated with breakdown of residential placements for people with intellectual disabilities. *Journal of Intellectual Disability Research* 51(4), 293-301.
- Bromley, J. and Emerson, E. (1995), Beliefs and emotional reactions of care staff working with people with challenging behaviour. *Journal of Intellectual Disability Research*, 39, 341–352.

- Brown R. I., Schalock R. L. and Brown I. (2009) Quality of life: its application to persons with intellectual disabilities and their families-introduction and overview. *Journal of Policy and Practice in Intellectual Disabilities* 6, 2–7.
- Browning-Wright, D., et al. (2007). A preliminary study on the effects of training using Behavior Support Plan Quality Evaluation Guide (BSP-QE) to improve positive behavioral support plans. *Education and Treatment of Children*, 30(3) 89-106.
- Browning-Wright, D., Mayer, G.R., Saren, D. (2003). *Behaviour Support Plan Quality Evaluation Scoring Guide II*. [Online]. Available from: <http://www.pent.ca.gov/beh/qe/bipscoreingrubric.pdf> [Accessed 17August 2015].
- Brulle, A. R. and Repp, A. C. (1984). An investigation of the accuracy of momentary time sampling procedures with time series data. *British Journal of Psychology* 75, 481-485.
- Buschbacher, P. W. and Fox, L. (2003). Understanding and intervening with the challenging behavior of young children with autism spectrum disorder. *Language Speech and Hearing Services in Schools*, 34(3), 217-227.
- Carpenter J., et al. (2000). From the asylum to the community: a longitudinal study of staff involved in the transition from Tone Vale Hospital to community-based services. *Journal of Mental Health* 9, 211–30.
- Carr, E. G., et al. (2002). Positive Behavior support: Evolution of an applied science. *Journal of Positive Behavior Interventions*, 4(1), 4-20.
- Carr, E. G. (2007). The expanding vision of positive behavior support: Research perspectives on happiness, helpfulness, hopefulness. *Journal of Positive Behavior Interventions*, 9(1), 3-14.
- Carr, E.G. (1988). Functional equivalence as a mechanism of response generalization. In Horner, R.H., Dunlap, G., Koegel, R.L. eds. *Generalization and Maintenance: Lifestyle changes in applied settings*. Baltimore: Paul Brookes, p.221-241.
- Carr, E.G., et al. (1999). *Positive Behavioral Support for People with Developmental Disabilities: A Research Synthesis*. Washington: American Association on Mental Retardation.
- Carr, J. E. and Sidener, T. M. (2002). On the relation between applied behavior analysis and positive behavioral support. *Behavior Analyst*, 25(2), 245-253.
- Cherniss, C. (1980). *Staff Burnout: Job Stress in the Human Services*. Beverly Hills, CA: Sage.
- Cicchetti, D.V. and Sparrow, S. (1981). Developing criteria for establishing interrater reliability of specific items: Applications to assessment of adaptive behavior. *American Journal of Mental Retardation*. 86, 127–137.
- Clark-Carter, D. (2010). *Quantitative Psychological Research*. New York: Psychology Press.

- Clarke, S. and Dunlap, G. (2008). A descriptive analysis of intervention research published in the Journal of Positive Behavior Interventions: 1999 through 2005. *Journal of Positive Behavior Interventions*, 10(1), 67-71.
- Clarke, S., et al. (2002). Using Multiple Measures to Evaluate Positive Behavior Support a Case Example. *Journal of Positive Behavior Interventions*, 4(3), 131-145.
- Clement, T. and Bigby, C. (2012). Competencies of front-line managers in supported accommodation: Issues for practice and future research. *Journal of Intellectual and Developmental Disability*, 37(2), 131-140.
- Clement, T. and Bigby, C. (2008). *Making life good in the community: As good as it gets? An overview of methodology and a synthesis of findings and recommendations*. La Trobe University: Victoria Department of Human Services.
- Clement, T. and Bigby, C. (2007). *Making life good in the community: The importance of practice leadership and the role of the house supervisor*. La Trobe University: Victoria Department of Human Services.
- Crates, N. and Spicer, M. (2012). Developing behavioural training services to meet defined standards within an Australian statewide disability service system and the associated client outcomes. *Journal of Intellectual and Developmental Disability*, 37(3), 196-208.
- Dagnan, D., Grant, F. and McDonnell, A. (2004). Understanding challenging behaviour in older people; the development of the Controllability Beliefs Scale. *Behavioural and Cognitive Psychotherapy*, 32(04), 501-506.
- Dagnan D., Trower P. and Smith R. (1998). Care staff responses to people with learning disabilities and challenging behaviour: a cognitive emotional analysis. *British Journal of Clinical Psychology*. 37, 59–68.
- Davison, S., et al. (2015). A national UK survey of peripatetic support teams for children and adults with intellectual and developmental disability who display challenging behaviour. *International Journal of Positive Behavioural Support* 5(1) 26-33.
- Daynes, S. and Baker, P. (2014). A pilot evaluation of positive behavioural support workshops for families of adults with intellectual disabilities who present challenging behaviour: ‘it should have been offered years ago’ *International Journal of Positive Behavioural Support* 4(2), 24-31.
- Dench, C. (2005). A model for training staff in positive behaviour support *Learning Disability Review* 10, 2.
- Denne, L., et al. (2013). Developing a core competencies framework for positive behavioural support: issues and recommendations. *International Journal of Positive Behavioural Support* 3(2), 24-31.

- Department of Health (2014). *Positive and Proactive Care: Reducing the need for restrictive interventions*. [Online]. London: HMSO. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/300293/JRA_DoH_Guidance_on_RP_web_accessible.pdf [Accessed 17 August 2015].
- Department of Health (2013). *Winterbourne View Review. Transforming Care One Year On*. [Online]. London: HMSO. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/265752/Winterbourne_View_One_Year_On_Report.pdf [Accessed 17 August 2015].
- Department of Health (2012). *Transforming Care: a national response to Winterbourne View Hospital*. [Online]. London: HMSO. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/213215/final-report.pdf [Accessed 17 August 2015].
- Department of Health (2007). *Services for people with learning disabilities and challenging behaviour or mental health needs*. [Online]. London: Department of Health. Available from: https://www.kent.ac.uk/tizard/research/research_projects/dh2007mansellreport.pdf [Accessed 17 August 2015].
- Department of Health (2001). *Valuing People: A new strategy for learning disability for the 21st Century*. [Online]. London: HMSO. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/250877/5086.pdf [Accessed 17 August 2015].
- Deveau, R. and McGill, P. (2013). Leadership at the front line: Impact of practice leadership management style on staff experience in services for people with intellectual disability and challenging behaviour. *Journal of Intellectual and Developmental Disability*, 39(1), 65-72.
- Deveau, R. and McGill, P. (2009). Physical Interventions for Adults with Intellectual Disabilities: Survey of Use, Policy, Training and Monitoring. *Journal of Applied Research in Intellectual Disabilities*, 22(2), 145-151.
- Didden, R., et al. (2006). Behavioral treatment of challenging behaviors in individuals with mild mental retardation: Meta-analysis of single-subject research. *American Journal on Mental Retardation*, 111(4), 290-298.
- Didden, R., Duker, P. C., and Korzilius, H. (1997). Meta-analytic study of treatment effectiveness for problem behaviours with individuals who have mental retardation. *American Journal on Mental Retardation*, 101(4), 387-399.
- Donnellan, A.M., et al. (1985). A Time-Limited Intensive Intervention Program Model to Support Community Placement for Persons with Severe Behavior Problems *JASH*, 10(3), 123-131.

- Dowey, A., et al. (2007). Can brief workshop interventions change care staff understanding of challenging behaviours? *Journal of Applied Research in Intellectual Disabilities*, 20(1), 52-57.
- Dunlap, G., et al. (2014). A Comment on the Term 'Positive Behavior Support'. *Journal of Positive Behavior Interventions*, 16(3), 133-136.
- Dunlap, G., et al. (2010). A descriptive, multiyear examination of positive behavior support. *Behavioral Disorders*, 259-279.
- Dunlap, G., et al. (2009). Overview and history of positive behavior support. In: Sailor, W., Dunlap, G., Sugai, G., and Horner, R. eds. *Handbook of Positive Behavior Support*. Springer Science and Business Media, p.3-16.
- Dunlap, G., et al. (2008). Positive behavior support and applied behavior analysis: a familial alliance. *Behavior Modification*, 32(5), 682-698.
- Dunlap, G. and Carr, E.G. (2007). Positive behavior support and developmental disabilities: A summary and analysis of research. In Odom, S.L., Horner, R.H., Snell, M.E. and Blacher, J. eds. *Handbook of Developmental Disabilities*. London: Guilford Press p.469-482.
- Dunlap, G. and Horner, R. H. (2006). The Applied Behavior Analytic Heritage of PBS: A Dynamic Model of Action-Oriented Research. *Journal of Positive Behavior Interventions*, 8(1), 58-60.
- Dunlap, G., et al. (2000). Essential elements of in-service training in positive behavior support. *Journal of Positive Behavior Interventions*, 2(1), 22-32.
- Dunlap, G., Clarke, S. and Steiner, M. (1999). Intervention Research in Behavioral and Developmental Disabilities 1980 to 1997. *Journal of Positive Behavior Interventions*, 1(3), 170-180.
- Durand, M., et al. (2013). Positive family intervention for severe challenging behavior 1: a multi-site randomized clinical trial. *Journal of Positive Behavior Interventions* 15(3), 133-143.
- Emerson, E. (1996). Community support teams for people with learning disabilities and challenging behaviours: Results of a national survey. *Journal of Mental Health*, 5(4), 395-406.
- Emerson, E. and Einfield, S. (2011). *Challenging Behaviour: Analysis and Intervention in People with Severe Learning Disabilities*. Cambridge: Cambridge University Press.
- Emerson, E. and Hatton, C. (1996). Deinstitutionalization in the UK and Ireland: Outcomes for service users. *Journal of Intellectual and Developmental Disability*, 21(1), 17-37.
- Emerson E and McGill P (1989). Normalization and applied behaviour analysis: values and technology in services for people with learning difficulties. *Behavioural Psychotherapy* 17, 101-17.
- Emerson, E., et al. (2013). *People with Learning Disabilities in England 2012*. Cambridge: The Learning Disabilities Observatory.

- Emerson, E., et al. (2001). The prevalence of challenging behaviors: a total population study. *Research in Developmental Disabilities*, 22(1), 77-93.
- Emerson, E., et al. (2000a). Treatment and management of challenging behaviour in residential settings. *Journal of Applied Research in Intellectual Disabilities* 13, 197-215.
- Emerson, E., et al. (2000b). The quality and costs of community-based residential supports and residential campuses for people with severe and complex disabilities. *Journal of Intellectual and Developmental Disabilities* 25, 263-279.
- Emerson, Eric, et al. (2001). The prevalence of challenging behaviors: A total population study. *Research in Developmental Disabilities* 22(1), 77-93.
- Emerson, E., et al. (1987). Challenging behaviour and community services: 1 Introduction and overview. *Mental Handicap*, 15(4), 166-169.
- Felce, D. and Emerson, E. (2004). Research on engagement in activity. In Emerson, E. Hatton, C. Thompson, T. Parmenter, T. eds. *The International Handbook of Applied Research in Intellectual Disabilities*. England: John Wiley and Sons, Ltd, 353-368.
- Felce D., et al. (2003). Rational resourcing and productivity: relationships among staff Input, resident characteristics, and group home quality. *American Journal on Mental Retardation* 108(3), 161-172.
- Felce, D. Lowe, K. and Jones, E. (2002). Staff activity in supported housing services. *Journal of Applied Research in Intellectual Disabilities* 15, 388-403.
- Felce, D., et al. (2000). Exploring the relationship between costs and quality of services for adults with severe intellectual disabilities and the most severe challenging behaviours in Wales: a multivariate regression analysis. *Journal of Intellectual and Developmental Disability*, 25, 307-326.
- Felce, D. et al. (1998). Service support to people in Wales with severe intellectual disabilities and the most severe challenging behaviour: processes, outcomes and costs. *Journal of Intellectual Disability Research* 42, 390-408.
- Felce, D. and Perry, J. (1995). The extent of support for ordinary living provided in staffed housing: the relationship between staffing levels, resident characteristics, staff: resident interactions and resident activity patterns. *Social Science and Medicine*, 40(6), 799-810.
- Felce, D., Lowe, K. and Beswick, J. (1993). Staff turnover in ordinary housing services for people with severe or profound mental handicaps. *Journal of Intellectual Disability Research*, 37, 143-152.
- Felce, D., et al. (1991). The relationship of staff: client ratios, interactions, and residential placement. *Research in Developmental Disabilities*, 12(3), 315-331.

- Filter, K.J. (2007). In Response. Positive Behavior Support: Considerations for the future of a model. *The Behavior Analyst* 30(1), 87-89.
- Ford, J. and Honnor, J. (2000). Job satisfaction of community residential staff serving individuals with severe intellectual disabilities. *Journal of Intellectual and Developmental Disability*, 25(4), 343-362.
- Freeman, R., et al. (2005). Building a state-wide plan for embedding positive behaviour support in human services organisations. *Journal of Positive Behaviour Interventions* 7,109-119.
- Friedman, L.M. Furberg, C.D. and De Mets, D.L. (1998). *Fundamentals of Clinical Trials* Springer New York.
- Gardner, R. M., et al. (2003). Intensive positive behavior supports for adolescents with acquired brain injury - Long-term outcomes in community settings. *Journal of Head Trauma Rehabilitation*, 18(1), 52-74.
- George, D., and Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference*. 11.0 update (4th ed.). Allyn and Bacon Boston.
- Gifford, J., Beadle-Brown, J. and Mansell, J. (2006). A study of the work of managers in residential services for people with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities* 19 (3), 229.
- Goh, A. E. and Bambara, L. M. (2012). Individualized positive behavior support in school settings: A meta-analysis. *Remedial and Special Education*.
- Gore, N. and Umizawa, H. (2011). Challenging behavior training for teaching staff and family carers of children with intellectual disabilities: A preliminary evaluation. *Journal of Policy and Practice in Intellectual Disabilities* [Online], 8(4), 266-275.
- Gore, N., et al. (2013). Definition and scope for positive behavioural support. *International Journal of Positive Behavioural Support* 3(2), 14-23.
- Gray, D., et al. (2013). Positive behavioural support as a clinical model within acute assessment and treatment services. *International Journal of Positive Behavioural Support*, 3(1), 40-46.
- Grey, I. M. and McClean, B. (2007). Service user outcomes of staff training in positive behaviour support using person-focused training: A control group study. *Journal of Applied Research in Intellectual Disabilities*, 20(1), 6-15.
- Grey I.M., McClean, B. and Barnes-Holmes, D. (2002). Staff attributions about the causes of challenging behaviours: effects of longitudinal training in multi-element behaviour support. *Journal of Learning Disabilities* 6, 297-312.
- Grey, I.M., et al. (2005). Evaluating the effectiveness of teacher training in applied behaviour analysis. *Journal of Intellectual Disabilities*, 9(3), 209-227.

- Gupta, S.K. (2011). Intention to treat concept: a review. *Perspectives in Clinical Research* 2(3) 109-112.
- Haring, N. and De Vault, G. (1996). Discussion: Family issues and family support. *Positive Behavioral Support: Including People with Difficult Behavior in the Community*, 115-120.
- Harris, P., Humphreys, J., and Thomson, G. (1994). A checklist of challenging behaviour: the development of a survey instrument. *Mental Handicap Research*, 7(2), 118-133.
- Harrop, A., Murphy, M. and Shelton, T. (1994). Momentary time sampling and partial interval recording: a rejoinder to Adams (1991). *Behavioural and Cognitive Psychotherapy*, 22(01), 97-98.
- Hassiotis, A., and I. Hall (2008). 'Behavioural and cognitive-behavioural interventions for outwardly-directed aggressive behaviour in people with learning disabilities.' *Cochrane Library* 2008, 3.
- Hassiotis, A., et al. (2012). Impact of applied behaviour analysis on carer burden and community participation in challenging behaviour: results from a randomised controlled trial. *Journal of Intellectual Disability Research* 56(3) 285-290.
- Hassiotis, A., et al. (2011). Applied behaviour analysis and standard treatment in intellectual disability: 2-year outcomes. *The British Journal of Psychiatry: The Journal of Mental Science*, 198(6), 490-491.
- Hassiotis, A., et al. (2009). Randomized, Single-Blind, Controlled Trial of a Specialist Behavior Therapy Team for Challenging Behavior in Adults With Intellectual Disabilities. *American Journal of Psychiatry*, 166(11), 1278-1285.
- Hassiotis, A., et al. (2008). Individual characteristics and service expenditure on challenging behaviour for adults with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities*, 21(5), 438-445.
- Hastings, R.P. (1997). Measuring staff perceptions of challenging behaviour: the Challenging Behaviour Attributions Scale (CHABA). *Journal of Intellectual Disability Research* 41, 495-501.
- Hastings, R. P. (1996). Staff strategies and explanations for intervening with challenging behaviours. *Journal of Intellectual Disability Research*, 40(2), 166-175.
- Hastings, R.P. (1995). Understanding factors that influence staff responses to challenging behaviours: an exploratory interview study. *Mental Handicap Research* 8(4), 296-320.
- Hastings, R.P. and Brown T. (2002). Behaviour problems of children with autism, parental self-efficacy and mental health. *American Journal on Mental Retardation* 107, 222-232.

- Hastings, R. P., and Remington, B. (1994a). Staff behaviour and its implications for people with learning disabilities and challenging behaviours. *British Journal of Clinical Psychology*, 33, 423–438.
- Hastings R. and Remington B. (1994b) Rules of engagement: towards an analysis of staff responses to challenging behaviour. *Research in Developmental Disability*. 15, 279–98
- Hasting, R.P., et al. (2013). A conceptual framework for understanding why challenging behaviours occur in people with developmental disabilities. *International Journal of Positive Behavioural Support*, 3(2), 5-13.
- Hastings R., Reed T. and Watts M. (1997). Community staff causal attributions about challenging behaviour in people with intellectual disabilities. *Journal of Applied Research in Intellectual Disability*. 10, 238–49.
- Hastings R., Remington B. and Hopper G. (1995). Experienced and inexperienced health care workers' beliefs about challenging behaviours. *Journal of Intellect Disability Research*. 39, 474–83.
- Hatton, C., et al. (2001). Factors associated with intended staff turnover and job search behaviour in services for people with intellectual disability. *Journal of Intellectual Disability Research*, 45(3), 258-270.
- Hatton, C. and Emerson, E. (1996). *Residential Provision for People with Learning Disabilities: A Research Review*. Manchester: Hester Adrian Research Centre, Manchester University.
- Hewitt, A. S., et al. (2004). Role and essential competencies of the frontline supervisors of direct support professionals in community services. *Journal Information*, 42(2).
- Hieneman, M. and Dunlap, G. (2000). Factors affecting the outcomes of community-based behavioral support: I. Identification of factor categories. *Journal of Positive Behavior Interventions*, 2(3), 161-169.
- Hieneman, M. and Dunlap, G. (2001). Factors affecting the outcomes of community-based behavioral support: II. Factor category importance. *Journal of Positive Behavior Interventions*, 3(2), 67.
- Hollis, S. and Campbell, F. (1999). What is meant by intention to treat analysis? Survey of published randomised controlled trials. *British Medical Journal* 319 p670-674.
- Holmes, L. (1994). *The relationship between community participation, self-esteem, perceptions of stigma and social comparison processes in adults with mild learning disabilities*. Unpublished Master's dissertation, Canterbury: University of Kent.
- Horner, R. H. (2000). Positive Behavior Supports. *Focus on Autism and Other Developmental Disabilities* 15(2), 97-105.

- Horner, R. H., et al. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, 71(2), 165-179.
- Horner, R. H., et al. (2000). Elements of Behavior Support Plans: A Technical Brief. *Exceptionality*, 8(3), 205-215.
- Horner, R.H., et al. (1990). Toward a technology of ‘non-aversive’ behavioral support. *Journal of the Association for Persons with Severe Handicaps*, 15, 125-132.
- House of Commons Committee of Public Accounts (2015). *Care services for people with learning disabilities and challenging behaviour*. [Online]. London: HMSO. Available from: <http://www.publications.parliament.uk/pa/cm201415/cmselect/cmpubacc/973/973.pdf> [Accessed 17 Aug. 15].
- Jacobson, N.S. and Truax, P. (1991). Clinical Significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology* 59(1) p12-19.
- Johnston, J.M., et al. (2006). Positive behavior support and applied behavior analysis. *The Behavior Analyst*, 29, 51-74.
- Jones, C. and Hastings, R. P. (2003). Staff reactions to self-injurious behaviours in learning disability services: Attributions, emotional responses and helping. *British Journal of Clinical Psychology*, 42, 189-203.
- Jones, E., et al. (2013). Active support as a primary prevention strategy for challenging behaviour. *International Journal of Positive Behavioural Support*, 3(1), 16-30.
- Jones, E., et al. (2001a). Evaluation of the dissemination of active support training in staffed community residences. *American Journal on Mental Retardation*, 106(4), 344-358.
- Jones, E., et al. (2001b). Evaluation of the dissemination of active support training and training trainers. *Journal of Applied Research in Intellectual Disabilities* 14, 79-99.
- Jones, E., et al. (1999). Opportunity and the promotion of activity among adults with severe mental retardation living in community residences: The impact of training staff in Active Support. *Journal of Intellectual Disability Research*, 43, 164–178.
- Jones, H. (1998). PSR: Progressively sustainable results. *Positive Practices Newsletter*, 3(4), 3-9. Los Angeles: Institute for Applied Behavioral Analysis.
- Jung, D (2013). *Review of the use of the Guernsey Community Participation and Leisure Assessment*. Unpublished Master’s dissertation. Canterbury: Tizard Centre.
- Kalsy, S., et al. (2007). Effects of training on controllability attributions of behavioural excesses and deficits shown by adults with Down syndrome and dementia. *Journal of Applied Research in Intellectual Disabilities* [Online], 20(1), 64-68.

- Kazdin, A. E. (1981). Drawing valid inferences from case studies. *Journal of Consulting and Clinical Psychology*, 49, 183-192.
- Kincaid, D., George, H. P. and Childs, K. (2006). Review of the positive behavior support training curriculum: Supervisory and direct support editions. *Journal of Positive Behavior Interventions*, 8(3), 183-188.
- Kincaid, D. (1996). Person-centered planning. In Koegel, L.K., Koegel, R.L., and Dunlap, G. (Eds) (1996) *Positive Behavioral Support: Including people with difficult behavior in the community*. Baltimore: Paul H. Brookes.
- Kincaid, D., et al. (2002). Measuring the impact of positive behavior support. *Journal of Positive Behavior Interventions*, 4, 109-117.
- Kinnear, P.R. and Gray, C.D. (2010). *PASW17 Statistics Made Simple*. New York: Psychology Press.
- Knapp, M., et al. (2005). Intellectual disability, challenging behaviour and cost in care accommodation: What are the links? *Health and Social Care in the Community*, 13(4), 297–306.
- Knoster, T. (1999a). *Behavioral outcomes survey*. Unpublished manuscript, Tri-State Consortium on Positive Behavior Support, Central Susquehanna Intermediate Unit, Lewisburg, PA.
- Knoster, T. (1999b). *Positive behavioral support satisfaction survey*. Unpublished manuscript, Tri-State Consortium on Positive Behavior Support, Central Susquehanna Intermediate Unit, Lewisburg, PA.
- de Kock, U., et al. (1987). Staff turnover in a small home service: a study of facilities for adults with severe and profound mental handicaps. *Journal of the British Institute of Mental Handicap*, 15(3), 97-101.
- Koegel, L.K., Koegel, R.L., and Dunlap, G. eds. (1996). *Positive Behavioral Support: Including people with difficult behavior in the community*. Baltimore: Paul Brookes.
- Kraemer, B.R., et al. (2008). Effects of training on the use of the Behavior Support Plan Quality Evaluation Guide with autism educators: a preliminary investigation examining positive behavior support plans. *Journal of Positive Behavior Interventions*, 10(3), 179-189.
- Larson, S. A. and Hewitt, A. S. (2005). *Staff Recruitment, Retention, and Training Strategies for Community Human Services Organizations*. Baltimore: Paul Brookes.
- La Vigna, G.W. and Donnellan A.M. (1989). *Behavior Assessment and Intervention Guide*. Los Angeles: Institute for Applied Behavioral Analysis.
- La Vigna, G. W., and Donnellan, A. M. (1986). *Alternatives to punishment: Solving behavior problems with non-aversive strategies*. New York: Irvington.

- La Vigna, G. W. and Willis, T. J. (2012). The efficacy of positive behavioural support with the most challenging behaviour: The evidence and its implications. *Journal of Intellectual and Developmental Disability*, 37(3), 185-195.
- La Vigna, G.W. and Willis, T.J. (2009). *Positive practices in behavioral support: Facilitator's Manual*. Los Angeles: Institute for Applied Behavior Analysis.
- La Vigna, G.W. and Willis, T.J. (2005a). A positive behavioral support model for breaking the barriers to social and community inclusion. *Learning Disability Review*, 10(2), 16-23.
- La Vigna, G.W. and Willis, T.J. (2005b). Episodic severity: an overlooked dependent variable in the application of behavior analysis to challenging behavior. *Journal of Positive Behavior Interventions*, 7, 47-54.
- La Vigna, G.W. and Willis, T.J. (1996). Behavioral Technology in Support of Values. *Positive Practices Newsletter*, 1(4), 1 and 7-16. Los Angeles: Institute for Applied Behavioral Analysis.
- La Vigna, G.W. and Willis, T.J. (1995). A positive behavioral support model for breaking the barriers to social and community inclusion *Positive Practices Newsletter*, 1(1), 1 and 8-14. Los Angeles: Institute for Applied Behavioral Analysis.
- La Vigna, G. W. and Willis, T. J. (1992). A model for multi-element treatment planning and outcome measurement. In Berkell, D.E. ed. *Autism: Identification, education, and treatment*. Hillsdale, NJ: Erlbaum, 135-149.
- La Vigna, G. W., Christian, L. and Willis, T. J. (2005). Developing behavioural services to meet defined standards within a national system of specialist education services. *Pediatric Rehabilitation*, 8(2), 144-155.
- La Vigna, G.W., et al. (2002). Training professionals in the use of positive methods for community integration of persons with developmental disabilities *Psychiatric Services* 53, 16-18.
- La Vigna, G.W., et al. (1994). *The Periodic Service Review: A total quality assurance system for human services and education*. Baltimore: Paul Brookes.
- La Vigna, G.W., et al. (1989). The role of positive programming in behavioral treatment. In Cipani, E. ed. *The Treatment of Severe Behavior Disorders*. Monographs of the American Association on Mental Retardation 12, 59-83.
- Leadbetter, D. (2002). Good practice in physical intervention. In: Allen, D. ed. *Ethical Approaches to Physical Interventions. Responding to challenging behaviour in people with intellectual disabilities*. Kidderminster: BILD.
- Lehotkay, R., et al. (2015). Factor validity and reliability of the Aberrant Behavior Checklist-Community (ABC-C) in an Indian population with intellectual disability. *Journal of Intellectual Disability Research* 59(3), 208-214.
- Leyin, A. (2008). Social inclusion: ten years on. *Tizard Learning Disability Review*, 13(1), 10-17.

- Little, R and Yau, L (1996). Intent-to-treat analysis for longitudinal studies with dropouts. *Biometrics* 52(4) 1324-1333.
- Local Government Association (2014). *Ensuring Quality Services*. [Online]. London: Local Government Association. Available from: <http://www.local.gov.uk/documents/10180/12137/L14-105+Ensuring+quality+services/085fff56-ef5c-4883-b1a1-d6810caa925f> [Accessed 17 Aug. 15].
- Local Government Association, Association of Directors of Adult Social Services, and NHS England (2015). *Supporting people with a learning disability and/or autism who have a mental health condition or display behaviour that challenges*. [Online]. Available from: <http://www.england.nhs.uk/wp-content/uploads/2015/07/ld-draft-serv-mod.pdf> [Accessed 17 August 2015].
- Lowe, K., et al. (2010). The evaluation of periodic service review (PSR) as a practice leadership tool in services for people with intellectual disabilities and challenging behaviour. *Tizard Learning Disability Rev*, 15(3), 17-28.
- Lowe, K., et al. (2007a). Challenging behaviours: Prevalence and topographies. *Journal of Intellectual Disability Research*, 51(8), 625-636.
- Lowe, K., et al. (2007b). Staff training in positive behaviour support: impact on attitudes and knowledge *Journal of Applied Research in Intellectual Disabilities* 20, 30-40.
- Lowe, K., Felce, D., and Blackman, D. (1996). Challenging behaviour: The effectiveness of specialist support teams. *Journal of Intellectual Disability Research*, 40, 336-347.
- Lowe, K. Felce, D. Blackman, D. (1995). People with learning disabilities and challenging behaviour: the characteristics of those referred and not referred to specialist teams *Psychological Medicine* 25 p595-603.
- Lucyshyn, J.M., et al. (2002). Positive behavior support with families. In Lucyshyn, J.M., Dunlap, G. and Albin, R.W. eds. *Families and Positive Behavior Support: Addressing Problem Behavior in Family Contexts*. Baltimore: Paul Brookes, 3-43.
- MacDonald, A., Shaw, C. and Thomson, M. (2013). Evaluating 'Positive Pathways', a training programme to address challenging behaviour for carers of adults with autism. *Good Autism Practice* 14(1), 120-130.
- MacDonald, A. and McGill, P. (2013). Outcomes of Staff Training in Positive Behaviour Support: A Systematic Review. *Journal of Developmental and Physical Disabilities*, 25(1), 17-33.
- MacDonald, A., McGill, P., and Deveau, R. (2011). 'You squeal and squeal but they just hold you down'. *International Journal of Positive Behavioural Support*, 1(1), 45-52.

- MacDonald, A., Hume, L. and McGill, P. (2010). The use of multi-element behaviour support planning with a man with severe learning disabilities and challenging behaviour. *British Journal of Learning Disabilities*, 38(4), 280-285.
- Macurik, K. M., et al. (2008). Video training of support staff in intervention plans for challenging behavior: Comparison with live training. *Behavioral Interventions*, 23(3), 143-163.
- Mansell, J. (1995). Staffing and staff performance in services for people with severe or profound learning disability and serious challenging behaviour. *Journal of Intellectual Disability Research*, 39(1), 3-14.
- Mansell J. (1985). Time sampling and measurement error: the effect of interval length and sampling pattern. *Journal of Behavior Therapy and Experimental Psychiatry* 16, 245–251.
- Mansell, J. and Beadle-Brown, J. (2012). *Active Support: Enabling and Empowering People with Intellectual Disabilities*. London: Jessica Kingsley.
- Mansell, J. and Beadle-Brown, J. (2011). Estimating activity duration by momentary time-sampling of part or all of the day *Journal of Applied Research in Intellectual Disabilities* 1-6.
- Mansell, J. and Ericsson, K. (1996). *Deinstitutionalization and community living: Intellectual disability services in Britain, Scandinavia and the USA*. London: Springer.
- Mansell, J., Beadle-Brown, J. and Bigby C. (2013). Implementation of active support in Victoria, Australia: an exploratory study. *Journal of Intellectual and Developmental Disability* 38(1), 48-58.
- Mansell J., et al. (2008). Effect of service structure and organisation on staff care practices in small community homes for people with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities* 21, 389-413.
- Mansell, J., Elliott, T. E., and Beadle-Brown, J. (2005a). *Active Support Measure (Revised)*. Canterbury, UK: Tizard Centre.
- Mansell J., et al. (2005b). *Person-Centred Active Support: A Multi-Media Training Resource for Staff to Enable Participation, Inclusion and Choice for People with Learning Disabilities*. Brighton: Pavilion.
- Mansell J., et al. (2003). Resident involvement in activity in small community homes for people with learning disabilities. *Journal of Applied Research in Intellectual Disabilities* 16, 63–74.
- Mansell, J., et al. (2002). Engagement in meaningful activity and ‘active support’ of people with intellectual disabilities in residential care. *Research in Developmental Disabilities*, 23(5), 342-352.

- Mansell, J., McGill, P. and Emerson, E. (2001). Development and evaluation of innovative residential services for people with severe intellectual disability and serious challenging behavior. *International Review of Research in Mental Retardation*, 24, 245-298.
- Mansell, J., Hughes, H., and McGill, P. (1994). Maintaining local, residential placements. In Emerson, E., McGill, P. and Mansell, J. (Eds.) *Severe learning disabilities and challenging behaviours: Designing high-quality services* (p.261–281). London: Chapman and Hall.
- Marquis, J. G., et al. (2000). A meta-analysis of positive behavior support. *Contemporary Special Education Research: Syntheses of the Knowledge Base on Critical Instructional Issues*, 137-178.
- Matson, J. L. and Gorman-Smith, D. (1986). A review of treatment research for aggressive and disruptive behavior in the mentally retarded. *Applied Research in Mental Retardation*, 7(1), 95-103.
- Matson, J.L. and Vollmer, T.R. (1995). *User's Guide: Questions About Behavioral Function (QABF)* Scientific Publishers, Baton Rouge, LA (1995).
- McAuley, E., Duncan, T.E. and Russell, D.W. (1992). Measuring causal attributions: The revised causal dimension scale (CDSII). *Personality and Social Psychology Bulletin*, 18, 566–573.
- McClellan, B. and Grey, I. (2012). A component analysis of positive behavior support plans. *Journal of Intellectual and Developmental Disability* 37(3) 221-231.
- McClellan, B., et al. (2005). Person Focused Training: a model for delivering positive behavioural supports to people with challenging behaviours. *Journal of Intellectual Disability Research*, 49, 340-352.
- McClellan, B., Grey, I. M. and McCracken, M. (2007). An evaluation of positive behavioural support for people with very severe challenging behaviours in community-based settings. *Journal of Intellectual Disabilities*, 11(3), 281-301.
- McClintock, K., Hall, S. and Oliver, C. (2003). Risk markers associated with challenging behaviours in people with intellectual disabilities: a meta-analytic study. *Journal of Intellectual Disability Research*, 47(6), 405-416.
- McDonnell, A. (2008a). Training care staff to manage challenging behaviour: an evaluation of a three-day training course. *British Journal of Developmental Disabilities*, 43(2), 156-162.
- McDonnell, A., et al. (2008b). The effects of staff training on staff confidence and challenging behavior in services for people with autism spectrum disorders. *Research in Autism Spectrum Disorders*, 2(2), 311-319.
- McGill, P. and Poynter, J. (2012). High Cost Residential Placements for Adults with Intellectual Disabilities. *Journal of Applied Research in Intellectual Disabilities*, 25(6), 584-587.

- McGill, P., et al. (in press). Preventing challenging behaviour of adults with complex needs in supported accommodation.
- McGill, P., Richardson, L., and Langthorne, P. (2010). *A comparative analysis of the quality of provision in the Hesley Group's adult residential care homes*. Canterbury: Tizard Centre.
- McGill, P., Bradshaw, J. and Hughes, A. (2007). Impact of extended education/training in positive behaviour support on staff knowledge, causal attributions and emotional responses. *Journal of Applied Research in Intellectual Disabilities*, 20(1), 41-51.
- McGillivray, J. and McCabe, M. (2005). The Relationship between Residence and the Pharmacological Management of Challenging Behavior in Individuals with Intellectual Disability. *Journal of Developmental and Physical Disabilities*, 17(4), 311-325.
- McKenzie, K., et al. (2006). Difficulties faced by social care staff when managing challenging behaviour: *Learning Disability Practice*, 9(2), 28-32.
- McKenzie K., MacLean H., Megson R. and Reid, K. (2005). Behaviours that challenge. *Learning Disability Practice*. 8(9) 16-19.
- McKenzie, K., et al. (1999). Health and social care workers' understanding of the meaning and management of challenging behaviour in learning disability services. *Journal of Learning Disabilities for Nursing, Health and Social Care*, 3, 98–105.
- Mental Welfare Commission (2010). *Consent to treatment: a guide for mental health practitioners*. [Online]. Available from: <http://www.mwcscot.org.uk/media/51774/Consent%20to%20Treatment.pdf> [Accessed 17 August 2015].
- Mitchell, G. and Hastings, R. P. (1998). Learning disability care staffs emotional reactions to aggressive challenging behaviours: Development of a measurement tool. *British Journal of Clinical Psychology*, 37, 441-449.
- Mores, B and Grant, GW (1976). Nurse expectations for accomplishment of mentally retarded patients. *American Journal of Mental Deficiency*, 80, 644–649.
- Munton, A. G., Silvester, J. and Stratton, P. (1999). *Attributions in Action: A Practical Approach to Coding Qualitative Data*. John Wiley and Sons Inc.
- Murphy, G. (1987). Direct observation as an assessment tool in functional analysis and treatment. *Assessment in Mental Handicap: A Guide to Assessment Practice, Tests and Checklists*, 190-238.
- NHS England (2014). *Winterbourne View – Time for Change: Transforming the commissioning of services for people with learning disabilities and/or autism*. [Online]. Available from: <http://www.england.nhs.uk/wp-content/uploads/2014/11/transforming-commissioning-services.pdf> [Accessed 17 August 2015].

- Newman, I. and Emerson, E. (1991). Specialised treatment units for people with challenging behaviours. *Journal of the British Institute of Mental Handicap*, 19(3), 113-119.
- Nihira, K., Leland, H., and Lambert, N. (1993). AAMR adaptive behavior scale—residential and community (2nd edn). Austin, Texas: Pro-Ed.
- O'Brien, J. (1987). *A Guide to Lifestyle Planning*. Baltimore: Paul Brookes.
- O'Brien, J. and O'Brien, C.L. (eds) (2002). *Implementing Person-Centred Planning, Voices of Experience*. Toronto: Inclusion Press.
- O'Dell, S. M., et al. (2011). JPBI 10 Years Later: Trends in Research Studies. *Journal of Positive Behavior Interventions*, 13(2), 78-86.
- O'Neill, R.E., et al. (1997). *Functional Assessment and Program Development for Problem Behaviour: A Practical Handbook* (2nd edn). Pacific Grove, CA: Brooks/Cole.
- Ono, Y. (1996). Factor validity and reliability for the Aberrant Behavior Checklist-Community in a Japanese population with mental retardation. *Research in Developmental Disabilities*, 17(4), 303-309.
- Oliver, C., et al. (1996). Self-injurious behaviour and people with intellectual disabilities: Assessing the behavioural knowledge and causal explanations of care staff. *Journal of Applied Research in Intellectual Disabilities*, 9(3), 229-239.
- Perry, J., et al. (2011). Resettlement outcomes for people with severe challenging behaviour moving from institutional to community living. *Journal of Applied Research in Intellectual Disabilities* 24, 1-17.
- Pilling, N., McGill, P., and Cooper, V. (2007). Characteristics and experiences of children and young people with severe intellectual disabilities and challenging behaviour attending 52-week residential special schools. *Journal of Intellectual Disability Research*, 51(3), 184-19.
- Potter, G. (2014). The impact of multi-element positive behavioural support on severe self-injurious behaviour: a 21-year follow-up case report. *International Journal of Positive Behavioural Support*, 4(2), 3-8.
- Public Health England (2015). *Prescribing of psychotropic drugs to people with learning disabilities and/or autism by general practitioners in England*. [Online]. Available from: http://www.improvinghealthandlives.org.uk/securefiles/150818_1849//Psychotropic%20medication%20and%20people%20with%20learning%20disabilities%20or%20autism.pdf [Accessed 17 August 2015].
- Rae, H., Murray, G. and McKenzie, K. (2011). Teaching staff knowledge, attributions and confidence in relation to working with children with an intellectual disability and challenging behaviour. *British Journal of Learning Disabilities*, 39(4), 295-301.

- Reid, D., et al. (2003). Training human service supervisors in aspects of PBS: evaluation of a state-wide, performance-based program. *Journal of Positive Behavior Interventions*, 5(1), 35-46.
- Remington, B., et al. (2007). Early intensive behavioural intervention: outcomes for children with autism and their parents after two years. *American Journal on Mental Retardation* 112(6), 418-438.
- Reynolds, S., Lynch, S. L. and Litman, S. (2011). Training care teams of children with autism spectrum disorders in positive behaviour support: an innovative approach. *Healthcare Quarterly* 14(3), 95-9.
- Robertson, J. Emerson, E. Pinkney, L. Caesar, E. Felce, D. Meek, A. Carr, D. Lowe, K. Knapp, M. Hallam, A. (2004). Quality and costs of community-based residential supports for people with mental retardation and challenging behavior *American Journal on Mental Retardation* 109, 332-344.
- Rojahn, J., Polster, L.M., Mulick, J.A., Wisniewski, J.J. (1989). Reliability of the behaviour problems inventory. *Journal of the Multi-handicapped Person* 2(4), 283-293.
- Rose, D. and Rose, J. (2005). Staff in services for people with intellectual disabilities: the impact of stress on attributions of challenging behaviour. *Journal of Intellectual Disability Research* [Online], 49, 827-838.
- Rose, J., et al. (2014). Staff training using positive behavioural support: the effects of a one-day training on the attributions and attitudes of care staff who work with people with an intellectual disability and challenging behaviour. *International Journal of Developmental Disabilities*, 60(1), 35-42.
- Rotholz, D. A. and Ford, M. E. (2003). Statewide system change in positive behavior support. *Mental Retardation*, 41(5), 354-364.
- Rothwell, N. A., Lavigna, G. W. and Willis, T. J. (1999). A non-aversive rehabilitation approach for people with severe behavioural problems resulting from brain injury. *Brain Injury*, 13(7), 521-533.
- Roy, A, Matthews, H, Clifford, P., Fowler, V and Martin, DM (2002). 'Health of the Nation Outcome Scales for People with Learning Disabilities (HoNOS-LD)'. *British Journal of Psychiatry*, 80, 61-6.
- Royal College of Psychiatrists, British Psychological Society and Royal College of Speech and Language Therapists (2007). *Challenging Behaviour: a unified approach. Clinical and service guidelines for supporting people with learning disabilities who are at risk of receiving abusive or restrictive practices*. [Online]. Available from: <http://www.rcpsych.ac.uk/files/pdfversion/cr144.pdf> [Accessed 17 August 2015].

- Rusch, R.G., Hall, J.C. and Griffen, H.C. (1986). Abuse provoking characteristics of institutionalised mentally retarded individuals. *American Journal of Mental Deficiency* 90, 618-24.
- Sailor, W., et al. (2008). *Handbook of Positive Behavior Support*. Springer Science and Business Media.
- Sailor, W., et al. (2000). Using information technology to prepare personnel to implement functional behavioral assessment and positive behavioral support. *Exceptionality*, 8(3), 217-230.
- Schalock, R., et al. (2002). Conceptualization, measurement and application of quality of life for persons with intellectual disabilities: report of an international panel of experts. *Mental Retardation* 40, 457-470.
- Schalock, R.L. and Keith K.D. (1993). *Quality of Life Questionnaire: Standardisation Manual*. Hastings: Mid Nebraska Mental Retardation Services.
- Scotti, J. R., et al. (1991). A meta-analysis of intervention research with problem behavior: Treatment validity and standards of practice. *American Journal of Mental Retardation*, 96, 233–256.
- Scottish Consortium for Learning Disability (2015). *Learning Disability Statistics Scotland 2014*. [Online]. Available from: <http://www.sclcd.org.uk/wp-content/uploads/2015/08/Learning-Disability-Statistics-Scotland-2014-report.pdf> [Accessed 17 August 2015].
- Scottish Government (2013). *The Keys to Life: improving quality of life for people with learning disabilities*. Edinburgh. [Online]. Available from: <http://www.gov.scot/resource/0042/00424389.pdf> [Accessed 17 August 2015].
- Scottish Government (2010). *Adults with incapacity (Scotland) Act 2000. Code of practice*. [Online]. Available from: <http://www.gov.scot/Resource/Doc/327864/0105906.pdf> [Accessed 17 August 2015].
- Scottish Government (2008). *Communication and Assessing Capacity; a guide for social work and health care staff*. Edinburgh.
- Scottish Government (2000). *The Same as You? A review of services for people with learning disabilities*. [Online]. Available from: <http://www.gov.scot/resource/doc/1095/0001661.pdf> [Accessed 17 August 2015].
- Scottish Government (2000). *Adults with incapacity (Scotland) Act 2000*. [Online]. Available from: http://www.legislation.gov.uk/asp/2000/4/pdfs/asp_20000004_en.pdf [Accessed 17 August 2015].
- Sheehan, R. et al (2015). Mental illness, challenging behaviour, and psychotropic drug prescribing in people with intellectual disability: UK population based cohort study. *British Medical*

Journal, 351:h4326. [Online]. Available from: <http://www.bmj.com/content/351/bmj.h4326> [Accessed 7th September 2015].

- Singer, G.H.S. and Wang, M. (2009). The intellectual roots of positive behavior support and their implications for its development. In: Sailor, W., Dunlap, G., Sugai, G. and Horner, R. (eds) *Handbook of Positive Behavior Support*. New York: Springer Science and Business Media, 17-46.
- Singh, N.S. (Ed) (in press). *Clinical Handbook of Evidence-based Practices for Individuals with Intellectual and Developmental Disabilities*. Springer.
- Smith, M and Nethell, G (2010) *Brief Behavioural Assessment Tool (BBAT)*, Cardiff: Special Projects Team, Abertawe Bro Morgannwg University Health Board.
- Snell, M. E., Voorhees, M. D. and Chen, L. (2005). Team Involvement in Assessment-Based Interventions with Problem Behavior 1997—2002. *Journal of Positive Behavior Interventions*, 7(3), 140-152.
- Stancliffe, R.J. (1997). Community living unit size, staff presence, and residents' choice making. *Mental Retardation* 35, 1-9.
- Stancliffe, R.J., et al. (2007). Australian implementation and evaluation of active support. *Journal of Applied Research in Intellectual Disabilities* 20, 211-227.
- Stancliffe R.J. and Lakin K.C. (1998). Analysis of expenditures and outcomes of residential alternatives for persons with developmental disabilities. *American Journal of Mental Retardation*. 102, 552–568.
- Stokes, T. F. and Baer, D. M. (1977). Implicit Technology of Generalization. *Journal of Applied Behavior Analysis*, 10(2), 349-367.
- Sturmey, P. (2009). Restraint, seclusion and PRN Medication in English services for people with learning disabilities administered by the National Health Service: An analysis of the 2007 National Audit Survey. *Journal of Applied Research in Intellectual Disabilities*, 22, 140–144.
- Steiner, D. and Geddes, J. (2001). Intention to treat analysis in clinical trials when there are missing data. *Evidence Based Mental Health* 4, 70-71. [Online]. Available from: <http://ebmh.bmj.com> [Accessed 18 January 2014].
- Suen, H. K., and Ary, D. (1989). *Analyzing quantitative behavioral observation data*. Hillsdale, NJ: Erlbaum.
- Sugai, G., et al. (2000). Applying Positive Behavior Support and Functional Behavioral Assessment in Schools. *Journal of Positive Behavior Interventions*, 2(3), 131.

- Taylor L, Oliver C, Murphy G. (2011) .The chronicity of self-injurious behaviour: a long-term follow-up of a total population study. *Journal of Applied Research in Intellectual Disabilities*, 24(2), 105-117.
- Thabane, L., et al. (2013). *A tutorial on sensitivity analyse in clinical trials: the what, why, when and how*. [Online]. Available from: www.biomedcentral.com/1471-2288/13/92 [Accessed 10 October 2014].
- Thackrey, M. (1987). Clinician confidence in coping with patient aggression: Assessment and enhancement. *Professional Psychology: Research and Practice*, 18(1), 57-60.
- Thompson, T., et al. (1996). Interdependence of architectural features and program variables in community residences for people with mental retardation. *American Journal on Mental Retardation*, 101, 315-327.
- Tierney, E. Quinlan, D. and Hastings, R.P. (2007). Impact of a 3-day training course on challenging behaviour on staff cognitive and emotional responses. *Journal of Applied Research in Intellectual Disabilities* 20, 58-63.
- Toogood, S, et al. C (1994). Meeting the challenge in Clwyd 1. *British Journal of Learning Disabilities*, 22(1), 18–24.
- Tossebro J. (1995). Impact of size revisited: relation of number of residents to self-determination and deprivatization. *American Journal on Mental Retardation*, 100, 59-67.
- Totsika, V., et al. (2008). Persistence of challenging behaviours in adults with intellectual disability over a period of 11 years. *Journal of Intellectual Disability Research*, 52(5), 446–457.
- Totsika, V., et al. (2010). The effect of active support interactive training on the daily lives of adults with an intellectual disability. *Journal of Applied Research in Intellectual Disabilities*, 23(2), 112-121.
- Tyrer P., et al. (2008). Risperidone, haloperidol, and placebo in the treatment of aggressive challenging behaviour in patients with intellectual disability: a randomised controlled trial. *The Lancet*, 371 57–63.
- van Oorsouw, W. M., et al. (2009). Training staff serving clients with intellectual disabilities: A meta-analysis of aspects determining effectiveness. *Research in Developmental Disabilities*, 30(3), 503-511.
- van Oorsouw, W. M. W., et al. (2010). Training Staff to Manage Challenging Behaviour. *Journal of Applied Research in Intellectual Disabilities*, 23(2), 192-196.
- Wacker, D. P. and Berg, W. K. (2002). PBS as a service delivery system. *Journal of Positive Behavior Interventions*, 4(1), 25-28.

- Wang M., Schalock R. L., Verdugo M. A. and Jenaro C. (2010). Examining the factor structure and hierarchical nature of the quality of life construct. *American Journal on Intellectual and Developmental Disabilities* 115, 218–33.
- Wanless, L. and Jahoda, A. (2002). Responses of staff towards people with mild to moderate intellectual disability who behave aggressively: a cognitive emotional analysis. *Journal of Intellectual Disability Research*, 46(6), 507-516.
- Wardale, S., Davis, F., Carroll, M. and Vassos, M. (2014a). Outcomes for staff participating in positive behavioural support training. *International Journal of Positive Behavioural Support* 4(1), 10-23.
- Wardale, S., Davis, F. and Dalton, C. (2014b). Positive behavioural support training in a secure forensic setting: the impact on staff knowledge and positive behavioural support plan quality. *International Journal of Positive Behavioural Support* 4(2), 9-13.
- Watson, D., Everett, G. and Webb, H. (2008). *An independent evaluation of the kingfisher's project*. Bristol: Norah Fry Research Centre.
- Weiner, B. (1980). A cognitive (attributional) – emotion – action model of motivated behavior: An analysis of judgments of help-giving. *Journal of Personality and Social Psychology*, 39, 186–200.
- West, E.A. and Patton, H.A. (2010). Positive behaviour support and supported employment for adults with severe disability. *Journal of Intellectual and Developmental Disability*, 35(2), 104-111.
- Willis T. and La Vigna G.W. (1990) *Behavioral Assessment Report and Intervention Plan*. Los Angeles: Institute for Applied Behavior Analysis.
- Wills, S., Shephard, J. and Baker, P. (2013). Evaluating the impact of positive behaviour support training on staff knowledge, attributions, emotional responses and helping behaviour: capturing hearts and minds. *International Journal of Positive Behavioural Support* 3(1), 31-39
- Wacker, D.P. and Berg, W.K. (2002). PBS as a service delivery system. *Journal of Positive Behavior Interventions*, 4(1), 25-28.
- Williams, S., et al. (2012). Changes in Attributions as a Consequence of Training for Challenging and Complex Behaviour for Carers of People with Learning Disabilities: A Systematic Review. *Journal of Applied Research in Intellectual Disabilities*, 25(3), 203-216.
- Wishart, J., et al. (2013). Staff responses to behaviour that challenges: The role of knowledge, attributions, and emotion regulation style. *Journal of Intellectual and Developmental Disability*, 38(3), 205-214.
- Wolf, M. M. (1978). Social validity: the case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis*, 11(2), 203-214.

Wolfensberger, Wolf (1972). 'Normalization' *Toronto: National Institute on Mental Retardation*.

Zeilinger, E. L., Weber, G., and Haveman, M. J. (2011). Psychometric properties and norms of the German ABC-Community and PAS-ADD Checklist. *Research in Developmental Disabilities*, 32(6), 2431-2440.

13 Appendices

Appendix One – Details of Search for Systematic Review (Chapter Three)

Web of Knowledge

Search took place using the following terms:

- **Topic “Positive Behav* Support” + training = 12 studies** (Rose et al, 2014; McClean & Grey, 2012; Crates & Spicer, 2012; Reynolds et al, 2011; Kraemer et al, 2008; Grey & McClean, 2007; Lowe et al, 2007b; McGill et al, 2007; McClean et al, 2005; La Vigna et al, 2005; Freeman et al, 2005; Reid et al, 2003) plus Rotholz & Ford, 2003 (not included as data reported in Reid et al, 2003).
- **Topic “Positive Behav* Support” + learning disab* OR intellectual disab* = 7 studies** (McClean & Grey, 2012; Macurik et al, 2008; Grey & McClean, 2007; McGill et al, 2007; McClean et al, 2005; Freeman et al, 2005; Reid et al, 2003) plus Rotholz & Ford, 2003, as above.
- **Topic “challenging behav*” + training = 10 studies** (Crates & Spicer, 2012; McClean & Grey 2012; Gore & Umizawa, 2011; Grey & McClean, 2007; Lowe et al, 2007; McGill et al, 2007; Tierney et al 2007; McClean et al, 2005; Allen & Tynan, 2000; Allen et al, 1997) plus Grey et al, 2005 (see below, reference list search).

Psych Info

Search took place using the following terms:

- **All text “Positive Behav* Support” + training = 11 studies** (Crates & Spicer, 2012; McClean & Grey, 2012; Browning-Wright et al, 2007; Kraemer et al, 2005; Grey & McClean, 2007; Lowe et al, 2007; McGill et al, 2007; McClean et al, 2005; Freeman et al, 2005; Grey et al, 2005; Reid et al, 2003)
- **All text “Positive Behav* Support” + all text learning disab* OR intellectual disab* = no additional studies**
- **All text “challenging behav* + all text training + all text learning disab* OR intellectual disab* = no additional studies**

PubMed

Search took place using the following terms:

- **“Positive Behaviour Support”** = no studies
- **“challenging behaviour” AND training** = no studies

Google Scholar

Search took place using the following terms:

- **Exact phrase “Positive Behaviour Support” with the word training** = 8 studies (no additional studies)
- **Exact phrase “Positive Behaviour Support” with the words “learning disability”** = 8 studies (no additional studies)

Reference List Searches

Search took place for any additional studies referenced in the above articles.

6 studies additional studies were found: Berryman et al, 1994 & Grey et al, 2002 (both cited in McGill et al, 2007); Dench, 2005 (cited in Lowe et al, 2007); Baker, 1998 (cited in McClean et al, 2005); La Vigna et al, 2002 (cited in La Vigna et al, 2005); Dowey et al, 2007 (cited in Grey et al, 2007; this is an editorial in a special issue journal focused on training in challenging behaviour, therefore it was read in detail and its reference list was reviewed). The reference list of each of the 6 additional studies was then also reviewed. No further studies were found.

Citation Searches

Search took place in Web of Knowledge and Psych Info for any additional studies which cited the above articles. No additional studies were found

Search of International Journal of Positive Behavioural Support

3 additional studies (Wardale et al, 2014a; Wardale et al, 2014b; Wills et al, 2013)

Appendix Two – Sample Information Sheet (Chapter Five)

Information Sheet

Invitation to Participate

I am the manager of The Richmond Fellowship Scotland's Positive Behaviour Support Team, and our team is carrying out an evaluation which we would like to invite you to be part of.

Before you decide whether to take part, we would like you to understand why the evaluation is being done and what it would involve for you. One of our team will go through this information sheet with you and answer any questions you have. This should take about 15 minutes.

(Part 1 tells you the purpose of this evaluation and what will happen to you if you take part. Part 2 gives you more detailed information about the conduct of the evaluation).

Ask us if there is anything that is not clear. Talk to others about the evaluation if you wish.

Thank you,
Anne MacDonald

PART ONE

What is the purpose of the evaluation?

We are carrying out this work in order to evaluate the effect of training managers in 'Management of Positive Behaviour Support' (PBS). The training lasts a year and is person-focused, i.e. each manager attending the training selects a service user that they currently support who has learning disabilities and challenging behaviour. The training requires managers to put the learning from the course into practice in their service and to pass on the learning to their staff teams.

The study will look at the impact of the training on managers, on their staff teams, and on the service users they support. The study will consider the following areas for managers, staff teams and service users. Is PBS training linked with:

- increased knowledge regarding PBS
- better management practice
- changes in beliefs about challenging behaviour
- better quality of staff support to service users
- reductions in severity & frequency of behaviour for service users
- improved quality of life for service users

Why have I been invited?

You have been invited to take part in the study because you are undertaking the training. In order to ensure that any effects noted after the training are actually linked to the training and not to other organisational or service changes, a group of staff whose managers are not participating in the training are also needed to be part of this study. You are being asked to take part so your results can be compared with this comparison group whose manager is not doing the training.

Do I have to take part?

It is up to you to decide to join the study. We will describe the study and go through this information sheet. If you agree to take part, we will then ask you to sign a consent form. You are free to withdraw at any time, without giving a reason. This would not affect your job in any way.

What will happen to me if I take part?

If you take part in the evaluation, some information will be asked for from you about the service user you support. This will include:

- questionnaires for you to complete regarding PBS – to be done 3 times over 18 months
- questionnaires for you to complete about challenging behaviour – to be done 3 times over 18 months
- assessment rating of your manager – to be used 3 times in 3 years
- recording of incidents of challenging behaviour – for 4 week periods, on 3 occasions over 3 years
- videoing of your support to service users– for 2 hours, on 3 occasions over 3 years

The evaluation will take place over 3 years. Timing for the different evaluation measures is shown in the table below.

Dates	What is done	Who is involved
November 2011 December 2012 June 2013	4 weeks incident forms recordings	Staff & service users
November 2011 December 2012 June 2013	Questionnaire re PBS	Staff & manager
November 2011 December 2012 June 2013	Questionnaire re causes of challenge	Staff & manager
November 2011 December 2012	Questionnaire re management practice	Staff & manager
November 2011 December 2012 June 2013	Video of service users – 2hours	Staff & service users

Questionnaires

When questionnaires are filled out, one of the research team will send these to your service and then will arrange for them to be picked up again.

Video

When video is taken it will be by one of the research team. They will have specific guidelines to follow when they video – these are available for you to look at.

Confidentiality

All information taken from you or your service will be kept confidential and will not have your name on it. It will only have an identifying code that will be kept separately, so your information cannot be identified. Any information taken from you or your service will be stored securely and only the research team will have access to it. It will not be passed on to your managers.

What will I have to do?

You will be asked to complete the questionnaires and other measures, and to allow the research team into the service in order to speak to staff and to take video of service users. They will also want to collect records of incidents of challenging behaviour.

What are the risks of taking part?

You may feel uncomfortable having video taken in the service. This may feel difficult for service users also. You may find that having research team in the service talking to staff and collecting questionnaires is a bit disruptive to the service. However, any participant in the study is able to withdraw their participation at any time – so if you or service users wish to stop, then you should just tell the researcher.

What are the benefits of taking part?

The benefit of taking part is that you are helping the organisation to evaluate its training. We cannot promise that this evaluation will help you directly, but it may help service users with

challenging behaviour and it may also help the staff teams that support them.

What happens when the evaluation is over?

Once the evaluation is over the results will be written up and you will have a chance to read them. They will also be published on the organisation's website although you will not be identified.

What if there is a problem?

Any complaint about the way you have been dealt with during the study or any possible harm you might suffer will be addressed. The detailed information on this is given in Part 2.

Will my taking part be confidential?

The fact your service is participating in the study will not be confidential, but any information given by you will be and you will not be individually identified in the study.

PART TWO

What if I don't want to carry on with the evaluation?

If you want to withdraw from the study, you can do this at any time. We will not use any of your data if you do not wish us to.

What if there is a problem?

If you have a concern about any aspect of this study, you should ask to speak to Anne MacDonald who is the Chief Investigator in this evaluation study. She will do her best to answer your questions. She can be contacted on 0141 779 6300. If you remain unhappy and wish to complain formally, you can do this to her academic supervisor, Peter McGill. His contact details can be obtained from Anne.

What happens to the results of the study?

The results of the study will be written up and will be published on the organisation's website. They may also be published elsewhere or presented at academic conferences. However, you will not be identified in the study and all results will be presented generally not specifically about individual people or services.

Who is organising the study?

The evaluation study is being carried out as part of a PhD at the University of Kent and therefore they are sponsoring the research.

Who has reviewed the study?

All research is looked at by independent group of people, called a Research Ethics Committee, to protect the interests of any participants. This study has been reviewed and given favourable opinion.

CONSENT FORM

Title of Project: Evaluation of Positive Behaviour Support Training

Name of Researcher: Anne MacDonald

Please read the following statements and then initial the box next to each.

I confirm that I have read and understand the information sheet dated 4.5.11 for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	
I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my employment being affected.	
I understand that data collected during the study, may be looked at by individuals from the University of Kent where it is relevant to my taking part in this research. I give permission for these individuals to have access to this data.	
I agree to take part in the above study.	

Name of staff member:

Date:

Signature:

Name of person taking consent:

Date:

Signature:

When completed: 1 for participant; 1 for researcher file.

Appendix Three – Ethics Approval Letter (Chapter Five)

Scotland A Research Ethics Committee

Secretariat
2nd Floor Waverley Gate
2-4 Waterloo Place
Edinburgh
EH1 3EG
Telephone: 0131 465 5680
Fax: 0131 465 5789
www.nres.nhs.uk



Ms Anne MacDonald
Flat 2/1
39 Airlie Street
Glasgow
G12 9TS

Date: 30 June 2011
Your Ref.:
Our Ref.: 11/AL/0290

Enquiries to: Walter Hunter
Extension: 35680
Direct Line: 0131 465 5680
Email: walter.hunter@nhslothian.scot.nhs.uk

Dear Ms MacDonald

Study title: Evaluating the outcomes of management training in Positive Behaviour Support for managers, staff, and service users
REC reference: 11/AL/0290

Thank you for your letter of 14 June 2011, responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information was considered in correspondence by a Sub-Committee of the Scotland A REC. A list of the Sub-Committee members is attached.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation [as revised], subject to the conditions specified below.

Ethical review of research sites

Non-NHS sites

The Committee has not yet been notified of the outcome of any site-specific assessment (SSA) for the non-NHS research site(s) taking part in this study. The favourable opinion does not therefore apply to any non-NHS site at present. We will write to you again as soon as one Research Ethics Committee has notified the outcome of a SSA. In the meantime no study procedures should be initiated at non-NHS sites.

Chairman Dr Ian Zealley
Vice-Chairman Dr Colin Selby

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission ("R&D approval") should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements.

Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at <http://www.rdforum.nhs.uk>.

Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of approvals from host organisations

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document	Version	Date
REC application IRAS Form	3.1	09 May 2011
Protocol	1	04 May 2011
Investigator CV: Anne MacDonald		
Investigator CV: Peter McGill		12 May 2011
Participant Information Sheet: Staff: Experimental	2	10 June 2011
Participant Consent Form: Staff: Experimental	2	10 June 2011

Participant Information Sheet: Staff: Control	2	10 June 2011
Participant Consent Form: Staff: Control	2	10 June 2011
Participant Information Sheet: Managers: Control	2	10 June 2011
Participant Consent Form: Managers: Control	2	10 June 2011
Participant Information Sheet: Managers: Experimental	2	10 June 2011
Participant Consent Form: Managers: Experimental	2	10 June 2011
Participant Information Sheet: Service Users' Representatives: Experimental	2	10 June 2011
Participant Consent Form: Service Users' Representatives: Experimental	2	10 June 2011
Participant Information Sheet: Service Users' Representatives: Control	2	10 June 2011
Participant Consent Form: Service Users' Representatives: Control	2	10 June 2011
Participant Information Sheet: Service Users': Control	1	10 June 2011
Participant Consent Form: Service Users': Control	1	10 June 2011
Participant Information Sheet: Service Users': Experimental	1	10 June 2011
Participant Consent Form: Service Users': Experimental	1	10 June 2011
'Minimising Bias'		
Response to request for further information		14 June 2011

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Now that you have completed the application process please visit the National Research Ethics Service website > After Review

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

The attached document "*After ethical review – guidance for researchers*" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Progress and safety reports
- Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

We would also like to inform you that we consult regularly with stakeholders to improve our service. If you would like to join our Reference Group please email referencegroup@nres.npsa.nhs.uk.

REC reference number: 11/AL/0290-Please quote this number on all correspondence

With the Committee's best wishes for the success of this project.

Yours sincerely


Dr Ian Zealley
Committee Chairman
cc: Ms Nicole Palmer
The Registry
University of Kent
Canterbury
CT2 7NZ

Scotland A REC

Attendance at Sub-Committee of the REC meeting on 29 June 2011

Committee Members:

Name	Profession	Present	Notes
Mrs Angela Macpherson	Retired	No	
Dr Rachel Smith	Project Manager	Yes	
Dr Ian Zealley	Consultant	Yes	

Also in attendance:

Name	Position (or reason for attending)
Mr Walter Hunter	Committee Coordinator

Written comments received from:

Name	Position
Dr Rachel Smith	Project Manager
Dr Ian Zealley	Consultant

Appendix Four: ANOVA Tables – Managers T1-T2 (Chapter 6)

Table 13.1: Practice Leadership ANOVA Results for Time and Group

Source	Type III Sum of Squares	<i>df</i>	Mean Square	F	Sig.	Partial Eta Squared
Time	4333.365	1	433.365	2.334	0.131	0.033
Group	925.132	1	925.132	2.948	0.090	0.041

Table 13.2: Managers' PBS Knowledge ANOVA Results for Time and Group

Source	Type III Sum of Squares	<i>df</i>	Mean Square	F	Sig.	Partial Eta Squared
Time	3833.563	1	3833.563	33.153	p<0.001*	0.334
Group	6473.949	1	6473.949	24.096	p<0.001*	0.267

*Significant at p<0.05

Table 13.3: Managers' CHABA Learned Positive ANOVA Results for Time and Group

Source	Type III Sum of Squares	<i>df</i>	Mean Square	F	Sig.	Partial Eta Squared
Time	0.204	1	0.204	0.532	0.468	0.008
Group	0.021	1	0.021	0.037	0.848	0.001

Table 13.4 Managers' CHABA Learned Negative ANOVA Results for Time and Group

Source	Type III Sum of Squares	<i>df</i>	Mean Square	F	Sig.	Partial Eta Squared
Time	0.011	1	0.011	0.027	0.871	0.000
Group	0.007	1	0.007	0.012	0.915	0.000

Table 13.5: Managers' CHABA Biomedical ANOVA Results for Time and Group

Source	Type III Sum of Squares	<i>df</i>	Mean Square	F	Sig.	Partial Eta Squared
Time	0.268	1	0.268	1.053	0.309	0.016
Group	0.024	1	0.024	0.051	0.822	0.001

Table 13.6: Managers' CHABA Stimulation ANOVA Results for Time and Group

Source	Type III Sum of Squares	<i>df</i>	Mean Square	F	Sig.	Partial Eta Squared
Time	0.413	1	0.413	1.116	0.294	0.016
Group	0.281	1	0.281	0.441	0.509	0.006

Table 13.7: Managers' CHABA Emotional ANOVA Results for Time and Group

Source	Type III Sum of Squares	<i>df</i>	Mean Square	F	Sig.	Partial Eta Squared
Time	0.053	1	0.053	0.292	0.591	0.004
Group	0.583	1	0.583	1.320	0.255	0.019

Table 13.8: Managers' CHABA Physical Environment ANOVA Results for Time and Group

Source	Type III Sum of Squares	<i>df</i>	Mean Square	F	Sig.	Partial Eta Squared
Time	0.362	1	0.362	1.080	0.302	0.016
Group	0.003	1	0.003	0.007	0.935	0.000

Appendix Five: ANOVA Tables – Staff T1-T2 (Chapter 7)

Table 13.9: MTS Assistance ANOVA Results for Time and Group

Source	Type III Sum of Squares	<i>df</i>	Mean Square	F	Sig.	Partial Eta Squared
Time	857.405	1	857.405	1.986	0.163	0.028
Group	5964.091	1	5964.091	7.562	0.008	0.100

Table 13.10: MTS Positive Contact ANOVA Results for Time and Group

Source	Type III Sum of Squares	<i>df</i>	Mean Square	F	Sig.	Partial Eta Squared
Time	414.315	1	414.315	1.067	0.305	0.015
Group	6455.922	1	6455.922	5.300	0.024	0.070

Table 13.11: MTS No Contact ANOVA Results for Time and Group

Source	Type III Sum of Squares	<i>df</i>	Mean Square	F	Sig.	Partial Eta Squared
Time	254.281	1	254.281	0.657	0.420	0.009
Group	4394.235	1	4394.235	3.393	0.070	0.046

Table 13.12: Staff PBS Knowledge ANOVA Results for Time and Group

Source	Type III Sum of Squares	<i>df</i>	Mean Square	F	Sig.	Partial Eta Squared
Time	464.122	1	464.122	4.658	0.035	0.069
Group	128.537	1	128.537	0.950	0.333	0.015

Table 13.13: Staff CHABA Learned Positive ANOVA Results for Time and Group

Source	Type III Sum of Squares	<i>df</i>	Mean Square	F	Sig.	Partial Eta Squared
Time	0.702	1	0.702	1.502	0.225	0.022
Group	0.394	1	0.394	1.192	0.279	0.017

Table 13.14: Staff CHABA Learned Negative ANOVA Results for Time and Group

Source	Type III Sum of Squares	<i>df</i>	Mean Square	F	Sig.	Partial Eta Squared
Time	1.508	1	1.508	4.436	0.039	0.063
Group	0.611	1	0.611	2.047	0.157	0.030

Table 13.15: Staff CHABA Biomedical ANOVA Results for Time and Group

Source	Type III Sum of Squares	<i>df</i>	Mean Square	F	Sig.	Partial Eta Squared
Time	0.775	1	0.775	3.860	0.054	0.056
Group	0.032	1	0.032	0.127	0.723	0.002

Table 13.16: Staff CHABA Stimulation ANOVA Results for Time and Group

Source	Type III Sum of Squares	<i>df</i>	Mean Square	F	Sig.	Partial Eta Squared
Time	0.001	1	0.001	0.002	0.963	0.000
Group	4.475E-5	1	4.475E-5	0.000	0.992	0.000

Table 13.17: Staff CHABA Emotional ANOVA Results for Time and Group

Source	Type III Sum of Squares	<i>df</i>	Mean Square	F	Sig.	Partial Eta Squared
Time	0.128	1	0.128	0.426	0.516	0.006
Group	0.045	1	0.045	0.187	0.667	0.003

Table 13.18: Staff CHABA Physical Environment ANOVA Results for Time and Group

Source	Type III Sum of Squares	<i>df</i>	Mean Square	F	Sig.	Partial Eta Squared
Time	0.001	1	0.001	0.002	0.960	0.000
Group	0.427	1	0.427	0.938	0.336	0.013

Appendix Six: ANOVA Tables – Service Users T1-T2 (Chapter 8)

Table 13.19: ABC Total ANOVA Results for Time and Group

Source	Type III Sum of Squares	<i>df</i>	Mean Square	F	Sig.	Partial Eta Squared
Time	3114.591	1	3114.591	10.567	0.002*	0.136
Group	8524.013	1	8524.013	12.535	0.001*	0.158

*Significant at $p < 0.025$

Table 13.20: GCPLA Range ANOVA Results for Time and Group

Source	Type III Sum of Squares	<i>df</i>	Mean Square	F	Sig.	Partial Eta Squared
Time	10.298	1	10.298	0.576	0.451	0.008
Group	0.850	1	0.850	0.008	0.929	0.000

Table 13.21: GCPLA Busy ANOVA Results for Time and Group

Source	Type III Sum of Squares	<i>df</i>	Mean Square	F	Sig.	Partial Eta Squared
Time	7.013	1	7.013	0.728	0.397	0.010
Group	2.525E-5	1	2.525E-5	0.000	0.999	0.000