Dissemination of evidence-based practice to frontline staff working in the field of intellectual disability: An evaluation of training protocols, intervention compliance, and generalisation.

Volume I

Laura Gormley

A thesis submitted to Trinity College Dublin, the University of Dublin, in partial fulfillment of the requirements for the Degree of Doctor of Philosophy (PhD) in Psychology

2018

Supervisor: Dr. Olive Healy (Trinity College Dublin)
Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work.

I agree to deposit this thesis in the University’s open access institutional repository or allow the library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College conditions of use and acknowledgement.

I have read, and I understand the plagiarism provisions in the General Regulations of the University Calendar for the current year, found at: http://www.tcd.ie/calendar. I have also completed the Online Tutorial on avoiding plagiarism ‘Ready, Steady, Write’, located at http://tcd-ie.libguides.com/plagiarism/ready-steady-write.

__________________________________________

Laura Gormley
# Table of Contents

## Volume I

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td></td>
<td>i</td>
</tr>
<tr>
<td>Funding</td>
<td></td>
<td>iii</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td></td>
<td>iv</td>
</tr>
<tr>
<td>Conference proceedings arising from this thesis</td>
<td></td>
<td>v</td>
</tr>
<tr>
<td>List of tables</td>
<td></td>
<td>vi</td>
</tr>
<tr>
<td>List of figures</td>
<td></td>
<td>xiii</td>
</tr>
<tr>
<td>List of abbreviations</td>
<td></td>
<td>xviii</td>
</tr>
</tbody>
</table>

### Chapter 1   General introduction

- Intellectual disability: 1
- Prevalence of intellectual disability in the Republic of Ireland: 2
- Support for people with an intellectual disability: 5
- Challenging behaviour: 6
- Conclusion: 9

### Chapter 2   Frontline staff in the intellectual disability sector

- The work environment: 14
- Burnout and frontline staff: 16
- Conclusion: 17

### Chapter 3   Evidence-based practice and Applied Behaviour Analysis

- Evidence-based practice: 28
- Applied Behaviour Analysis: 29
- Conclusion: 32

### Chapter 4   Dissemination of evidence-based practice and Behaviour Skills Training

- Factors influencing the use of evidence-based practice in an intellectual disability setting: 42
- Disseminating evidence-based practice: 43
- Conclusion and research aims: 49
<table>
<thead>
<tr>
<th>Chapter 5</th>
<th>Staff training in intellectual and developmental disability settings: A systematic review</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td></td>
<td>61</td>
</tr>
<tr>
<td>Introduction</td>
<td></td>
<td>62</td>
</tr>
<tr>
<td>Method</td>
<td></td>
<td>65</td>
</tr>
<tr>
<td>Results</td>
<td></td>
<td>69</td>
</tr>
<tr>
<td>Discussion</td>
<td></td>
<td>82</td>
</tr>
<tr>
<td>Chapter 6</td>
<td>An examination of the adaptive functioning and behavioural profiles of service users, in addition to the training needs of staff providing direct care within a large intellectual disability service provider organisation</td>
<td>86</td>
</tr>
<tr>
<td>Abstract</td>
<td></td>
<td>87</td>
</tr>
<tr>
<td>Introduction</td>
<td></td>
<td>88</td>
</tr>
<tr>
<td>Phase I – Administration of measures</td>
<td></td>
<td>91</td>
</tr>
<tr>
<td>Method</td>
<td></td>
<td>91</td>
</tr>
<tr>
<td>Results</td>
<td></td>
<td>103</td>
</tr>
<tr>
<td>Phase II- Practical observations</td>
<td></td>
<td>124</td>
</tr>
<tr>
<td>Method</td>
<td></td>
<td>124</td>
</tr>
<tr>
<td>Results</td>
<td></td>
<td>126</td>
</tr>
<tr>
<td>Phase III – Evaluation of extant behaviour support plans</td>
<td></td>
<td>128</td>
</tr>
<tr>
<td>Method</td>
<td></td>
<td>128</td>
</tr>
<tr>
<td>Results</td>
<td></td>
<td>129</td>
</tr>
<tr>
<td>Discussion</td>
<td></td>
<td>130</td>
</tr>
<tr>
<td>Chapter 7</td>
<td>An evaluation of the evidence-base supporting the use of specific behavioural interventions with individuals with intellectual and developmental disabilities: A series of systematic literature reviews</td>
<td>137</td>
</tr>
<tr>
<td>Abstract</td>
<td></td>
<td>138</td>
</tr>
<tr>
<td>Introduction</td>
<td></td>
<td>139</td>
</tr>
<tr>
<td>Method</td>
<td></td>
<td>143</td>
</tr>
<tr>
<td>Results</td>
<td></td>
<td>150</td>
</tr>
<tr>
<td>Discussion</td>
<td></td>
<td>166</td>
</tr>
<tr>
<td>Chapter 8</td>
<td>Dissemination of evidence-based practice to frontline staff working in the field of intellectual disability: A pilot study</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>---------------------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Abstract</td>
<td>174</td>
<td></td>
</tr>
<tr>
<td>Introduction</td>
<td>175</td>
<td></td>
</tr>
<tr>
<td>Phase I – Behaviour Skills Training</td>
<td>179</td>
<td></td>
</tr>
<tr>
<td>Method</td>
<td>179</td>
<td></td>
</tr>
<tr>
<td>Results</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td>Phase II – Skill Generalisation</td>
<td>212</td>
<td></td>
</tr>
<tr>
<td>Method</td>
<td>212</td>
<td></td>
</tr>
<tr>
<td>Results</td>
<td>215</td>
<td></td>
</tr>
<tr>
<td>Discussion</td>
<td>215</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 9</th>
<th>Dissemination of evidence-based practice to frontline staff working in the field of intellectual disability: A randomised control trial study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>223</td>
</tr>
<tr>
<td>Introduction</td>
<td>224</td>
</tr>
<tr>
<td>Method</td>
<td>227</td>
</tr>
<tr>
<td>Results</td>
<td>261</td>
</tr>
<tr>
<td>Discussion</td>
<td>278</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 10</th>
<th>Dissemination of evidence-based practice to frontline staff working in the field of intellectual disability: An indepth analysis of the Behaviour Skills Training intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>284</td>
</tr>
<tr>
<td>Introduction</td>
<td>285</td>
</tr>
<tr>
<td>Phase I – Behaviour Skills Training</td>
<td>289</td>
</tr>
<tr>
<td>Method</td>
<td>289</td>
</tr>
<tr>
<td>Results</td>
<td>299</td>
</tr>
<tr>
<td>Phase I – Skill Generalisation</td>
<td>312</td>
</tr>
<tr>
<td>Method</td>
<td>312</td>
</tr>
<tr>
<td>Results</td>
<td>314</td>
</tr>
<tr>
<td>Discussion</td>
<td>316</td>
</tr>
<tr>
<td>Chapter 11</td>
<td>Dissemination of evidence-based practice to frontline staff working in the field of intellectual disability: An examination of service user impact</td>
</tr>
<tr>
<td>------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Abstract</td>
<td>324</td>
</tr>
<tr>
<td>Introduction</td>
<td>325</td>
</tr>
<tr>
<td>Method</td>
<td>327</td>
</tr>
<tr>
<td>Results</td>
<td>330</td>
</tr>
<tr>
<td>Discussion</td>
<td>345</td>
</tr>
<tr>
<td><strong>Chapter 12</strong></td>
<td><strong>General Discussion</strong></td>
</tr>
<tr>
<td>Contributions of current research programme and overview of findings</td>
<td>350</td>
</tr>
<tr>
<td>Limitations, challenges, and future directions</td>
<td>356</td>
</tr>
<tr>
<td>Final conclusions</td>
<td>362</td>
</tr>
</tbody>
</table>

**Volume II**

**References**

<table>
<thead>
<tr>
<th>Appendices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendix A</td>
</tr>
<tr>
<td>Appendix B</td>
</tr>
<tr>
<td>Appendix C</td>
</tr>
<tr>
<td>Appendix D</td>
</tr>
<tr>
<td>Appendix E</td>
</tr>
<tr>
<td>Appendix F</td>
</tr>
<tr>
<td>Appendix G</td>
</tr>
<tr>
<td>Appendix H</td>
</tr>
<tr>
<td>Appendix I</td>
</tr>
</tbody>
</table>
Summary

Research findings indicate that the practice of hiring personnel without suitable professional qualifications, is a significant concern within the social care sector in the Republic of Ireland (Joint Committee on Social Care Professionals, 2002; Power & Lavelle, 2011). Therefore, the purpose of the current programme of research was to examine the training needs of frontline staff employed by one of the largest intellectual disability service providers in the Republic of Ireland and subsequently develop an effective and socially valid method for disseminating a package of evidence-based practices, in order to prepare these frontline staff to support their service users with intellectual disabilities.

Chapters 1-4 present a comprehensive review of the extant literature, which is relevant to the current research programme. Chapter 1 discusses the intellectual disability diagnosis, with a particular focus on quality of life implications, prevalence rates in the Republic of Ireland, supports available to those with a diagnosis, and the impact of engaging in challenging behaviour. Chapter 2 reviews the literature pertaining to variables that impact the psychological well-being of frontline staff working in the intellectual disability sector. Chapter 3 examines the definition of evidence-based practice as well as methods for evaluating the evidence-based status of interventions. Finally, Chapter 4 discusses the factors that influence the dissemination of evidence-based practice within the intellectual disability sector. The role of effective training, and Behaviour Skills Training in particular, is described.

A systematic literature review of the training provided to staff working with people with intellectual disabilities is reported in Chapter 5. The review comprised 107 studies and results show that while the range of practices targeted for training were relatively broad, training in the application of many established evidence-based practices was not addressed.

Chapter 6 describes the results of a training needs analysis conducted among a representative sample of frontline staff and their service users within RehabCare, one of the largest providers of intellectual disability services in the Republic of Ireland. Results demonstrated a disconnect between the training received by staff and the adaptive and behavioural needs of the service users. Communication, social skills, and daily living skills were identified as priority areas for service users in the
RehabCare organisation. In order to develop a training package that addressed these priority areas, Chapter 7 presents a series of systematic literature reviews, which demonstrate the evidence-base supporting the use of reinforcement, systematic prompting, functional communication training, and task analysis with people with intellectual and developmental disabilities.

Chapters 8-11 investigate the effectiveness of Behaviour Skills Training in disseminating the four evidence-based practices, which were reviewed in Chapter 7, to frontline staff working with adults with intellectual disabilities in the RehabCare organisation. Chapter 8 presents the results of a pilot study, which was conducted to assess the feasibility of implementing an evidence-based Behaviour Skills Training intervention in the host organisation. Chapter 9 compares the Behaviour Skills Training intervention against a wait list control condition to evaluate the differential impact on target knowledge acquisition and psychological well-being among frontline staff, while Chapter 10 introduces a more thorough analysis of outcomes for the intervention group only. Finally, Chapter 11 examines outcomes for service users, whose frontline staff had participated in the Behaviour Skills Training intervention, reported in Chapters 9 and 10.

Results from the pilot study reported in Chapter 8 show that the proposed experimental design and associated methodology were robust in measuring outcomes related to staff participants. Results from the randomised control trial presented in Chapters 9 and 10 reveal that participants in the intervention condition demonstrated improved performance on knowledge assessments and this improvement maintained over a 4-week period. However, participants in the control condition did not demonstrate such an improvement. Target skill acquisition and subsequent generalisation to the workplace environment were also observed among participants in the intervention condition. Finally, observations from four case studies, reported in Chapter 11, show evidence of improvement in service users’ functional communication repertoires, as well as the quality of support they received from frontline staff, post-intervention.

Chapter 12 discusses the role of Behaviour Skills Training in the effective dissemination of evidence-based practice. Additional directions for research, which include the role of organisational factors, such as supervisory support, are also discussed.
Funding

This research programme was funded by the Irish Research Council’s Employment Based Programme.
Acknowledgements

Firstly, I would like to thank my supervisor Dr. Olive Healy. It has been a privilege to have had you as a mentor. I have learned so much from you over the past 3 years and this PhD process has been such an enjoyable and positive experience because of that. Thank-you for your support, patience and availability when I have needed your advice most.

I would also like to thank the RehabCare organisation for facilitating the research programme. The frontline staff and service managers were a pleasure to work with and are among the most dedicated and caring professionals that I have ever worked with. I am also indebted to the behaviour therapists, who provided the training package with enthusiasm and motivation; this PhD would have been impossible without your support. Finally, I would like to express my gratitude to Darragh O’Regan, the employment mentor, and Michael O’Connor, my line manager for your support and guidance over the past three years.

I am lucky to have had Brona O’Sullivan as a wonderful research assistant. You always went above and beyond to contribute to this research and no matter what the situation, you approached it with amazing enthusiasm and optimism.

Finally, I owe a wealth of gratitude to my family and friends for supporting me throughout this process. To my unbelievable husband Mark, thank-you for your patience, understanding and love. I would not be at this point without you. To my parents, thank-you for all your support and encouragement throughout the years and to my sister Olive, you believed I could do it, even when I did not.
Conference Proceedings Arising from this Thesis

Presentations:


## List of Tables

<table>
<thead>
<tr>
<th>Chapter 5</th>
<th>Staff Training in Intellectual and Developmental Disability Settings: A Systematic Review</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1:</td>
<td>A Summary of the Studies Providing Training in Positive Behaviour Support (PBS)</td>
<td>Appendix A</td>
</tr>
<tr>
<td>Table 2:</td>
<td>A Summary of the Studies Providing Training in Applied Behaviour Analysis-Based Assessment Practices</td>
<td>Appendix A</td>
</tr>
<tr>
<td>Table 3:</td>
<td>A Summary of the Studies Providing Training in Applied Behaviour Analysis-Based Antecedent Practices</td>
<td>Appendix A</td>
</tr>
<tr>
<td>Table 4:</td>
<td>A Summary of the Studies Providing Training in Applied Behaviour Analysis-Based Consequence Practices</td>
<td>Appendix A</td>
</tr>
<tr>
<td>Table 5:</td>
<td>A Summary of the Studies Providing Training in Applied Behaviour Analysis-Based “Mixed” Practices</td>
<td>Appendix A</td>
</tr>
<tr>
<td>Table 6:</td>
<td>A Summary of the Studies Providing Training in “Other” Practices</td>
<td>Appendix A</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 6</th>
<th>An examination of the adaptive functioning and behavioural profiles of service users, in addition to the training needs of staff providing direct care within a large intellectual disability service provider organisation</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 7:</td>
<td>Participant Characteristics According to Gender and Length of Service with the Organisation</td>
<td>92</td>
</tr>
<tr>
<td>Table 8:</td>
<td>Service user Characteristics According to Age and Gender (Upper Panel) and Type of Placement and Length of Time with the Host Organisation (Lower Panel)</td>
<td>93</td>
</tr>
<tr>
<td>Table 9:</td>
<td>Summary of Missing and Valid Cases across Staff-Related Measures</td>
<td>101</td>
</tr>
<tr>
<td>Table 10:</td>
<td>Summary of Missing and Valid Cases across Service User-Related Measures</td>
<td>101</td>
</tr>
</tbody>
</table>
Table 11: Spearman Correlations Between Service User-Related Measures of Challenging Behaviour and Adaptive Functioning

Table 12: Spearman Correlations Between Service User-Related Measures of Challenging Behaviour and Quality of Life

Table 13: Spearman Correlations Between Service User-Related Measures of Adaptive Functioning and Quality of Life

Table 14: Spearman Correlations Between Service User-Related Measures of Adaptive Behaviour

Table 15: Spearman Correlations Between Staff-Related Measures

Table 16: Spearman Correlations Between the Vineland-II Measure of Externalised Challenging Behaviour and Staff-Related Measures (Upper Panel); Spearman Correlations Between the Behaviour Problems Inventory – Short Form (BPI-S) Measure of Challenging Behaviour and Staff-Related Measures (Lower Panel)

Table 17: Spearman Correlations Between the Vineland-II Adaptive Functioning Domains and Staff-Related Measures (Upper Panel); Spearman Correlations Between the Life Skills Assessment (LSA) domains (Hanley, Heal, Tiger, & Ingvarsson, 2007) and Staff-Related Measures (Lower Panel)

Table 18: Spearman Correlation Results Between the Personal Outcomes Scale (POS) domains and Staff-Related Measures

Chapter 7 An evaluation of the evidence-base supporting the use of specific behavioural interventions with individuals with intellectual and developmental disabilities: A series of systematic literature reviews

Table 19: Strength of Evidence Classification System – Criteria for “Established” Interventions

Table 20: A Summary of the Studies Focusing on Systematic Prompting

Appendix C
| Table 21: | Systematic Prompting Studies Scoring at Least an SMRS Score of 3 | 151 |
| Table 22: | Total Number of Studies for which a PND Rating could be Calculated, Total Number of Studies with at Least a Minimally Effective PND and the Status of the Target Interventions for a Population of Children with Intellectual and Developmental Disabilities | 155 |
| Table 23: | Total Number of Studies for which a PND Rating could be Calculated, Total Number of Studies with at Least a Minimally Effective PND and the Status of the Target Interventions for a Population of Adults with Intellectual and Developmental Disabilities | 156 |
| Table 24: | A Summary of the Studies Focusing on Task Analysis | Appendix C |
| Table 25: | Task Analysis Studies Scoring at Least an SMRS Score of 3 | 157 |
| Table 26: | Total Number of Studies for which a PND Rating could be Calculated, Total Number of Studies with at Least a Minimally Effective PND and the Status of the Target Interventions for a Population of Children with Intellectual and Developmental Disabilities | 159 |
| Table 27: | Number of Studies (SMRS $\geq 3$) Employing Each Chaining Procedure | 159 |
| Table 28: | Total Number of Studies for which a PND Rating could be Calculated, Total Number of Studies with at Least a Minimally Effective PND and the Status of the Target Interventions for a Population of Adults with Intellectual and Developmental Disabilities | 160 |
| Table 29: | Number of Studies (SMRS $\geq 3$) Employing Each Chaining Procedure | 161 |
| Table 30: | A Summary of the Studies Focusing on Functional Communication Training | Appendix C |
| Table 31: | Functional Communication Training Studies Scoring at Least an SMRS Score of 3 | 162 |
| Table 32: | Total Number of Studies for which a PND Rating could be Calculated, Total Number of Studies with at Least a Minimally Effective PND and the Status of the Target Interventions for a Population | 164 |
of Children with Intellectual and Developmental Disabilities

Table 33: Breakdown of Behavioural Function According to Treatment Package

Table 34: Total Number of Studies for which a PND Rating could be Calculated, Total Number of Studies with at Least a Minimally Effective PND and the Status of the Target Interventions for a Population of Adults with Intellectual and Developmental Disabilities

Table 35: Breakdown of Behavioural Function According to Treatment Package

Table 36: Studies from the Systematic Prompting Review (SMRS ≥3) that Used Reinforcement as Part of the Intervention

Table 37: Studies from the Task Analysis Review (SMRS ≥3) that Used Reinforcement as Part of the Intervention

Table 38: Studies from the FCT Review (SMRS ≥3) that Used Reinforcement as Part of the Intervention

Table 39: Percentage of Studies Using Reinforcement, Percentage of Studies Using Positive Reinforcement and Percentage of Studies for which a PND Rating could be Calculated that Produced a Minimally Effective Outcome

Chapter 8 Dissemination of evidence-based practice to frontline staff working in the field of intellectual disability: A pilot study

Table 40: Summary of Participant Characteristics: Upper Panel - Geographic Location, Service Type, and Length of Service with the Organisation; Lower Panel – Gender, Education, and Job Title (N=17)

Table 41: Mean Percentage of Steps Correctly Implemented on First and Final Observations across Each Evidence-Based Practice Module (N=17)

Table 42: Number of Participants from the Overall Sample Scoring Low, Moderate or High on the MBI-HSS Dimensions (N=17) and the Number of Participants, with Both Baseline and Follow-Up
Data, who Scored Low, Moderate or High on the MBI-HSS Dimensions (N=6)

Table 43: Potential Score Range, Actual Score Range and the Mean and Median Scores for the Overall Sample (N=17) and the Participants with Both Baseline and Follow-Up Data (N=6), on the Minnesota Satisfaction Questionnaire (MSQ-SF) 207

Table 44: Potential Score Range, Actual Score Range and the Mean and Median Scores for the Overall Sample (N=17) and the Participants with Both Baseline and Follow-Up Data (N=6), on the Short Version of the Occupational Self-Efficacy Scale 208

Table 45: Potential Score Range, Actual Score Range and the Mean and Median Scores for the Overall Sample (N=17) and the Participants with Both Baseline and Follow-Up Data (N=6), on the Shortened Ways of Coping Questionnaire (SWC-R) 209

Table 46: Potential Score Range, Actual Score Range and the Mean and Median Scores for the Participants with Data on the Perceptions of Supervisory Support Scale (N=6) 210

Table 47: Summary of the Percentage of Participants Selecting a Specific Response on Each Item of the Training Acceptability Questionnaire Appendix F

Table 48: Potential Score Range, Actual Score Range and the Mean and Median Scores for Participants’ Evaluations of the Four Evidence-Based Practice Modules (N=17) 211

Table 49: Percentage of Participants Selecting a Specific Response on Each Item of the Attitudes to Evidence-Based Practice Questionnaire (N=17) Appendix G
Chapter 9  Dissemination of evidence-based practice to frontline staff working in the field of intellectual disability: A randomised control trial study

Table 50:  Participant Characteristics According to Service Site and Length of Service with the Host Organisation (N=104)  229

Table 51:  Summary of Participant Characteristics: Upper Panel - Gender, Education, and Job Title for the Intervention Group (N=54); Lower Panel - Gender, Education, and Job Title for the Control Group (N=50)  230

Table 52:  Summary of Participant Characteristics: Upper Panel - Gender, Education and Job Title for the Intervention Condition (N=54); Lower Panel – Gender, Education and Job Title for the Control condition (N=45)  244

Table 53:  Summary of Participant Characteristics: Upper Panel - Gender, Education and Job Title for the Intervention Condition (N=46); Lower Panel – Gender, Education and Job Title for the Control condition (N=46)  246

Table 54:  Summary of Missing and Valid Cases across the Measures Assessing Knowledge Acquisition: Upper Panel - Intervention Condition; Lower Panel – Control Condition  251

Table 55:  Summary of Missing and Valid Cases across the Predictor Variables: Upper Panel - Intervention Condition; Lower Panel – Control Condition  255

Table 56:  Pre-Training, Post-Training and Follow-up Breakdown of the Percentage of Intervention Group Participants Selecting a Specific Response on Each Item of the Attitudes to Evidence-Based Practice Questionnaire (N=54)  Appendix H

Table 57:  Pre-Training, Post-Training and Follow-up Breakdown of the Percentage of ‘Treatment as Usual’ Group Participants Selecting a Specific Response on Each Item of the Attitudes to Evidence-Based Practice Questionnaire (N=50)  Appendix H
Chapter 10  Dissemination of evidence-based practice to frontline staff working in the field of intellectual disability: An indepth analysis of the Behaviour Skills Training intervention

Table 58:  Interobserver agreement (IOA) and interrater reliability (Kappa) across Service Sites for all Rehearsal Sessions 295

Table 59:  Summary of Missing and Valid Cases across the Rehearsal and Feedback Sessions for the Intervention Group 297

Table 60:  Summary of Missing and Valid Cases on the Training Acceptability Survey for each Evidence-Based Practice Module 298

Table 61:  Mean Percentage of Steps Correctly Implemented on First and Final Observations Across Each Evidence-Based Practice Module 311

Table 62:  Percentage of Participants Selecting a Specific Response on Each Item of the Training Acceptability Questionnaire Appendix I

Table 63:  Potential Score Range, Actual Score Range and the Mean and Median Scores for Participants’ Evaluations of the Four Evidence-Based Practice Modules 312

Table 64:  Interobserver Agreement (IOA) and interrater reliability (Kappa) across Service Sites for all Observations 314
## List of Figures

<table>
<thead>
<tr>
<th>Chapter 5</th>
<th>Staff Training in Intellectual and Developmental Disability Settings: A Systematic Review</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1:</td>
<td>Flow diagram showing inclusion/exclusion of studies identified during database search process.</td>
<td>67</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 6</th>
<th>An examination of the adaptive functioning and behavioural profiles of service users, in addition to the training needs of staff providing direct care within a large intellectual disability service provider organisation</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 2:</td>
<td>The percentage of service users scoring “low”, “moderately low” or “adequate” across the communication domains, as measured by the Vineland-II.</td>
<td>104</td>
</tr>
<tr>
<td>Figure 3:</td>
<td>The percentage of service users scoring “low”, “moderately low” or “adequate” across the daily living skills domains, as measured by the Vineland-II.</td>
<td>105</td>
</tr>
<tr>
<td>Figure 4:</td>
<td>The percentage of service users scoring “low”, “moderately low” or “adequate” across the social skills domains, as measured by the Vineland-II.</td>
<td>105</td>
</tr>
<tr>
<td>Figure 5:</td>
<td>The percentage of service users displaying “clinically significant”, “elevated” or “average” levels of externalised challenging behaviour, according to participant reports on the Vineland-II.</td>
<td>106</td>
</tr>
<tr>
<td>Figure 6:</td>
<td>The percentage of staff receiving behaviour management and skills teaching training before working with service users and the percentage of staff receiving ongoing behaviour management and skills teaching training.</td>
<td>114</td>
</tr>
<tr>
<td>Figure 7:</td>
<td>The percentage of participants demonstrating target skills as outlined in the Discrete Trials Teaching Evaluation Form (DTTEF).</td>
<td>127</td>
</tr>
<tr>
<td>Figure 8:</td>
<td>The percentage of participants demonstrating target skills as outlined in the Discrete Trials Teaching Evaluation Form (DTTEF).</td>
<td>127</td>
</tr>
</tbody>
</table>
Figure 9: The percentage of behaviour plans rated according to the categories of the Behaviour Intervention Plan - Quality Evaluation Tool (BIP-QEII).

Chapter 7 An evaluation of the evidence-base supporting the use of specific behavioural interventions with individuals with intellectual and developmental disabilities: A series of systematic literature reviews

Figure 10: Flow diagram showing inclusion/exclusion of studies identified during database search process for systematic prompting procedures.

Figure 11: Flow diagram showing inclusion/exclusion of studies identified during database search process for systematic prompting procedures.

Figure 12: Flow diagram showing inclusion/exclusion of studies identified during database search process for functional communication training.

Chapter 8 Dissemination of evidence-based practice to frontline staff working in the field of intellectual disability: A pilot study

Figure 13: Flow diagram showing the timeline of study implementation.

Figure 14: The number of participants who received previous training in each of the four evidence-based practices (N=17).

Figure 15: Mean number of observations taken by participants to reach mastery during rehearsal and feedback for each evidence-based practice module (N=17).

Chapter 9 Dissemination of evidence-based practice to frontline staff working in the field of intellectual disability: A randomised control trial study

Figure 16: Flow diagram showing the timeline of study implementation.

Figure 17: Flow diagram showing participant attrition during the randomised control trial.

Figure 18: The percentage of participants in the intervention (N=54) condition who had already received training in each of the four evidence-based practices.
Figure 19: The percentage of participants (N=45) in the control condition who had already received training in each of the four evidence-based practices.

Figure 20: Average percentage of items answered correctly on the reinforcement knowledge assessment (based on the estimated marginal mean) for participants in the intervention condition at Time 1 (N=54) and Time 3 (N=44) and participants in the control condition at Time 1 (N=45) and Time 3 (N=46).

Figure 21: Average percentage of items answered correctly on the systematic prompting knowledge assessment (based on the estimated marginal mean) for participants in the intervention condition at Time 1 (N=54) and Time 3 (N=44) and participants in the control condition at Time 1 (N=45) and Time 3 (N=46).

Figure 22: Average percentage of items answered correctly on the functional communication training knowledge assessment (based on the estimated marginal mean) for participants in the intervention condition at Time 1 (N=54) and Time 3 (N=44) and participants in the control condition at Time 1 (N=45) and Time 3 (N=46).

Figure 23: Average percentage of items answered correctly on the task analysis knowledge assessment (based on the estimated marginal mean) for participants in the intervention condition Time 1 (N=54) and Time 3 (N=44) and participants in the control condition at Time 1 (N=45) and Time 3 (N=46).

Figure 24: Average percentage of items answered correctly on the Test of Knowledge (based on the estimated marginal mean) for participants in the intervention condition at Time 1 (N=54) and Time 3 (N=44) and participants in the control condition at Time 1 (N=45) and Time 3 (N=46).

Chapter 10 Dissemination of evidence-based practice to frontline staff working in the field of intellectual disability: An in-depth analysis of the Behaviour Skills Training intervention
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 25:</td>
<td>Flow diagram showing the timeline of study implementation.</td>
</tr>
<tr>
<td>Figure 26:</td>
<td>Average percentage of items answered correctly on the reinforcement knowledge assessment (based on the estimated marginal mean) at Time 1 (N=54), Time 2 (N=54), and Time 3 (N=44).</td>
</tr>
<tr>
<td>Figure 27:</td>
<td>Average percentage of items answered correctly on the systematic prompting knowledge assessment (based on the estimated marginal mean) at Time 1 (N=54), Time 2 (N=54), and Time 3 (N=44).</td>
</tr>
<tr>
<td>Figure 28:</td>
<td>Average percentage of items answered correctly on the functional communication training knowledge assessment (based on the estimated marginal mean) at Time 1 (N=54), Time 2 (N=53), and Time 3 (N=44).</td>
</tr>
<tr>
<td>Figure 29:</td>
<td>Average percentage of items answered correctly on the task analysis knowledge assessment (based on the estimated marginal mean) at Time 1 (N=54), Time 2 (N=50), and Time 3 (N=44).</td>
</tr>
<tr>
<td>Figure 30:</td>
<td>Average percentage of items answered correctly on the Test of Knowledge (based on the estimated marginal mean) at Time 1 (N=54), Time 2 (N=50), and Time 3 (N=44).</td>
</tr>
<tr>
<td>Figure 31:</td>
<td>Average number of observations taken by participants in the intervention group to reach mastery during the rehearsal and feedback sessions for each evidence-based practice module.</td>
</tr>
<tr>
<td>Figure 32:</td>
<td>Percentage of participants reaching mastery after the first, second and third observation during the implementation of functional communication training with a service user in the typical work environment.</td>
</tr>
<tr>
<td>Chapter 11</td>
<td>Dissemination of evidence-based practice to frontline staff working in the field of intellectual disability: An examination of service user impact</td>
</tr>
<tr>
<td>Figure 33:</td>
<td>Pre- and post-intervention scores on the first eight items of the Active Support Measure.</td>
</tr>
<tr>
<td>Figure 34:</td>
<td>Pre- and post-intervention scores on the final seven items of the Active Support Measure.</td>
</tr>
<tr>
<td>Figure 35:</td>
<td>Pre- and post-intervention scores on the first eight items of the Active Support Measure.</td>
</tr>
<tr>
<td>Figure 36:</td>
<td>Pre- and post-intervention scores on the final seven items of the Active Support Measure.</td>
</tr>
<tr>
<td>Figure 37:</td>
<td>Pre- and post-intervention scores on the first eight items of the Active Support Measure.</td>
</tr>
<tr>
<td>Figure 38:</td>
<td>Pre- and post-intervention scores on the final seven items of the Active Support Measure.</td>
</tr>
<tr>
<td>Figure 39:</td>
<td>Pre- and post-intervention scores on the first eight items of the Active Support Measure.</td>
</tr>
<tr>
<td>Figure 40:</td>
<td>Pre- and post-intervention scores on the final seven items of the Active Support Measure.</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>ABA</td>
<td>Applied Behaviour Analysis</td>
</tr>
<tr>
<td>AAC</td>
<td>Alternative and augmentative communication</td>
</tr>
<tr>
<td>AS</td>
<td>Active Support</td>
</tr>
<tr>
<td>BIP-QEII</td>
<td>Behaviour Intervention Plan – Quality Evaluation Tool II</td>
</tr>
<tr>
<td>BPI-S</td>
<td>Behaviour Problems Inventory – Short Form</td>
</tr>
<tr>
<td>BST</td>
<td>Behaviour Skills Training</td>
</tr>
<tr>
<td>CBT</td>
<td>Cognitive Behaviour Therapy</td>
</tr>
<tr>
<td>DTT</td>
<td>Discrete trial teaching</td>
</tr>
<tr>
<td>DTTEF</td>
<td>Discrete Trial Training Evaluation Form</td>
</tr>
<tr>
<td>EBP</td>
<td>Evidence-based practice</td>
</tr>
<tr>
<td>FCT</td>
<td>Functional communication training</td>
</tr>
<tr>
<td>HSE</td>
<td>Health Service Executive</td>
</tr>
<tr>
<td>ICC</td>
<td>Interclass correlation</td>
</tr>
<tr>
<td>ICESCR</td>
<td>International Covenant on Economic, Social and Cultural Rights</td>
</tr>
<tr>
<td>ICF</td>
<td>International Classification of Functioning, Disability, and Health</td>
</tr>
<tr>
<td>IOA</td>
<td>Interobserver agreement</td>
</tr>
<tr>
<td>KBMQ</td>
<td>Knowledge Behaviour Management Principles Questionnaire</td>
</tr>
<tr>
<td>M</td>
<td>Mean</td>
</tr>
<tr>
<td>MAR</td>
<td>Missing at Random</td>
</tr>
<tr>
<td>MBI-HSS</td>
<td>Maslach Burnout Inventory – Human Services Survey</td>
</tr>
<tr>
<td>MCAR</td>
<td>Missing Completely at Random</td>
</tr>
<tr>
<td>MSQ-SF</td>
<td>Minnesota Satisfaction Questionnaire – Short Form</td>
</tr>
<tr>
<td>N</td>
<td>Number of participants</td>
</tr>
<tr>
<td>NA</td>
<td>Non-applicable</td>
</tr>
<tr>
<td>NDA</td>
<td>National Disability Authority</td>
</tr>
<tr>
<td>NIDD</td>
<td>National Intellectual Disability Database</td>
</tr>
<tr>
<td>PBS</td>
<td>Positive Behaviour Support</td>
</tr>
<tr>
<td>PDC-HS</td>
<td>Performance Diagnostic Checklist – Human Services</td>
</tr>
<tr>
<td>PECS</td>
<td>Picture Exchange Communication System</td>
</tr>
<tr>
<td>PEM</td>
<td>Percentage of Data Exceeding the Median</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>PND</td>
<td>Percentage of non-overlapping data</td>
</tr>
<tr>
<td>POS</td>
<td>Personal Outcomes Scale</td>
</tr>
<tr>
<td>PRT</td>
<td>Pivotal Response Training</td>
</tr>
<tr>
<td>PSS</td>
<td>Perceptions of Supervisory Support Scale</td>
</tr>
<tr>
<td>SMRS</td>
<td>Scientific Merit Rating Scale</td>
</tr>
<tr>
<td>SSED</td>
<td>Single-subject experimental design</td>
</tr>
<tr>
<td>SWC-R</td>
<td>Shortened Ways of Coping - Revised</td>
</tr>
<tr>
<td>UK</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organisation</td>
</tr>
<tr>
<td>yo</td>
<td>Year old</td>
</tr>
</tbody>
</table>
Chapter 1

General Introduction
Chapter 1

**Intellectual Disability**

Within the intellectual and developmental disability literature, there is much debate surrounding the construct of disability, how the concept of intellectual disability fits within this construct and the use and understanding of the term intellectual disability (Schalock et al., 2007). The term disability specifies limitations in human functioning and the origins of this term and its interpretation is strongly based within the medical model (Wehmeyer et al., 2008). The International Classification of Impairments, Disabilities, and Handicaps (1980) extended the medical model’s definition of human functioning, which views disability as “a problem within the person” to include “all the life activities of a person” (Wehmeyer et al., 2008, p. 314). Subsequently, the International Classification of Functioning, Disability, and Health (ICF) broadened this perspective to account for the impact of contextual factors, such as environmental and personal influences (World Health Organization, 2001).

According to the ICF, human functioning encompasses all activities in an individual’s life, including bodily function, personal pursuits and participation (World Health Organization, 2001). Limitations to this functioning are considered a disability (World Health Organization, 2001), with intellectual disability being characterised as a sub-category of this broader grouping. Currently, the American Association on Intellectual and Developmental Disabilities (2017) operationally defines intellectual disability as a developmental “disability characterised by significant limitations both in intellectual functioning (reasoning, learning, problem solving) and in adaptive behavior, which covers a range of everyday social and practical skills… and originates before the age of 18”. However, in line with the ICF perspective on human functioning, there is a growing consensus that operational (i.e.,
diagnostic) definitions of intellectual disability must be supplemented and supported by reference to the wider influential contextual factors (Luckasson et al., 2003; Schalock et al., 2007; Wehmeyer et al., 2008). As such, intellectual disability should be conceptualised as the degree of congruency between an individual’s abilities and the context in which they function, rather than a state internal to the individual (Wehmeyer et al., 2008).

Wehmeyer et al. (2008) distinguished between a topographical definition of intellectual disability that provides specific criteria for diagnosis and classification, and a constitutive definition, which includes the broader context of intellectual disability and provides a basis for planning necessary supports. The authors list five dimensions, which they argue should be included in any constitutive definition of intellectual disability: intellectual abilities, adaptive behaviour, health, participation, and context.

The first dimension, “intellectual abilities”, refers to general mental functioning, that includes reasoning, problem-solving, forward planning, learning from experience, abstract thinking and comprehension abilities. Wehmeyer et al. (2008) stress that intelligence represents much more than academic skill or test performance. Rather, it is the ability to adapt and learn from one’s environment, engage in complex reasoning and problem-solving. The second dimension, “adaptive behaviour”, encompasses the skills (i.e., social, practical and cognitive (Luckasson et al., 2003) that an individual requires to function effectively in their daily lives. The “health” dimension is a holistic one, referring to social, mental and physical well-being. According to Wehmeyer et al. (2008), this dimension has a global influence on an individual’s functioning, given its potential to directly or indirectly impact one or more of the other four dimensions. The fourth dimension,
“participation”, characterises an individual’s ability to actively partake and contribute to the social community in which they reside (i.e., home, work, education, leisure, spiritual, cultural). Finally, “context” is divided into two components that includes environmental factors and personal factors. Environmental factors refer to the “physical, social and attitudinal environment in which people live and conduct their lives” (p. 316), while personal factors encompass elements such as gender, age, ethnicity, lifestyle, education, employment status and coping style.

The constitutive definition of intellectual disability, outlined by Wehmeyer et al. (2008), is congruent with the general conceptualisation of quality of life, which is defined by the World Health Organisation (WHO) as: “individuals’ perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns” (The World Health Organization Quality of Life Group, 1995, p. 1405). Similar to the concept of intellectual disability, quality of life is considered to be a multidimensional construct, which encompasses a range of health and non-health related domains (Eapen, Crncec, Walter, & Tay, 2014). The core domains of emotional well-being, interpersonal relationships, material well-being, personal development, physical well-being, self-determination, social inclusion and civil rights (Schalock, 2004) are considered integral to the quality of life construct and are consistent with the dimensions that underpin the constitutive definition of intellectual disability.

Schalock, Lemanowicz, Conroy, and Feinstein (1994) examined the impact of 18 predictors on the quality of life scores for a sample of 715 people with an intellectual disability. The 18 predictors were divided into three categories; personal characteristics, objective life conditions, and perceptions of significant others. A hierarchical regression analysis found that personal characteristics accounted for the
greatest amount of variance in the quality of life scores. The authors also found that within this category, level of adaptive functioning was a significant predictor of participants’ quality of life. In addition, Maes, Lambrechts, Hostyn, and Petry (2007) found that personal ability is an influential predictor of quality of life, with individuals who display lower levels of adaptive skills achieving less personal outcomes than their peers with higher functioning levels. However, according to Luckasson et al. (2003), an individual’s ability to function in daily living activities is modifiable and can be developed or improved through the implementation of personalised supports. Luckasson et al. (2003) state that when personalised supports, which foster growth, education, interests and well-being, are in place for an extended period, the functioning of a person with an intellectual disability will subsequently improve. This underpins the importance of defining and interpreting intellectual disability through constitutive terms, as the motivation behind categorising an individual’s limitations occurs in the context of identifying and developing effective, targeted supports.

**Prevalence of Intellectual Disability in the Republic of Ireland**

According to the most recent annual report, there were 28,275 people registered on the Irish National Intellectual Disability Database (NIDD) at the end of December 2016 (Doyle, Hourigan, & Fanagan, 2017). This represents a sizeable portion of the Irish population, with the figures translating into a prevalence rate of approximately 6 people per 1,000. This breaks down into 2 per 1,000 of the population presenting with a mild intellectual disability and 3.59 per 1,000 being diagnosed with a moderate, severe or profound intellectual disability.

Of the 28,275 individuals registered with the NIDD, 18,133 (64.1%) are adults, aged 20 years and older. Analysis of these data trends have shown a steady
increase in the proportion of individuals aged 35 years or over, who present with moderate, severe or profound intellectual disability. Specifically, in 1974 the percentage was 29%, rising to 38% in 1996 and 49% in 2014 (Doyle et al., 2017). Doyle et al. (2017) theorises that this trend most likely reflects an increase in the life expectancy of this cohort, a finding which has significant implications for service provision, currently and in the future.

Furthermore, 98.5% of persons registered with the NIDD (including children and adults) are currently in receipt of some type of services, and of those availing of services, 27,813 (99.8%) have participated in at least one “day programme” in 2016. This represents the highest percentage since the register was first developed in 1996. In addition, 7,582 of these individuals are also receiving full-time residential care (Doyle et al., 2017). Therefore, with a growing population of adults with moderate to severe intellectual disabilities and an increasing demand for day services and residential care, there is an ever-growing need for high quality, evidence-based supports for people with an intellectual disability in the Republic of Ireland.

**Support for People with an Intellectual Disability**

**An International Context**

The International Covenant on Economic, Social and Cultural Rights (ICESCR) states that everyone, regardless of disability, has the right to the highest quality of physical and mental health care available (United Nations Human Rights Office of the High Commissioner, 1976). The ICESCR is an international human rights law, which obliges Governments to promote and protect the human rights of people with an intellectual disability. However, over 30 years following the ICESCR, the United Nations’ Convention on the Rights of Persons with Disabilities
reported that people with disabilities were still subject to substantial inequality within society (United Nations Division of Social Policy and Development, 2008).

As a result, the overarching objective of the United Nations Convention (2008) was to move away from perceiving people with disabilities as passive recipients of charity, medical intervention and social protection to people with rights, who were entitled to ownership over their own lives and active participation within society. Furthermore, the Convention sought to reaffirm and extend the ICESCR by reinforcing people with disabilities’ entitlement to enjoy all human rights. The Convention on the Rights of Persons with Disabilities, was the first comprehensive human rights treaty enforced in the 21st century (May 2008).

An Irish Context

In 2000, the National Disability Authority (NDA) was established as the state body responsible for coordinating and developing practices and policies that directly affect people with disabilities in the Republic of Ireland (National Disability Authority, 2003). The content and nature of their policies is largely governed by international guidelines in best practice, including the ICESCR (United Nations Human Rights Office of the High Commissioner, 1976) and the Convention on the Rights of Persons with Disabilities (United Nations Division of Social Policy and Development, 2008). In 2005, the National Disability Authority reported that within Irish society, individuals with disabilities were subject to higher levels of inequality and lower levels of participation, than people without disabilities (National Disability Authority, 2005). Since 2005, it is reported that no significant improvements in attitudes towards people with disabilities has been observed, and that there is a growing trend of “less positive attitudes towards children with disabilities in mainstream education” (National Disability Authority, 2011b, p. 4). These
observations have been echoed by Watson, Lunn, Quinn, and Russell (2012), who found that Irish adults with disabilities also experience significant disadvantage in terms of social inclusion, quality of living and access to employment.

Currently the disability sector in the Republic of Ireland is undergoing a transformation. According to a report from the National Economic and Social Council (2012), “services have been free to develop as independent republics, providing care as they see fit, with little protection for vulnerable groups” (p. xiii). In recent years, there has been a movement towards establishing greater regulation. However, to-date, relatively little progress has been made, with this sector continuing to remain largely self-regulatory (National Economic and Social Council, 2012). As a consequence, the quality of service provision can vary substantially across providers (National Economic and Social Council, 2012). This is in direct conflict with international guidelines on best practice for people with disabilities (e.g., United Nations Division of Social Policy and Development, 2008; United Nations Human Rights Office of the High Commissioner, 1976).

The National Working Group was established by the Health Service Executive (HSE) in 2007 to conduct a review of day facilities for people with disabilities in the Republic of Ireland. The focus of this review was to examine practices in care provision, outcomes for service users and cost-efficiency. The resulting report strongly recommended a thorough overhaul in the provision of day services to ensure individualised supports that are flexible and responsive to each service user’s needs. Furthermore, it was recommended that these support systems should be outcome-based to ensure the highest quality of service provision (Health Service Executive, 2012a).
Similar to the National Working Group, Genio (2009) has sought to identify good practice within the disability sector in the Republic of Ireland. According to reports from both of these groups, service provider organisations for intellectual disability in the Republic of Ireland, which have been identified as supporting good practice are those also providing tailored, individualised services (Genio, 2009; Health Service Executive, 2012a). The ultimate goal of these individualised services is to facilitate service users to achieve specific outcomes, which are compatible with their ambitions and aspirations. This includes supporting each individual to live as independent a life as possible, while maximising their full potential (Department of the Taoiseach, 2006; Genio, 2009; Health Information And Quality Authority, 2009b).

**Challenging Behaviour**

Although the diagnosis of intellectual disability is characterised by deficits in intellectual and adaptive functioning (American Association on Intellectual and Developmental Disabilities, 2017), people with intellectual disability also face additional challenges that are not explicitly part of their diagnosis. One such issue is the elevated tendency for people with an intellectual disability to engage in challenging behaviour, when compared to the general population (Emerson et al., 2001; National Disability Authority, 2003; National Institute for Health and Care Excellence, 2015).

Emerson and colleagues (Emerson et al., 2001) reported results from a total population study, which examined the prevalence of challenging behaviour among people with an intellectual disability in England. The authors found that 10-15% of those in contact with educational, health or social supports for people with an intellectual disability engaged in a form of challenging behaviour. Two-thirds of this
group were male and the most frequently displayed topographies of challenging behaviour were aggression, destructive behaviour and self-injury. Similarly, the Review of Access to Mental Health Services for People with Intellectual Disabilities (National Disability Authority, 2003) reported that, between 900 and 2,400 of people with an intellectual disability in the Republic of Ireland exhibit severe challenging behaviour. However, based on a number of unpublished survey studies conducted across Irish intellectual disability services, this figure could be closer to 8,000 people engaging in some form of challenging behaviour (National Disability Authority, 2003).

Challenging behaviour is a widely recognised term but can be difficult to operationally define (Michail, 2011). It encompasses a wide range of possible topographies, including aggression, property destruction, self-injury, excessive activity, incongruous social responses, problematic sexual actions, bizarre mannerisms and ingestion of inappropriate objects (Emerson, 2001; Michail, 2011).

One operational definition of challenging behaviour presented in the literature (e.g., Dunlap et al., 2006; Michail, 2011) includes: “any repeated pattern of behaviour, or perception of behaviour, that interferes with or is at risk of interfering with optimal learning or engagement in prosocial interactions” (Smith & Fox, 2003, p. 5). It has been argued that having an intellectually disability significantly impacts all aspects of life (Maes et al., 2007; Schalock et al., 1994). However, the comorbid presence of challenging behaviour exponentially exacerbates the problems and obstacles faced by affected individuals (Emerson, 2001; Schalock et al., 1994). Furthermore, the presence of challenging behaviour substantially complicates the delivery of services and supports by organisations and frontline staff (Emerson, 2001).
In a recent study, Smyth, Healy, and Lydon (2015) found that for a sample of 138 disability support workers in the UK, there was a significant relationship between service user engagement in aggressive/destructive behaviour and perceived stress and burnout for staff. In addition, the frequency and severity of these behaviours were positively predictive of emotional exhaustion and depersonalisation. To date, the impact of exposure to challenging behaviour on the psychological well-being of staff has been well-documented and there is a growing body of research focusing on the consequences of this for those service users who display challenging behaviour (e.g., Hastings, 2002; Skirrow & Hatton, 2007; Smyth et al., 2015). For example, these individuals may be at an increased risk for negative interactions with staff (Finn & Sturmey, 2009) and inappropriate handling of challenging situations (Hastings, 2002; White, Holland, Marsland, & Oakes, 2003).

Kiernan et al. (1997) conducted a study to examine the patterns of persistence and change in challenging behaviour for 398 people with an intellectual disability in England. Their findings replicated the results from prior research by demonstrating that once challenging behaviour is established it can persist over extended periods of time. Of the 179 people identified as initially displaying more demanding challenging behaviour, 63% continued to engage in these challenging behaviours seven years later. However, Kiernan et al. (1997) also found that individuals with more demanding challenging behaviour can show improvements in this challenging behaviour over time, given the right circumstances.

There is growing evidence that the tendency to engage in challenging behaviour is a product of a complex interaction between the intellectual and adaptive limitations experienced by an individual with an intellectual disability and the environment in which they function (Emerson et al., 2001; Kiernan et al., 1997;
National Institute for Health and Care Excellence, 2015). Research supports a relationship between restricted expressive and receptive communication, deficits in social skills and a need for assistance with adaptive behaviours and the tendency to display more demanding challenging behaviour (Emerson et al., 2001; Kiernan et al., 1997).

Kiernan et al. (1997) showed that participants who developed better social and communication skills, concurrently displayed lower levels of challenging behaviour. These participants also reported improvements in their own ability to occupy themselves. A systematic literature review, published over a decade later Kurtz, Boelter, Jarmolowicz, Chin, and Hagopian (2011) supported these results by demonstrating that functional communication training, which involves explicitly replacing an inappropriate behaviour with a functionally equivalent communication behaviour, is an empirically supported treatment for people with intellectual disabilities. Therefore, programmes and supports that effectively foster adaptive skills, including communication and social skills, are likely to result in reductions in challenging behaviour.

**Conclusion**

Despite much debate, there is an emerging consensus on the definition of intellectual disability and the importance of accounting for wider contextual and environmental factors in its interpretation. In line with this development, the Convention on the Rights of Persons with Disabilities (United Nations Division of Social Policy and Development, 2008) sought to remove the perception of people with disabilities as passive recipients of support and move towards a global acceptance that people with disabilities are entitled to ownership over their own lives and active participation within society.
This movement is of particular relevance in modern Ireland, with figures indicating substantial growth in the number of adults presenting with an intellectual disability, as well the number of people availing of intellectual disability services (Doyle et al., 2017). Considering that a sizeable portion of this population are likely to engage in challenging behaviour (National Disability Authority, 2003), there is an ever-growing need for high quality, evidence-based supports for people with an intellectual disability in the Republic of Ireland. However, although research indicates that challenging behaviour is more likely to emerge in the absence of socially appropriate adaptive behaviour, interventions that effectively address these deficits should result in lower levels of challenging behaviour over time. Therefore, given the rising demand for support and the opportunity for innovation within this sector, it is the optimal time to examine the potential for adopting best-practice within the intellectual disability sector in the Republic of Ireland.
Chapter 2

Frontline Staff in the Intellectual Disability Sector
In 2010, the National Disability Authority (NDA) commissioned a report to document the views of relevant stakeholders on the potential for independence, community inclusion and enhanced quality of life for individuals with disabilities in the Republic of Ireland (Weafer, 2010). Information was gathered through the provision of 15 focus groups, which included people with an intellectual disability and frontline staff within this sector. While the intellectual disability group covered a relatively broad spectrum of ability, all participants reported requiring some element of support to achieve their own personal ambitions and aspirations. However, this group listed limited assistance from frontline staff as one of the most important impediments to this goal (Weafer, 2010).

Despite recognising the need for progressive approaches within the intellectual disability sector (Weafer, 2010) many frontline staff fail to demonstrate an understanding that their own behaviour can have a significant impact on service user outcomes and the overall standard of service provision (Campbell, 2010). This is in direct contrast to the widespread acceptance among the academic research community of the critical influence of frontline staff on the service users they support (Finn & Sturmey, 2009; Jahr, 1998; Schepis, Reid, Ownbey, & Parsons, 2001). Frontline staff are an extremely valuable asset within the intellectual disability sector, with the quality of their work dictating the overall standard of care within an organisation (Devereux, Hastings, & Noone, 2009; Salyers et al., 2015). Research has shown that compromised skill sets among staff can adversely affect staff-service user relations (Finn & Sturmey, 2009), the learning opportunities available (Schepis et al., 2001) and quality of life outcomes for service users (Jahr, 1998).

Allen (1999) analysed the differential experiences of two groups of individuals with intellectual disabilities; one group successfully maintained their
community placement, while the second group experienced placement breakdown. One important factor, which was shown to help sustain the first group in their less restrictive placement, was the ability of frontline staff to support their service users through acute challenging situations. As such, the behaviour of the frontline staff played a vital role in placement success, despite the obstacles encountered.

**The Work Environment**

Hamilton and Atkinson (2009) reported a qualitative study that examined the life stories of adults with intellectual disabilities in the Republic of Ireland. Qualitative analyses revealed a consensus among this cohort that positive support from frontline staff “sustained them during… difficult times” (Hamilton & Atkinson, 2009, p. 316). In addition, McManus, Feyes, and Saucier (2011) found that positive experiences among the general population with people with intellectual disabilities consistently predicted better attitudes towards this group, as well as increased support for their private and civil rights, including their integration into educational and workplace environments. However, consistent positive interactions may be difficult to achieve, given that the literature repeatedly classifies intervention settings for people with intellectual disabilities as stressful environments (Hatton et al., 1999; Hensel, Lunsky, & Dewa, 2012; Mitchell & Hastings, 2001; Robertson et al., 2005; Rose & Rose, 2005).

According to Maslach (1999), the demands associated with working with people with an intellectual disability are often particularly stressful for frontline staff. Situations, in which stress is not successfully managed, may result in staff burnout over time (Devereux, Hastings, & Noone, 2009). This can be costly for service provider organisations, in terms of financial impact and impaired service delivery.
Staff in intellectual disability settings are at an increased risk for illness, absenteeism and employee turnover (Rose, 1995), as well as reduced empathy for service users (Åström, Nilsson, Norberg, Sandman, & Winblad, 1991), deteriorating interpersonal relationships (Lawson & O'Brien, 1994), maladaptive practices (White et al., 2003) and inappropriate handling of challenging situations with service users (Hastings, 2002),

Smyth et al. (2015) analysed stress, burnout and work commitment among 138 disability support workers in the UK. Participants, who were employed as frontline staff in residential homes for adults with intellectual and developmental disabilities, reported their own levels of perceived stress, burnout and commitment to work. The association between these variables and the frequency and severity of the challenging behaviour that staff were exposed to was analysed. Results showed a significant relationship between stress, burnout and an individual’s commitment to work, with staff exposed to challenging behaviour being at an inflated risk for emotional exhaustion and depersonalisation.

**Burnout and Frontline Staff**

Burnout has been described as a multidimensional construct, defined by emotional exhaustion, depersonalisation and low personal accomplishment (Maslach, 1993). According to Maslach (1993), emotional exhaustion develops as a result of prolonged exposure to stress and the experience of having inadequate emotional resources to cope with a stressful situation. Depersonalisation, on the other hand, is linked to dysfunctional and cynical, interpersonal relationships, while personal achievement depends upon an individual’s own self-perception and self-evaluation. In 1999, Maslach proposed a theory of burnout, which claims that staff who do not
have the necessary resources to support them in their work with service users, are likely to experience elevated levels of stress. If these levels increase in intensity over time, the result is emotional exhaustion, which can subsequently lead to negative and demeaning interactions with service users (Maslach, 1999). At the core of this theory is the premise that the relationship between staff and service users may function as a major source of stress, particularly when perceived work demands exceed resources. Although Maslach’s theory to date has not been rigorously tested, there is growing support for this theoretical position within the literature.

A consistent correlation between exposure to challenging behaviour and staff stress and emotional exhaustion has been demonstrated (e.g. Hastings, 2002; Jenkins, Rose, & Lovell, 1997; Smyth et al., 2015). Jenkins et al. (1997) examined levels of stress among a sample of 78 direct-care staff, working with people with intellectual disabilities in a residential setting. Those staff who were employed to work with residents who engaged in challenging behaviour displayed higher levels of anxiety and reported lower levels of job satisfaction and perceived support than those staff who were not exposed to challenging behaviour. Similarly, Dyer and Quine (1998) assessed the occupational stress of 80 frontline staff, supporting individuals with intellectual disabilities. Within this sample there was a significant relationship between staff burnout and various service user characteristics, including adaptive functioning, communication difficulties and engagement in self-injury and other forms of challenging behaviour. In recent years a growing body of literature on burnout among frontline staff in the intellectual disability sector has emerged as a key measure of psychological well-being (Kozak, Kersten, Schillmöller, & Nienhaus, 2013).
The Maslach Burnout Inventory (Maslach, Jackson, & Leiter, 1996), has been used to quantify the three proposed dimensions of burnout (e.g. Gibson, Grey, & Hastings, 2009; Griffith, Barbakou, & Hastings, 2014; Smyth et al., 2015). To date, research involving this measure, has focused on identifying predictors of burnout in intellectual disability support staff. The most influential factors identified, include staffs’ own personal characteristics, such as coping style (Hastings & Brown, 2002a) and job satisfaction (Kozak et al., 2013; Skirrow & Hatton, 2007); organisational influences, such as supervisory support (Gibson et al., 2009); and service user characteristics, including levels of adaptive functioning (Buckhalt, Marchetti, & Bearden, 1990) and engagement in challenging behaviour (Hastings, 2002; Jenkins et al., 1997; Smyth et al., 2015).

**Development of Burnout**

According to work stress theories, perceived work-related demands are strongly associated with the development of stress and burnout (Devereux, Hastings, & Noone, 2009). However, this relationship can be positively or negatively influenced by the resources available to an individual. These resources can be physical, psychological, social or organisational in nature (Devereux, Hastings, Noone, Firth, & Totsika, 2009; Lazarus, 1999; Lazarus & Folkman, 1984).

**Work-related demands.** As previously discussed, frontline staff in the intellectual disability sector play a significant role in the lives of the people they support. It has been proposed that they shape the experiences and behaviour of their service users through the opportunities, feedback and consequences they deliver (Finn & Sturmey, 2009; Hamilton & Atkinson, 2009; Jahr, 1998; Mansell & Beadle-Brown, 2012). Furthermore, interactions with service users can have a pervasive
impact on the psychological well-being and functioning of frontline staff (e.g., McManus et al., 2011; Smyth et al., 2015; Zaharia & Baumeister, 1978). For example, Zaharia and Baumeister (1978) found service user skill acquisition to be positively correlated with staffs’ psychological well-being, with greater service user progress associated with higher levels of staff satisfaction. On the other hand, “emotionally charged client contacts” (Kozak et al., 2013, p. 111) have been consistently linked to burnout among intellectual disability staff (Kozak et al., 2013). As a result, interactions with service users are considered one of the principal work-related demands faced by intellectual disability staff.

Statistics indicate that people with an intellectual disability are at an increased risk for engagement in challenging behaviour, with published figures reported at approximately 10-15% of this population (Emerson et al., 2001; National Disability Authority, 2003). Challenging behaviour is any behaviour that places an individual or others in danger and/or interferes significantly with social interactions and restricts access to the community (Emerson, 2001; Smith & Fox, 2003). Therefore, given its prevalence within the intellectual disability population and short- and long-term consequences of engagement, challenging behaviour represents a major issue for frontline staff in their role as service providers (Devereux, Hastings, Noone, et al., 2009; Kozak et al., 2013).

Recent research has begun to investigate the effect of exposure to challenging behaviour on the psychological well-being of staff. For example, Hastings (2002) has provided a critical evaluation of the potential relationship between service user challenging behaviour and staff stress. While strong evidence was provided for a correlation between these variables, the empirical support for a causal relationship was limited. Therefore, alternative explanations could not be disregarded (Hastings,
2002). Subsequently, Skirrow and Hatton (2007) provided a further systematic literature review on burnout in intellectual disability staff. Findings revealed inadequate evidence for a predictive relationship between staff exposure to challenging behaviour and symptoms of burnout. However, both Hastings (2002) and Skirrow and Hatton (2007) caution against uncritically accepting the non-existence of a relationship. Both reviews (2002; 2007) claim that many included studies, which reported no association, were weakened by a variety of methodological issues.

While the research in this area is restricted by the number of high quality studies evaluating causation (Hastings, 2002; Skirrow & Hatton, 2007), studies supporting a correlational relationship continue to emerge. Kozak et al. (2013) examined potential predictors of burnout in intellectual disability staff. Their sample comprised 409 frontline staff, working across 30 residential settings in Germany. Although a predictive relationship between staff burnout and service user challenging behaviour was not shown, there was evidence of a significant correlation between perceived stress among frontline staff and aggression displayed by service users. In total, 48% of participants reported moderate to high levels of stress as a direct result of their service users’ engagement in challenging behaviour.

More recently, Smyth et al. (2015) analysed the impact of service user challenging behaviour on the stress, burnout and work commitment of 138 frontline staff. The study found that staff exposed to challenging behaviour were at an increased risk for emotional exhaustion and depersonalisation. Previous research has shown that engagement in challenging behaviour potentially exposes service users to negative interactions with staff (Finn & Sturmey, 2009) and inappropriate handling of challenging situations (Hastings, 2002; White et al., 2003). Therefore, although
findings from Kozak et al. (2013) and Smyth et al. (2015) were correlational, the association between exposure to challenging behaviour and the development of emotional exhaustion and depersonalisation could explain why service users displaying challenging behaviour may encounter reduced empathy and inappropriate treatment from staff (Åström et al., 1991; Hastings, 2002; Lawson & O’Brien, 1994; White et al., 2003).

**Resources.**

*Coping style.* The psychological resource of coping can be defined as “the individual’s response to a demand, and is related to… psychological flexibility” (Devereux, Hastings, Noone, et al., 2009, p. 368). The theory proposed by the cognitive-behavioural approach to work stress claims that an individual’s coping ability plays a pivotal role in stress development (Lazarus & Folkman, 1984). It is proposed that there is a two-step process, which involves the appraisal of the environmental threat, followed by an evaluation of resources available to manage this threat (Lazarus, 1999). Within this framework, it is suggested that a person’s coping style plays an important role in mediating the impact of the stressor on the individual’s emotional state (Lazarus, 1999).

According to Lazarus (1999), there are two styles of coping, practical and emotional. Practical coping skills are a solution-based coping style that focus on altering the situation for the better, while emotional coping strategies tend to focus on ways to reduce any distress experienced. The value of each coping style is largely dependent on the nature of the situation encountered and the resources available at that time. In recent years, a number of studies have examined coping styles, within the intellectual disability sector and their impact on staff and service users.
Mitchell and Hastings (2001) investigated the coping strategies employed by staff working in community-based services for people with intellectual disabilities. The authors found that emotional coping styles predicted higher levels of exhaustion and lower levels of personal accomplishment among staff. Conversely, problem-solving coping strategies were associated with elevated levels of personal accomplishment. Similarly, Griffith et al. (2014) demonstrated that a wishful thinking, emotionally-based coping style positively predicted emotional exhaustion and depersonalisation, while negatively predicting personal accomplishment, among a sample of 45 Applied Behaviour Analysis therapists working with children with autism. The results from Mitchell and Hastings (2001) and Griffith et al. (2014) are consistent with those of Devereux, Hastings, Noone, et al. (2009) who also found that the use of wishful thinking, when faced with work demands (e.g., service user engagement in challenging behaviour), is a significant risk factor for burnout in intellectual disability support staff.

**Job satisfaction.** A second psychological factor, that can serve a protective function in the development of burnout, is job satisfaction. It has been defined in the literature as a positive or negative emotional state that arises through an appraisal of one’s job (Locke, 1969) and has been widely researched in organisational psychology (Connolly & Viswesvaran, 2000). As a result, there is a general consensus that job satisfaction has a role to play in the interaction between work demands and the development of burnout.

For example, Dyer and Quine (1998) investigated the relationship between job satisfaction, burnout and work demands. The authors surveyed a cross section of 80 staff, who were employed by a government funded, community-based service for people with an intellectual disability in the UK. Results indicated that there was a
significant negative correlation between job satisfaction and burnout, with higher demands in the work environment (e.g., supporting demanding service users) adversely impacting both variables. Finally, there was also evidence that social support fosters job satisfaction and protects against burnout.

Kozak et al. (2013) also examined the relationship between job satisfaction and burnout among 409 intellectual disability staff. The results of this cross-sectional study showed a negative correlation between burnout and job satisfaction, with participants who scored higher on the burnout measure also reporting a greater intention to leave their current job. These findings correspond to the results of studies, which focus on the psychological well-being of staff within the health sector, in general. For example, Liakopoulou et al. (2008) explored the potential relationship between job satisfaction ratings and burnout levels among staff working in paediatric oncology units. Results showed low job satisfaction as a result of burnout. The congruency in results across disciplines consolidates the limited findings in relation to these variables within the intellectual disability literature.

Perceived supervisory support. According to Devereux, Hastings, Noone, et al. (2009) support or perceived support is a social and organisational resource that can help individuals manage demands in the workplace. Given the tendency for frontline staff within the intellectual disability sector to work as part of a team (Devereux, Hastings, Noone, et al., 2009; Rose, Ahuja, & Jones, 2006), social support and perceived supervisory support, specifically, could play an important role in psychological well-being. Fukui, Rapp, Goscha, Marty, and Ezell (2014) claim that there is growing evidence that the quality of staff engagement with service users can be significantly influenced by the quality of supervision that frontline staff receive. In addition, a series of evidence-based practice implementation reports from
the United States have shown that high quality supervision is integral to the success of programme outcomes (Rapp et al., 2010). Supervision is a multi-dimensional process that involves managing social and emotional challenges that confront staff, as well as attending to issues of productivity, efficiency and skill development (Fukui et al., 2014). In their development of the Perceptions of Supervisory Support Scale, Fukui et al. (2014) focused on the following types of supervisory support: emotional support for staff, support to help staff foster service users’ abilities to achieve their goals and support to promote staffs’ professional development.

Research has shown reduced burnout in staff who perceived that they could confide in their supervisors (Ito, Kurita, & Shiiya, 1999) and who reported experiencing support from their colleagues (Dyer & Quine, 1998). Gibson et al. (2009) surveyed 81 behaviour therapists, who worked with children with autism and found that perceived supervisory support significantly predicted reduced emotional exhaustion and depersonalisation and elevated personal accomplishment and self-efficacy. Furthermore, supervisory support appeared to act as a protective factor from low personal accomplishment, when work demands were high. In contrast, Griffith et al. (2014) reported no relationship between supervisory support and staff psychological well-being, including burnout, for their sample of 45 behaviour therapists.

Devereux, Hastings, Noone, et al. (2009) also found no evidence that social support, more generally, functions as a mediator between work demands and burnout. In addition, they reported an unusual pattern in the moderating relationship between social support and personal accomplishment. In situations of low demand and high levels of perceived support, participants reported high levels of accomplishment. However, in situations of high demand, participants who reported
low levels of support scored higher on personal accomplishment. The authors hypothesised that in this case, participants were attributing success to their own abilities rather than collegiate support. Considering the limited research and contrasting findings in the literature, additional studies are required to establish what role if any, that social support and supervisory support in particular play in staffs’ psychological well-being.

**Conclusion**

Frontline staff are considered one of the most valuable resources within the intellectual disability sector. They play an important role in the standard of service provided by an organisation and are an integral part of their service users’ social network (Devereux, Hastings, Noone, et al., 2009). However, the nature of this work is considered stressful (e.g., Rose & Rose, 2005). Daily interactions can be emotionally and psychologically draining (Schaufeli & Enzman, 1998), leading to stress and potentially burnout, over time (Devereux, Hastings, & Noone, 2009). Work-stress theories propose that an individual’s perception of work-related demands, plays an important role in the development of stress and burnout. However, the trajectory of this relationship can be altered depending on the nature of the psychological and organisational resources available to an individual (Devereux, Hastings, & Noone, 2009; Lazarus, 1999; Lazarus & Folkman, 1984).

Interactions with service users have been identified as one of the most influential work-related demands faced by intellectual disability staff. People with intellectual disabilities experience substantial deficits in their adaptive behaviour repertoire (American Association on Intellectual and Developmental Disabilities, 2017), which has been associated with an increased tendency to engage in
challenging behaviour (Emerson et al., 2001; Kiernan et al., 1997). In turn, exposure to service user challenging behaviour has been correlated with higher levels of burnout among intellectual disability staff (Hastings, 2002; Kozak et al., 2013; Smyth et al., 2015) and places the service users themselves at an increased risk for substandard practice from frontline staff (White et al., 2003).

However, these outcomes are not inevitable. The psychological and organisational resources available to an individual can regulate the impact of work-related demands, such as challenging behaviour, on staff and service user well-being. For example, Dyer and Quine (1998) found that higher levels of job satisfaction were associated with lower burnout levels. Furthermore, social support was shown to promote job satisfaction and protect against burnout. It is also important to note that deficits in adaptive skills and engagement in challenging behaviour by service users can be managed and improved (Kiernan et al., 1997; Kurtz et al., 2011), thereby reducing the work-related demands faced by frontline staff. Consequently, staff training in practices that have been shown to systematically improve service user outcomes (e.g., Carr et al., 2002; Kurtz et al., 2011; McClean et al., 2005; Wehmeyer & Schalock, 2001) should enhance staff-service user interactions (Hastings & Brown, 2002b; Jenkins et al., 1997), which subsequently creates an opening for additional service user learning opportunities (Rose, Jones, & Fletcher, 1998) and positive experiences for staff and service users, alike.
Chapter 3

Evidence-Based Practice and Applied Behaviour Analysis
While each health care profession (e.g., medicine, psychology, speech and language therapy) has their own definition for evidence-based practice, a common theme across disciplines has emerged; “Evidence-based practice… is the integration of the best available research with clinical expertise in the context of patient characteristics, culture and preferences” (American Psychological Association, 2006, p. 273). According to the American Psychological Association, evidence-based practice has been at the forefront of psychology since the birth of applied psychology, circa 1896 (American Psychological Association, 2006). Within this field, the goal of evidence-based practice is to “promote effective psychological practice and enhance public health” (p. 273) by using assessments, programmes and interventions, which are empirically validated (American Psychological Association, 2006). In 1995, the American Psychological Association approved a set of guidelines, to support the integration of best available research with clinical practice. According to these guidelines, all psychological interventions must be evaluated in terms of their efficacy and clinical utility. Efficacy refers to the degree to which causal relationships between interventions and outcomes can be established, while clinical utility involves a consideration of the generalisability, feasibility and efficiency of such research-based interventions (American Psychological Association, 1995).

The American Psychological Association does not advocate for any one type of research evidence. It is the position that multiple research designs (e.g., single-case experimental design, randomised control trials, meta-analyses) contribute to the evidence-base and are differentially suited to particular problems and areas of interest. The association maintains that psychologists must be skilled in recognising
the strengths and weaknesses of evidence obtained from various modes of research, while maintaining an open attitude to data and a willingness to test alternative hypotheses that may challenge existing theoretical positions (American Psychological Association, 2006). As far back as 1947, the American Psychological Association had developed policy mandating that doctoral level psychologists should have both scientific and practitioner training (American Psychological Association, 2006). As such, the field of psychology strives to encourage professionals to use evidence-based practice as a tool, to enhance service provision, through an “atmosphere of mutual respect, open communication and collaboration among all stakeholders” (p. 281) (American Psychological Association, 2006).

Criteria for Establishing Evidence-Based Practice

Numerous approaches to the evaluation of the scientific evidence supporting psychological treatments and interventions have been proposed in recent times (e.g., American Psychological Association, 1995; Chambless & Hollon, 1998; Chambless & Ollendick, 2001; Kratochwill & Stoiber, 2002; National Autism Center, 2009; 2015; Reichow, 2011; Shernoff, Kratochwill, & Stoiber, 2002; Task Force on Promotion and Dissemination of Psychological Procedures, 1995; Wong et al., 2014). The system of evaluation developed by Chambless and Hollon (1998) has had a particularly far-reaching impact within the field of Psychology and Applied Behaviour Analysis (ABA), more specifically (Carr, Severtson, & Lepper, 2009; Kerwin, 1999; Kurtz et al., 2011; Newsom & Hovanitz, 2005; Sturmey & Didden, 2014). Building directly on the work of the Task Force on Psychological Intervention Guidelines (American Psychological Association, 1995) and the Task Force on Promotion and Dissemination of Psychological Procedures (Task Force on Promotion and Dissemination of Psychological Procedures, 1995), Chambless and
Hollon (1998) developed a comprehensive structure for the evaluation of psychological intervention research. In line with the American Psychological Association guidelines (American Psychological Association, 1995), this structure accounted for two principal areas: efficacy and effectiveness.

In determining treatment efficacy, Chambless and Hollon (1998) describe criteria for research design, including the need for replication by independent research teams and detailed descriptions of outcome assessments, selected samples, and treatment implementation. In particular, the authors argue for the necessity of treatment manuals, stating that “research projects for which a treatment manual was not written and followed are of limited utility in terms of assessment of treatment efficacy” (p. 11). While not as detailed as the ‘efficacy’ recommendations, the section pertaining to ‘effectiveness’ stresses the importance of determining the clinical value of an intervention and establishing the usability of the practice in clinical and applied settings. In order to achieve this, Chambless and Hollon (1998) advise assessing the generalisability of the intervention, in terms of both population and setting, the cost-effectiveness of package implementation and the treatment feasibility, including ease of dissemination.

The Chambless and Hollon (1998) criteria also allow for treatments to be classified into one of three categories: well-established, probably efficacious and experimental. An intervention, which is considered “well-established” must be supported by a substantial amount of evidence (at least nine studies for single-subject experimental designs or two well-designed between-group designs), prove more efficacious than a comparison treatment or no treatment, and be replicated by at least two independent research teams. In addition, the sample characteristics must be comprehensively detailed and the treatment manualised (Carr et al., 2009; Chambless
& Hollon, 1998; Chambless & Ollendick, 2001). Criteria also require that single-subject experiments demonstrate a stable baseline trend over a sufficient period to “rule out pre-existing trends” (Newsom & Hovanitz, 2005, p. 39). Acceptable single-subject designs include for example reversal (ABAB) designs and multiple-baseline designs across behaviours, settings, or participants.

A treatment classified as “probably efficacious” has less supporting evidence than a “well-established” intervention but must still be validated as superior to a wait-list control group, in two group design experiments or 3-8 single-subject experimental design studies. In addition, the treatment must meet all other criteria specified for “well-established” interventions. Any treatments that do not meet the criteria for “well-established” or “probably efficacious” are classified as “experimental” (Carr et al., 2009; Chambless & Hollon, 1998; Chambless & Ollendick, 2001). According to the criteria, a treatment can change its classification should additional evidence become available (Carr et al., 2009).

**Applied Behaviour Analysis**

Applied Behaviour Analysis is defined as the science, which employs tactics derived from the principles of behaviour to systematically produce meaningful and significant changes in the adaptive and maladaptive behaviour of an individual (Cooper, Heron, & Heward, 2014). Applied Behaviour Analysis is one of three major divisions of behaviour analysis (Baer, Wolf, & Risley, 1987; Cooper et al., 2014) and its emergence has been partially linked to the publication of a study examining the impact of strategies, which were based on behavioural principles and employed by frontline staff, on the functioning of patients with intellectual disabilities and psychotic disorders (Ayllon & Michael, 1959). In 1968, a seminal
Chapter 3

paper was published outlining the standards for research and practice in Applied Behaviour Analysis (Baer, Wolf, & Risley, 1968). The authors proposed seven dimensions for judging applied behaviour analytic activity, which remain relevant in modern practice (Baer et al., 1968; 1987; Cooper et al., 2014).

These dimensions specify that any applied behaviour analytic research or practice must be applied and analytical in nature, focused on observable behaviour, technological in description, linked to the conceptual systems of the science and produce effective and generalisable outcomes (Baer et al., 1968; 1987). The first dimension stipulates that the activity must be applied in nature. As such, the behaviour of interest should be selected primarily because of its social importance, rather than its importance to theoretical development. This links with the third behavioural dimension, which dictates that the target behaviour must be observable and measurable.

Baer et al. (1968) also insist that experimental control is demonstrated when manipulating behaviour (i.e., the activity is analytic in nature). They recommend using either a reversal or multiple-baseline design to achieve this. With a reversal design the objective is to clearly show that when a particular variable is introduced the behaviour occurs, and when this variable is removed the behaviour disappears. Multiple-baseline designs are useful when behaviours are considered irreversible (e.g., skill acquisition). According to this experimental design, a number of behaviours are measured across time to establish baseline levels. Subsequently, the target variable is systematically introduced across each behaviour. For experimental control to be demonstrated there should be a change in the behaviour, compared to baseline, for which the target variable was implemented but little to no change across the behaviours that were not yet subject to the target variable. This pattern should be
replicated as the target variable is introduced with the remaining behaviours (Baer et al., 1968; Cooper et al., 2014).

The next two dimensions address the technological description of the behaviour analytic activity, as well as the conceptual link to scientific theory. In this context, a description is considered technological if a trained individual can replicate the procedure by simply reading the instructions provided. Furthermore, these technological descriptions must be linked back to the conceptual principles that provide the foundations for the behavioural science (e.g., reinforcement, punishment, extinction). Finally, the implementation of a behavioural intervention must produce an outcome that is large enough to have practical value (i.e., be effective) and this outcome must maintain over time and generalise to novel environments or related behaviours (i.e., display generalisable outcomes) (Baer et al., 1968).

**Applied Behaviour Analysis and Evidence-based Practice**

The seven dimensions outlined by Baer et al. (1968) remain functionally relevant in current behaviour analytic activity (Baer et al., 1987; Cooper et al., 2014) and provide the foundation and structure for high quality research and practice in this science (Morris, Smith, & Altus, 2005). As such, Applied Behaviour Analysis is ideally positioned as a scientific framework to develop, deliver and evaluate evidence-based services and supports within the intellectual disability sector. Within this context, an intervention will only be considered evidence-based if there is repeated, substantial demonstrations of a causal relationship between an operationally (i.e., technologically) defined intervention and a socially significant outcome of practical value (Fisher, Piazza, & Roane, 2010; Horner & Kratochwill, 2012).
However, applied behavior analytic research typically relies on single-subject experimental design to establish functional relations and experimental control. This approach can be limited in terms of the generalisability of the findings and the need for systematic replication (Horner & Kratochwill, 2012). Applied Behaviour Analysis has tackled these criticisms resolutely through the volume of systematic experimental investigations and replications that have been published since the 1960’s. Given the quantity of research produced, many groups have begun to assimilate the findings to help clarify the evidence-base within this field (e.g., Grey & Hastings, 2005; Kurtz et al., 2011; Peters-Scheffer, Didden, Korzilius, & Sturmey, 2011; Sturmey & Didden, 2014).

For example, in 2005, Grey and Hastings published a critical review of the evidence-based practices used with people with an intellectual disability and co-occurring challenging behaviour. Findings demonstrated that Applied Behaviour Analysis had a substantial research base, with strategies based on this science (e.g., functional communication training, choice provision, reinforcement) consistently producing positive results for individuals with intellectual disabilities and behavioural issues. Furthermore, the authors found evidence to support the use of applied behaviour analytic strategies as a service model for this population (Grey & Hastings, 2005).

Recently, a number of independent studies (e.g., Carr et al., 2009; Kerwin, 1999; Kurtz et al., 2011) employed the established guidelines that govern the evidence-based status of psychological interventions (Chambless & Hollon, 1998; Task Force on Promotion and Dissemination of Psychological Procedures, 1995) to examine practices within the field of Applied Behaviour Analysis. For example, Kerwin (1999) reviewed 79 psychosocial and behavioural studies, which involved
treatments to target paediatric feeding difficulties. Using an adapted version of these standards they identified two well-established interventions for severe feeding problems: (1) positive reinforcement of appropriate feeding responses, combined with ignoring inappropriate responses, and (2) differential reinforcement with manual guidance.

Similarly, Carr et al. (2009) and Kurtz et al. (2011) used the Chambless and Hollon (1998) criteria to systematically assess the empirical support for the use of non-contingent reinforcement and functional communication training with people with intellectual disabilities. Carr et al. (2009) found sufficient evidence to classify non-contingent reinforcement (i.e., response-independent fixed-time reinforcer delivery, combined with extinction and schedule thinning) as a well-established intervention in the treatment of challenging behaviour displayed by individuals with an intellectual disability. Kurtz et al. (2011) also identified functional communication training, which involves explicitly replacing a challenging behaviour with a functionally equivalent communication response, as an evidence-based strategy for the reduction of challenging behaviour and the development of adaptive behaviour, within this population.

Most recently, Sturmey and Didden (2014) published an edited volume that reviewed the evidence base for interventions used in the treatment of behavioural and psychiatric disorders, which frequently co-occur with intellectual disabilities. The Chambless and Hollon (1998) criteria for evidence-based practice was employed to evaluate proposed interventions for each of the disorders discussed. Healy, Lydon, and Murray (2014) reported that function-based behavioural interventions, including non-contingent reinforcement, functional communication training, and differential reinforcement strategies were effective treatments for aggression among individuals.
with an intellectual disability. Likewise, Sigafoos, O'Reilly M, Lancioni, Lang, and Didden (2014) found that applied behaviour analytic approaches, which address the function of self-injurious behaviour, were consistently effective in reducing the occurrence of this disorder. The authors also reported that intervention packages that included extinction of self-injurious behaviour and differential reinforcement of an alternative behaviour were the most “well-established” treatments for self-injurious behaviour within the intellectual disability population. Finally, behaviour modification procedures (i.e., extinction and reinforcement), which were based on a pre-treatment functional assessment, were also classified as effective interventions for individuals with an intellectual disability who experience sleep disturbances (Didden et al., 2014) or food refusal (Williams, Seiverling, & Field, 2014).

**Comprehensive Analyses of the Evidence-Base within Applied Behaviour Analysis**

The rationale behind the identification of evidence-based practice is to provide service users and practitioners with the necessary information to allow them make informed treatment decisions, either for themselves or the people they support. (National Autism Center, 2009). While research has repeatedly and consistently shown the pivotal role played by Applied Behaviour Analysis in the development of adaptive behaviour among children and adults with intellectual disabilities (e.g., Bailey & Angell, 2005; Cavkaytar, 2012; Kurtz et al., 2011; Peters-Scheffer et al., 2011), the volume of studies can make the process of identifying effective and suitable practices extremely difficult for consumers to navigate.

As a result, the National Professional Development Centre was funded by the United States Department of Education to promote the use of evidence-based
practices for children and young adults with autism spectrum disorder (birth-22 years; Wong et al., 2014). As part of this process, rigorous criteria were employed to conduct an extensive evaluation of the relevant peer-reviewed research. The investigation highlighted 27 interventions that met criteria for evidence-based practice. The majority of these interventions, including extinction, reinforcement, prompting, task analysis, functional behaviour assessment, discrete trial training and functional communication training, were developed and tested within the field of Applied Behaviour Analysis (Wong et al., 2014).

Around the same time, the National Autism Centre commissioned a two-phase quantitative analysis and evaluation of interventions for children, adolescents and adults with autism (National Autism Center, 2009; 2015). Phase 1 of this project assessed the integrity of interventions claiming to treat the primary dimensions of autism spectrum disorder in children, adolescents and adults below the age of 22 years. Phase 2 focused on extending the findings related to young adults, while simultaneously evaluating interventions designed for adults over the age of 22 years. According to the National Standards Report “just because an article has been published does not mean that the outcomes are critically important” (National Autism Center, 2009, p. 16). As a result, the “Scientific Merit Rating Scale” (SMRS) was developed to objectively evaluate the strength of a study’s experimental design and determine if there was enough experimental rigor demonstrated to adequately assess intervention effects (National Autism Center, 2009; 2015).

According to the SMRS, a study is rated across the following dimensions, using a six-point scale: (1) research design; (2) measurement of dependent variable; (3) procedural fidelity; (4) participant ascertainment, and (5) generalisation and maintenance of effects. A composite score, known as the SMRS score, is then
calculated. Scores of 3, 4 or 5 indicate that sufficient methodological rigor has been demonstrated, which allows definite conclusion to be drawn from a study’s results (National Autism Center, 2009; 2015). However, the results from one study alone is insufficient in determining the evidence-based status of an intervention (Chambless & Hollon, 1998). As a result, the “Strength of Evidence Classification System” was designed to evaluate the “quality, quantity and consistency of research findings” (National Autism Center, 2015, p. 34). According to this system, an intervention can be considered “established” if the following criteria are met: (1) there must be at least two group or four single-subject experimental design studies (published and peer-reviewed) with at least 12 participants and no conflicting results, or three group or six single-subject experimental design studies (published and peer-reviewed) with at least 18 participants and less than 10% of studies reporting conflicting results; (2) all studies must score 3, 4 or 5 on the SMRS, and (3) the intervention must produce beneficial effects (National Autism Center, 2015).

Employing the SMRS and the Strength of Evidence Classification System, the National Standards Report found that the overwhelming majority of interventions that were identified as “established” for children with autism, were generated from applied behaviour analytic research (National Autism Center, 2009; 2015). Furthermore, there was substantial overlap with findings from the National Professional Development Centre project (Wong et al., 2014). With regards to the adult literature, only 27 studies met the inclusion criteria for review by the National Standards Project. An evaluation of these studies showed behavioural intervention to be the only “well-established” treatment for adults with autism (over the age of 22 years). These behavioural interventions include, but are not limited to, prompting, extinction, differential reinforcement, and functional communication training.
These large-scale quantitative reviews of the literature provide a valuable resource to help professionals and service providers make informed decisions when selecting interventions for their service users with autism (National Autism Center, 2009; 2015; Wong et al., 2014). To-date, however, there has been no similar attempt to quantify and evaluate the research addressing interventions for children and adults with intellectual disabilities.

**Conclusion**

The quality of life of people with intellectual disabilities is significantly impacted by their levels of adaptive functioning and whether or not they engage in challenging behaviour. Individuals with impaired adaptive behaviour are at an increased risk for resorting to challenging behaviour (Emerson, 2001; Kiernan et al., 1997). In turn, this negatively impacts quality of life outcomes, including independence and social participation (Schalock et al., 1994). However, with suitable supports, adaptive functioning can improve, resulting in a reduced reliance on challenging behaviour (Grey & Hastings, 2005; Kiernan et al., 1997; Kurtz et al., 2011) and improvements in quality of life (Dollar, Fredrick, Alberto, & Luke, 2012).

Evidence-based practices are highly effective in facilitating the timely attainment of personal goal, increased independence and improved quality of life outcomes for people with intellectual disabilities (Brown, Schalock, & Brown, 2009; Maes et al., 2007; National Institute for Health and Care Excellence, 2015). However, the National Standards Report (2009) cautions against uncritical acceptance of published results and findings (National Autism Center, 2009). In order to be confident of a study’s conclusions, the methodology must be subject to rigorous evaluation of its scientific merit. Furthermore, the results of a single study...
are insufficient to establish a practice as evidence-based; there must be a critical synthesis and comparison of results across studies that all display the same high standard of experimental control (Chambless & Hollon, 1998).

Applied Behaviour Analysis, the science of behaviour, provides a scientific framework for the systematic development and evaluation of evidence-based practice, as well as the delivery of these practices within the intellectual disability sector. Two large scale projects undertaken in the United States of America, sought to promote the use of evidence-based practices for people with autism spectrum disorders (National Autism Center, 2009; 2015; Wong et al., 2014). To achieve this goal, both studies set out to quantitatively analyse and evaluate the interventions currently employed with this population. Across these two independent studies, the majority of interventions that met criteria for being empirically “well-established” practices originated within the field of Applied Behaviour Analysis (Wong et al., 2014). The most striking limitation to these findings, however, is their specificity to the autism population.
Chapter 4

Dissemination of Evidence-Based Practice and Behaviour Skills Training
Over recent years, the importance of embedding evidence-based practice within the service provision framework for people with intellectual disabilities has been repeatedly emphasised (Health Service Executive, 2012a; National Autism Center, 2009; 2015; National Institute for Health and Care Excellence, 2015; Wong et al., 2014). However, Lowe et al. (2007) argue that the delivery of intervention based on evidence-based practice within services should not be exclusively within the remit of specialist professional teams (e.g., behaviour specialists or those who have formal qualifications in psychology and related areas). It has been argued that frontline staff may be upskilled and supported to deliver evidence-based practice and can supplement specialist professional teams to ensure consistent application of these practices within service providers (Devereux, Hastings, Noone, et al., 2009; Lowe et al., 2007).

**Factors Influencing the Use of Evidence-Based Practice in an Intellectual Disability Setting**

It has been argued that there is a significant disconnect between the emerging knowledge pertaining to evidence-based practice for people with intellectual disabilities and the competencies and skills displayed by staff in applied settings (Campbell, 2010; Hile & Walbran, 1991; Rapp et al., 2010; Swain, Whitley, McHugo, & Drake, 2010). Campbell (2010) refers to this inconsistency in dissemination as a “theory-practice gap”. This theory-practice gap can arise as a result of a number of factors, including but not limited to insufficient basic knowledge (Hastings, 1996), inadequate training (Campbell, 2007; Jahr, 1998) and the perceived social validity of the treatments and interventions (Callahan et al., 2016).
From an Irish perspective, it is estimated that 71% (19,975) of persons registered with the National Intellectual Disability Database (NIDD) will require new or improved specialised supports during the 2017-2021 period (Doyle et al., 2017). However, restricted access to skilled professionals, including speech and language therapists, psychologists and psychiatrists (Fallon et al., 2017; McClean et al., 2005; Yacoub, 2015), may result in insufficient support and training for frontline staff and a continued reliance on the “caring” approach rather than active support of service users (Linehan et al., 2014; Tatlow-Golden et al., 2014). Furthermore, with the Irish government yet to ratify the United Nations Convention on the Rights of Persons with Disabilities (Tatlow-Golden et al., 2014), there is no formal obligation on service provider organisations to ensure that frontline staff are adequately qualified and trained in the delivery of relevant evidence-based interventions (National Economic and Social Council, 2012).

**Qualifications of Frontline Staff**

In 2002, the Joint Committee on Social Care Professionals in the Republic of Ireland published a report outlining the qualifications of staff employed in residential and community-based childcare settings, as well as residential settings for people with intellectual disabilities (Joint Committee on Social Care Professionals, 2002). Over half of the 2,791 participants sampled were employed by intellectual disability providers and 36% of these individuals did not possess a qualification in social care. Extant research demonstrates that compromised skill sets among intellectual disability staff can have significant adverse repercussions for service users (e.g., Finn & Sturme y, 2009; Jahr, 1998; Schepis et al., 2001), therefore the lack of professional qualifications in this sector in the Republic of Ireland is concerning.
Similarly, Power and Lavelle (2011) examined the qualifications of 3,878 frontline staff employed in private and public adult residential settings in the Republic of Ireland. While the original sample included residential settings for people with a mild intellectual disability, data from these service providers was excluded from analyses because of the exceptionally low response rate. Results revealed that 55% of staff who were employed by residential services that supported older adults (e.g., nursing homes and hospitals) did not hold a professional qualification nor were they pursuing such a qualification within the Irish National Framework of Qualifications. Of those staff who did possess a qualification under this framework, 90% had the “baseline qualification” (Level 5) and 91% of the staff in training were pursuing this basic qualification. Although these figures cannot be readily generalised to intellectual disability providers, they do indicate that the practice of hiring personnel without suitable professional qualifications, continues to be problematic within the social care sector in the Republic of Ireland.

**Access to Effective Training and Support**

Even though a sizeable portion of the frontline staff, who are supporting people with intellectual disabilities, may not be considered adequately qualified, research has shown that staff with varying backgrounds and educational qualifications can be effectively trained to carry out procedures in line with evidence-based practice (Allen & Tynan, 2000; Dench, 2005; Grey & Hastings, 2005). For instance, McClean et al. (2005) successfully designed a competency-based training programme to teach intellectual disability staff how to implement evidence-based behavioural assessments and interventions. One hundred and thirty two staff participated in this training programme; the majority (92%) were employed as direct caregivers and had diploma or degree level qualifications in nursing,
psychology, social studies or childcare. In total, 138 service users were involved in the study. All engaged in challenging behaviour (i.e., aggression or self-injury), 76% were adults and 51% were diagnosed with a severe or profound intellectual disability. The impact of staff training was significant, with 77% of cases achieving criteria for success (at least a 70% reduction in challenging behaviour from baseline levels). Longitudinal data show that this reduction maintained over a 2-year period.

The results from McClean et al. (2005) demonstrate that a well-designed, comprehensive, in-house training programme could overcome the challenges posed by hiring staff without what may be considered adequate qualifications. However, organisations and service providers are often hesitant to absorb the outlay associated with training, due to high staff turnover (Addis, 2002) and the potential for compromised service provision in the short-term (e.g., service closures during training delivery; Foa, Gillihan, & Bryant, 2013). This reluctance to invest in staff training may be compounded by the didactic instructional nature of standard continuous educational courses (van Oorsouw, Embregts, Bosman, & Jahoda, 2009), which has been shown to be largely ineffective in producing enduring change to practitioners’ work (Addis, 2002).

In addition, impactful and enduring change within an organisation cannot be achieved by targeting frontline staff alone (Cullen, 2000; Rapp et al., 2010; Swain et al., 2010). Ongoing training is required for professionals at all levels, particularly for those in managerial and supervisory positions (DiGennaro Reed & Henley, 2015). It is these professionals who are typically responsible for guiding teaching procedures, treatments and interventions, as well as monitoring and supporting the skills and behaviours of staff. Therefore, without leaders competent in the implementation of evidence-based practice and active collaboration between managers, programme
supervisors and frontline staff, the sustainability of such practices may be considered unviable (McCabe et al., 2015; Rapp et al., 2010).

For example, Swain and her colleagues (Swain et al., 2010) investigated the sustainability of practices promoted by the “National Implementing Evidence Based Practices Project” within 49 organisations in America. The results of this study revealed that organisations, which maintained the delivery of evidence-based practice in mental health services reported consistently higher mean scores on a number of measures. However, the largest differences between “sustainers” (continued use of evidence-based practice over a 2-year period) and “non-sustainers” (failure to implement evidence-based practices 2 years after initial implementation) were evident on staff training measures, followed closely by scores on staff supervision measures.

This finding is further supported by Rapp and his colleagues (Rapp, Etzel-Wise, Marty, Coffman, Carlson, Asher et al., 2008), who revealed that the impact of high quality supervision could often outweigh specific methods of staff training. The authors argue that programme leaders play a critical and unique role in the context of evidence-based service provision and have a responsibility to guide, develop, supervise and enforce evidence-based practice on a daily basis. Therefore, it is clear that high standard, empirically validated training and supervision for frontline staff and programme leaders is vital not only to ensure the fidelity and consistency of evidence-based interventions but also to ensure the best possible outcomes for the service users they support.
Social Validity of Evidence-Based Practice

Over recent years there has been a growing interest in the factors that influence the success of training for staff working directly with people with intellectual disabilities (e.g., Allen & Tynan, 2000; Brookman-Frazee, Drahota, & Stadnick, 2012; Browder, Trela, & Jimenez, 2007; Felce et al., 2000; Luiselli & St. Amand, 2005; Marks, Sisirak, & Chang, 2013; Roscoe, Fisher, Glover, & Volkert, 2006) and research is showing that social validity is positively correlated with effective use in applied settings (Callahan, Henson, & Cowan, 2008). Simply identifying evidence-based strategies is insufficient for successful dissemination, as many professionals and service providers appear to be unaffected by the developments around effective practice (Callahan et al., 2016). Instead, they continue with the implementation of unestablished techniques, which can adversely impact the progress of their vulnerable service users (Callahan et al., 2008).

According to Carter (2010) “just because a program is considered effective does not mean that it will be considered appropriate by those closely involved in the implementation of the program” (p. 7).

As previously described in Chapter 3, evidence-based practice is typically identified through evaluations of experimental rigor. However, there is generally little consideration given to the social validity of these treatments (i.e., consumer satisfaction with intended procedures and outcomes; Callahan et al., 2016). As a result, Callahan et al. (2016) investigated which of the evidence-based interventions (i.e., “established” and “emerging”), that were identified by the National Standards Project and the National Professional Development Centre report (National Autism Center, 2009; 2015; Wong et al., 2014), displayed social validity. According to the results, all of the evidence-based practices identified demonstrated at least a minimal
level of social validation, thereby supplementing the existing empirical support for their effectiveness. Although this research focused primarily on the social validity of these treatments for autism populations, the findings may also apply to other groups presenting with intellectual and developmental disabilities.

**Disseminating Evidence-Based Practice**

Although an important consideration, social validity is just one factor influencing effective dissemination of evidence-based practice to frontline staff. In order to synthesise the growing literature, van Oorsouw et al. (2009) conducted a meta-analysis to identify and examine effective structural components of staff training. The authors concluded that optimal results are achieved through the combined use of in-service training and on-the-job coaching. For in-service training, the authors recommended delivering target information to staff in a number of different ways (e.g., didactic instruction, modelling, and role-play) as findings indicated that staff performance was enhanced under these type of training conditions. On the other hand, on-the-job coaching benefited from a singular approach, in the form of verbal feedback.

The strategies highlighted as effective in van Oorsouw et al. (2009) are consistent with the four components of Behaviour Skills Training. Behaviour Skills Training is a type of training package, which employs a combination of instruction, modelling, rehearsal and feedback (Ward-Horner & Sturmey, 2012) and can be used to teach anyone a new skill (Miltenberger, 2016). This training approach originated within the discipline of Applied Behaviour Analysis and developed from a motivation within the field to successfully disseminate effective practices (Baer et al., 1968; 1987) to staff working in applied settings.
Chapter 4

**Behaviour Skills Training**

One of the first references to a training procedure involving the components of Behaviour Skills Training involved the work of Yeaton and Bailey (1978). The authors successfully employed a multi-component training package to teach 24 children six pedestrian safety skills. The package involved the following steps: Tell them; Show them; Ask them; Let them. Data from generalisation tests and a 1-year maintenance check supported the viability of this training approach. A follow-up study (Yeaton & Bailey, 1983), which involved a component analysis of this package found that when the steps were combined they were more effective than when used independently. For example, there was an inconsistent impact on the street-crossing behaviour of the children when instruction and modelling were used alone but performance greatly improved with the opportunity for practice with feedback. These findings are consistent with those of Ward-Horner and Sturmey (2012), which also supported the critical role of feedback within a Behaviour Skills Training package.

In 2000, Iwata and colleagues used all four components of Behaviour Skills Training to teach a complex skill, the administration of a functional analysis, to 11 undergraduate students. These trainees received written instruction, watched videotaped simulations and received feedback on their performance during rehearsal sessions. Results showed that the trainees reached an accuracy level of at least 95% correct implementation following training of approximately two hours (Iwata et al., 2000). Similarly, Lavie and Sturmey (2002) successfully trained three staff to carry out a paired-stimulus preference assessment with eight children with autism. Although it was not yet described as Behaviour Skills Training, this training package employed all of the relevant components.
While the individual components of Behaviour Skills Training have been used in isolation and in combination since the 1970’s, Behaviour Skills Training in its current form was formally identified by Sarokoff and Sturmey, in 2004. In this study, the authors set out to train three special education teachers to implement discrete-trial teaching correctly. Participants were instructed on how to effectively teach a 3-year old child with autism a matching task. Participant implementation of discrete-trial teaching was assessed across 10 areas, including waiting for the child to show readiness behaviour before delivering an instruction, making eye contact and using clear articulation when issuing an instruction, implementing the designated error correction procedure when necessary, providing reinforcement as specified, using behaviour specific praise throughout the session and maintaining a 5-second inter-trial interval between tasks.

Prior to receiving Behaviour Skills Training, baseline measurements of each participant’s performance across the 10 target areas were conducted. Immediately following baseline sessions, individual training commenced. Participants were given a written copy of the discrete-trial teaching procedures and reviewed the 10 areas with the experimenter. They were also provided with a graph of their performance during the baseline sessions and associated feedback. Modelling sessions involved the experimenter demonstrating the correct implementation of discrete-trial teaching with the 3-year old child. During rehearsal sessions each participant was observed, by the experimenter, delivering discrete-trial instruction to the 3-year old child. Following three trials participants were given positive and corrective feedback by the experimenter. Modelling and rehearsal sessions were repeated until 10 minutes had elapsed. Participants were then required to deliver 10 uninterrupted discrete-trials,
while their performance was assessed. Criteria for training completion was 90% correct implementation across three consecutive training sessions.

During baseline, procedural fidelity scores averaged 43%, 49% and 43% for participants 1, 2 and 3, respectively. However, post-training, participants 1, 2 and 3 achieved mean scores of 97%, 98% and 99%, respectively. These results demonstrated that a training package involving instruction, modelling, rehearsal, and feedback could produce substantial, rapid, positive changes in the teachers’ ability to implement the target skill appropriately.

**Training Professionals Using Behaviour Skills Training**

Since 2004, Behaviour Skills Training has been effectively implemented to disseminate skills to professionals working with children and adults with intellectual and developmental disabilities. For example, Nabeyama and Sturmey (2010) employed Behaviour Skills Training and self-recording to increase the correct posture and guarding responses of three staff working with students with multiple physical disabilities. Participants were employed as teaching aides in special education classrooms. Six posture and guarding responses, specific to each student, were identified for staff training. During baseline sessions, each participant was matched with a student and the six responses for correct posture and guarding were outlined. Participants were then expected to assist the student in ambulating between two locations, while their implementation of these responses was assessed.

During Behaviour Skills Training, participants were given a self-recording checklist, the six correct posture and guarding responses were described and explained, and questions were facilitated by the experimenter. Participants then assisted student ambulation, self-assessed their own performance, received positive
and corrective feedback, and compared their own performance checklist to that of the experimenter. Modelling and rehearsal sessions were implemented for participants failing to achieve six correct response components. During modelling, the experimenter demonstrated the correct responses with a student, while describing what the participant had done incorrectly. Participants were then required to rehearse the target responses with their designated student. The experimenter provided vocal and physical prompts as necessary and delivered positive and corrective feedback immediately after the trial. Criteria for training completion was the correct implementation of the six target posture and guarding responses across two consecutive trials. Findings demonstrated an improvement in staff behaviour, which generalised to new students. Furthermore, the distance that the students ambulated increased as a result of the change in the approach of staff (Nabeyama & Sturmey, 2010).

Behaviour Skills Training has also facilitated the effective dissemination of functional behaviour assessment techniques to frontline staff working with both children and adults with disabilities. For example, Lambert, Bloom, Kunnavatana, Collins, and Clay (2013) produced a partial replication of the study reported by Iwata et al. (2000) by successfully training six supervisors of residential service providers to teach nine house managers how to implement trial-based functional analyses with adult service users. Training involved verbal and written descriptions of each functional analysis condition, followed by modelling and rehearsal sessions with service user confederates. Rehearsal sessions included a feedback component. Similarly, Loman and Horner (2014) used all four components of Behaviour Skills Training to teach school professionals to conduct a functional behaviour assessment to assess the function of problem behaviour displayed by the children they supported.
Post-training results showed that all trainees could correctly identify the antecedents and consequences controlling their students' inappropriate behaviour.

The success of Behaviour Skills Training is also evident within research focusing on augmentative and alternative communication systems, such as the Picture Exchange Communication System (PECS; Homlitas, Rosales, & Candel, 2014; Wood, Luiselli, & Harchik, 2007) and manual signing (Parsons, Rollyson, & Reid, 2012). Homlitas et al. (2014) tested the effectiveness of Behaviour Skills Training to teach accurate use of PECS to staff working at a therapeutic centre for children with autism. All three teachers demonstrated the target skills in the natural environment and with multiple children from their classroom, without additional training. Likewise, Wood et al. (2007) used Behaviour Skills Training to teach four trainees to implement PECS with an adult service user. Results showed that three of the four trainees were able to maintain close to 100% instructional accuracy following the intervention. The importance of including all components of Behaviour Skills Training is emphasised by Barnes, Dunning, and Rehfeldt (2011) who found that training consisting of verbal instruction and an instructional video had little impact on staff’s ability to implement PECS. The authors argue that all four components of Behaviour Skills Training are necessary to facilitate the learning of a complex skill, such as PECS.

In line with these recommendations, Gianoumis, Seiverling, and Sturmey (2012) employed all elements of Behaviour Skills Training to successfully train staff in the appropriate application of the Natural Language Paradigm, a naturalistic teaching method aimed at increasing vocalisations in children with autism. The teachers’ ability to implement the Natural Language Paradigm increased significantly and the skill generalised to use with other children. The authors reported that all
teachers achieved mastery criterion in a relatively short period of time (20-30 minutes), which is consistent with other studies employing Behaviour Skills Training (Lavie & Sturmey, 2002; Sarokoff & Sturmey, 2004; Ward-Horner & Sturmey, 2008). Furthermore, the improved teacher performance resulted in a greater frequency of vocalisations and reduced challenging behaviour for the majority of children.

Behaviour Skills Training has also been repeatedly used to effectively teach the correct implementation of discrete trial teaching (Bolton & Mayer, 2008; Dib & Sturmey, 2007; Gilligan, Luiselli, & Pace, 2007; Nosik & Williams, 2011; Sarokoff & Sturmey, 2004; 2008). Discrete trial teaching is a method of individualising instruction to suit the learner’s needs (Sarokoff & Sturmey, 2004) and has shown to be effective when used with people with developmental disabilities (Green, 1996). As previously discussed in this chapter, Sarokoff and Sturmey (2004) used all four components of Behaviour Skills Training to coach staff in discrete trial teaching. Results demonstrated rapid and substantial improvement in procedural fidelity, with levels of correct implementation rising from an average of 45% in baseline to 98% post-intervention.

These finding were replicated and extended by Sarokoff and Sturmey (2008). Staff demonstrated significant gains following Behaviour Skills Training in the implementation of discrete trial teaching. In addition, these skills maintained over time and generalised to use with novel students and programmes. In a comparison study, Nosik, Williams, Garrido, and Lee (2013) demonstrated that participants who had engaged in Behaviour Skills Training consistently outperformed their colleagues (who had received computer-based instruction) on procedural fidelity measures of discrete trial implementation with adult service users. This effect remained evident
at a 6-week follow-up assessment. Most recently, Maffei-Almodovar, Feliciano, Fienup, and Sturmey (2017) successfully trained three community-based teachers to analyse graphs for the purpose of monitoring service user progress on discrete trial-based programmes. The three teachers were taught how to implement data decision rules to identify when instructional changes should be made to their assigned service users’ discrete trial tasks. Results showed an increase in correct data-based decisions and a decrease in error for all participants following Behaviour Skills Training.

Finally, in order to consolidate a portion of the Behaviour Skills Training literature, Fetherston and Sturmey (2014) provided a systematic replication of previous studies which had demonstrated the efficacy of Behaviour Skills Training across the implementation of discrete trial teaching (Dib & Sturmey, 2007; Lafasakis & Sturmey, 2007), incidental teaching (Ryan, Hemmes, Sturmey, Jacobs, & Grommet, 2008) and teaching activity schedules (Ryan et al., 2008; Sarokoff & Sturmey, 2004). There was substantial evidence of skill acquisition and positive behaviour change for staff trainees and their assigned service users, across all three experiments. In addition, staff generalised the new practices to novel situations and rated the training procedure as highly acceptable (Fetherston & Sturmey, 2014).

Based on this body of research, Behaviour Skills Training may be considered a robust, efficacious method for teaching new skills. It has been successfully applied to train a wide range of groups, including people with intellectual disabilities, undergraduate students, frontline staff and other professionals and there is repeated support for the generalisation of the newly learned skills to novel environments, without additional training (e.g., Homlitas et al., 2014; Maggin, Fallon, Sanetti, & Ruberto, 2012; Sarokoff & Sturmey, 2008). While the rehearsal and feedback components have been shown to be critically important (van Oorsouw et al., 2009;
Ward-Horner & Sturmey, 2012), it also seems that training results are optimised when all four components of Behaviour Skills Training are used in combination (Barnes et al., 2011; Homlitas et al., 2014). In conclusion, this training approach presents as a viable, effective package to teach a wide range of skills to different cohorts across diverse settings.

**Conclusions and Research Aims**

Frontline staff shape the experiences and behaviour of the service users who they work with, through the opportunities, feedback, and consequences they deliver (Finn & Sturmey, 2009; Jahr, 1998; Mansell & Beadle-Brown, 2012). However, the intellectual disability sector continues to hire what may be considered “professionally unqualified” individuals as their frontline staff (Department of Health, 2001; Joint Committee on Social Care Professionals, 2002). Therefore, in-house training is one of the most accessible methods for educating staff in the strategies and approaches that can optimise service user outcomes.

Studies have demonstrated that training in evidence-based practice can enhance service user outcomes by improving the interactions between staff and service users (Emerson et al., 2000) and increasing the learning opportunities available (Rose et al., 1998). Furthermore, effective training programmes are associated with lower levels of staff stress (Graber et al., 2008) and higher levels of job satisfaction (Zwijsen et al., 2015). Although substantial progress has been made in identifying evidence-based practice for people with intellectual disabilities, there is still scope for additional investigation. Furthermore, the factors influencing the successful dissemination of these practices to frontline staff warrant further exploration.
These issues are particularly relevant within the Irish context, considering the increasing number of individuals availing of intellectual disability services (Doyle et al., 2017). As discussed in Chapter 1, the intellectual disability sector in the Republic of Ireland is currently in transition, thereby providing an opportunity for innovation and positive change. As a result, the proposed research aims to examine the training needs of frontline staff employed by one of the largest intellectual disability service providers in the Republic of Ireland and subsequently develop an effective and socially valid method for disseminating a package of evidence-based practices, in order to prepare these frontline staff to support service users with intellectual disabilities. The objectives of the present programme of research are as follows: (1) to conduct a systematic literature review of the training provided to staff working in the intellectual disability sector; (2) to establish a training needs analysis among a representative sample of frontline staff and their service users within the host organisation. This analysis seeks to identify areas of need for service users, by acquiring information on current levels of adaptive behaviour skills, quality of life and management of challenging behaviour through staff reports of training and work-related experiences; (3) to design and employ a Behaviour Skills Training intervention, matched to the training needs of the host organisation and founded in evidence-based practice, and (4) evaluate the impact of this intervention on a range of staff- and service user-based outcomes.

Finally, to address the issue of social validity, the proposed research aims to use the following indicators to guide the experimental evaluation of the training package: (1) were socially important dependent variables selected?; (2) were consumers satisfied with the results?; (3) was it a time and cost effective training package?; (4) was the impact of the training package assessed in the natural context?
(5) did the training package impact the behaviour of frontline staff, and (6) were clinically significant levels of behaviour change achieved (Callahan et al., 2016)?
Chapter 5

Staff Training in Intellectual and Developmental Disability Settings: A
Systematic Review
Abstract

Frontline staff are a valuable asset within an intellectual disability service. Their work dictates the overall standard of care delivered by the organisation. This review examines staff training in practices to support people with intellectual and developmental disabilities and treat skill deficits and challenging behaviour. Systematic searches of relevant databases identified 107 papers, for inclusion in the review. Practices in which staff were trained, were categorised as: a) Positive Behaviour Support (PBS) interventions; b) Applied Behaviour Analysis (ABA) interventions, or c) other interventions. The Applied Behaviour Analysis category was sub-divided into: a) assessment-based practices; b) antecedent-based practices; c) consequence-based practices and, d) mixed interventions. Results showed that although staff were trained in a range of evidence-based practices, many empirically supported interventions were not utilised. Despite rigorous scientific support for strategies such as functional communication training and non-contingent reinforcement, the literature did not demonstrate effective protocols to disseminate these practices to frontline staff. The review also highlighted the scarcity of studies that compared the effectiveness of different training protocols. These findings provide a potential explanation for the apparent disconnect between theoretical advancements and practice in the applied setting and are discussed in relation to approaches to staff training.
Research has shown that training protocols that may be characterised as inadequate and of low quality, for frontline staff supporting people with intellectual disabilities, have been associated with a number of service user outcomes. These include impaired quality of life (Jahr, 1998), reduced learning opportunities (Schepis et al., 2001), and compromised relationships between staff and service users (Finn & Sturmey, 2009). Intervention settings for individuals with intellectual disabilities have been described as stressful environments (Hatton et al., 1999; Hensel et al., 2012; Mitchell & Hastings, 2001; Robertson et al., 2005; Rose & Rose, 2005) and staff employed within these settings are susceptible to illness, absenteeism, and employee turn-over (Rose, 1995). In addition, staff working with this population may not be equipped to demonstrate appropriate handling of challenging situations with service users (Hastings, 2002) and as a result these individuals may be at an increased risk for inappropriate practices within organisations (White et al., 2003).

International guidelines on best practice for people with disabilities state that all persons, regardless of disability, have the right to the highest quality of physical and mental health care that is currently available (National Disability Authority, 2003). This includes access to evidence-based treatments and interventions when necessary (National Institute for Health and Care Excellence, 2015). Research has consistently shown that high quality, empirically supported programmes are effective in facilitating the timely attainment of personal goals, increased independence and improved quality of life for service users (Brown et al., 2009; Maes et al., 2007). Furthermore, training and supervision in evidence-based practice is also associated with lower levels of staff stress and burnout (Graber et al., 2008) and higher levels of job satisfaction (Zwijsen et al., 2015).
Applied Behaviour Analysis is a science that systematically employs strategies derived from the scientific principles of behaviour to produce socially significant, meaningful changes in the adaptive and maladaptive behaviour of individuals (Cooper et al., 2014). At present, function-based behavioural interventions are one of the principle, empirically supported approaches used in the treatment of challenging behaviour (Carr & Durand, 1985; Carr, Horner, et al., 1999; Ellingson, Miltenberger, Stricker, Galensky, & Garlinghouse, 2000; Filter & Horner, 2009; Grey & Hastings, 2005; Ingram, Lewis-Palmer, & Sugai, 2005; McClean et al., 2005; Newcomer & Lewis, 2004; Payne, Scott, & Conroy, 2007). Functional communication training, for example, is a well-established intervention for challenging behaviour displayed by people with intellectual disabilities (Kurtz et al., 2011). This treatment involves replacing challenging behaviour with an alternative functionally equivalent communication response and it has been shown to be more effective in reducing challenging behaviour observed in individuals with intellectual disabilities than nonfunction-based communication training (Carr & Durand, 1985; Filter & Horner, 2009).

Likewise, non-contingent reinforcement has been repeatedly employed to effectively treat a range of inappropriate response topographies and reinforcement functions exhibited by this population (Carr et al., 2009). Ellingson et al. (2000) demonstrated that an intervention package, which included frequent non-contingent delivery of attention, effectively reduced the challenging behaviour of a child with an intellectual disability. This was in contrast to an alternative treatment, which did not specifically address the function of the challenging behaviour.

Positive Behaviour Support, which evolved from Applied Behaviour Analysis (Johnston, Foxx, Jacobson, Green, & Mulick, 2006), is also a widely cited approach
for treating challenging behaviour and improving quality of life outcomes for people with intellectual disabilities (National Institute for Health and Care Excellence, 2015; Royal College of Psychiatrists, 2007). The technology of Positive Behaviour Support employs system change methods, based on the fundamental principles of behaviour analysis, to produce robust changes in service user quality of life and instances of challenging behaviour (Carr et al., 2002; Wehmeyer & Schalock, 2001). McClean and colleagues (McClean et al., 2005) successfully trained frontline intellectual disability staff to design and implement positive behavioural support interventions for their service users. Longitudinal data showed a significant reduction in the frequency of challenging behaviour for 77% of service users (n=138), a result that maintained over a 22.5 month period. In addition, anecdotal reports revealed improvements across a number of quality of life variables. These included increased choice, communication, and opportunities for service users to engage in preferred activities.

To support practitioners and consumers in making informed treatment choices, research has begun to focus on the identification and operationalisation of evidence-based practice (Grey & Hastings, 2005; Grow, Carr, & LeBlanc, 2009; National Autism Center, 2009; Rapp et al., 2010; Sturmey & Didden, 2014). While progress is evident (e.g. National Autism Center, 2009; Sturmey & Didden, 2014), substantial disconnect remains between these theoretical advancements and the competencies of frontline staff in applied settings (Hile & Walbran, 1991; Rapp et al., 2010; Swain et al., 2010).

Intellectual disability frequently co-occurs with other developmental disabilities (American Association on Intellectual and Developmental Disabilities, 2017; Matson & Shoemaker, 2009). For instance, 50-70% of individuals with autism
are also diagnosed with an intellectual disability (Matson & Shoemaker, 2009). As a result, staff working in the intellectual disability sector often work with people diagnosed with both intellectual and developmental disabilities (American Association on Intellectual and Developmental Disabilities, 2017). Therefore, to investigate potential explanations for the inadequate dissemination of evidence-based practice in this sector, published literature, which focused on both intellectual and developmental disabilities, was reviewed and synthesised in order to answer the following questions: a) what practices are staff, who work with people with intellectual and developmental disabilities, being trained in to support them in treating skill deficits and challenging behaviour?; b) what method is used to deliver staff training?; and c) what is the impact of current staff training practices on service user, staff and organisational outcomes?

**Method**

**Search Procedures**

Comprehensive literature searches were carried out across the following electronic databases: EBSCO Academic Search Complete, PsycINFO, PsycArticles, ERIC, Web of Science, and PubMed. Systematic searches were conducted by combining the search terms *staff* and *training* with the following keywords: *evidence based practice, applied behavior* analysis, *positive behavior* support, *disability, disorder, autism, developmental disability, intellectual disability, and mental retardation*. In addition, a review of the reference lists of all included articles was conducted to identify other studies meeting the inclusion criteria.

**Inclusion and exclusion criteria.** Searches were limited by year of publication (2000-2015), with studies prior to 2000 being excluded from the review.
In addition, only those papers, which were written in the English language, involved original research and were published in peer-reviewed journals, were considered for inclusion. Studies, included in the current review met the following criteria: a) a specified training programme was provided to staff during the study; b) staff worked with individuals with intellectual and developmental disabilities; c) staff were trained in practices used to treat skill deficits and/or challenging behaviour, and d) the impact of staff training was evaluated for at least one of the following groups: staff; service users, and/or the wider organisation. Articles which described the provision of training to a student or parent/caregiver population only were excluded from the review. In addition, studies that focused exclusively on service users with a physical developmental disability only were excluded. A total of 107 studies met the criteria for inclusion following a filtering process outlined in Figure 1.
Figure 1. Flow diagram showing inclusion/exclusion of studies identified during database search process.

**Study Classification**

Studies were categorised according to the practices in which staff were trained. Practices were classified as either: (1) Positive Behaviour Support (PBS), which comprised all studies that explicitly mentioned Positive Behaviour Support in relation to the interventions being trained; (2) Applied Behaviour Analysis (ABA), which covered all studies with training content that utilised principles and/or interventional strategies derived from the science of behaviour, or (3) “other”, which
comprised treatments or intervention strategies drawn from disciplines other than Positive Behaviour Support or Applied Behaviour Analysis.

The Applied Behaviour Analysis category was further divided into four sub-categories including studies which provided training in: (1) behavioural assessment practices only; (2) antecedent-based practices only; (3) consequence-based practices only, or (4) mixed interventions. Mixed interventions comprised training packages that included a combination of techniques from the other three sub-categories (i.e., assessment, antecedent, and consequence-based practices).

**Data Extraction**

For each included study, data were extracted on the following variables: (1) training content and method of delivery; (2) method(s) used to assess staff learning as a result of training (e.g., knowledge acquisition specific to training content, ability to implement target practices with fidelity); (3) impact of training on staff learning (e.g., knowledge acquisition specific to training content, ability to implement target practices with fidelity); (4) evaluation of job performance; (5) service user population supported by staff receiving training (adults (>18 years), children (<18 years), combination); (6) service user outcomes assessed (skill acquisition, challenging behaviour) and the corresponding findings, and (7) impact of training for the organisation.

**Interrater Agreement**

The accuracy of the data extracted was evaluated by having a second rater review 55 (51.4%) of the included studies. Inter-rater reliability was found to be 99.2% (range 77.78% - 100%). In cases where discrepancies were evident, discussion between raters was used to achieve consensus.
Results

A total of 107 articles published between the years 2000 – 2015 met the inclusion criteria (Appendix A). The variables extracted across each category of practices were summarised and the raw data are available from the author on request.

Positive Behaviour Support (PBS) Practices

Sixteen studies met the criteria for inclusion in this category (Appendix A).

**Characteristics of staff training.** Eight of the studies in this category trained staff in Positive Behaviour Support principles and strategies. Two studies (Bradshaw, Mitchell, & Leaf, 2010; Bradshaw, Reinke, Brown, Bevans, & Leaf, 2008) examined school-wide Positive Behaviour Support and two (Allen & Tynan, 2000; Tierney, Quinlan, & Hastings, 2007) focused on Positive Behaviour Support strategies to address challenging behaviour, specifically. The remaining four studies (Gore & Umizawa, 2011; LaVigna, Christian, & Willis, 2005; Salmento & Bambara, 2000; Singh et al., 2015) provided training in one of the following areas: a) Positive Behaviour Support and functional communication training (FCT) principles; b) mindfulness-based Positive Behaviour Support; c) developing a positive behaviour support plan and, d) providing choice opportunities to service users using Positive Behaviour Support principles.

Four of the 16 studies, employed a train-the-trainer model in their protocols, while three used a Person Focused Training approach (i.e., developing a behaviour assessment and support plan for a specific service user). One study (Reid et al., 2003) used Behaviour Skills Training (i.e., instruction, modelling, rehearsal and feedback). Seven of the eight remaining studies employed individualised training packages that included instruction and practical application. One study (Singh et al.,
which focused on mindfulness-based Positive Behaviour Support, used classroom instruction alone.

**Method of assessing staff learning.** Seven studies in this category evaluated the impact of training on staff skill acquisition, six examined changes to staff knowledge, with three of those focusing on challenging behaviour attributions. Finally, three of the studies did not assess staff learning in any format.

**Staff learning.** A total of 13 studies in this category examined the impact of training on staff knowledge or skill acquisition and all but one study (Tierney et al., 2007) reported positive results post-training. However, one of the 13 studies (Lowe et al., 2007), showed that although there was a significant change to challenging behaviour attributions immediately following training, this was not maintained at a 1-year follow-up.

**On-the-job performance assessment.** For the seven studies, which examined staff skill acquisition, all seven evaluated the ability of staff to demonstrate their newly acquired skills in their typical work environment.

**Service user population.** The service users supported by staff in five of the studies were children, three studies examined staff working with adults and three looked at those assisting both adults and children. Five studies did not specify whether the service users were children, adults or a combination of both.

**Service user outcomes assessed and corresponding findings.** Eight studies examined the impact of training on service user outcomes. Six studies in this category evaluated the impact of staff training on service user challenging behaviour and one study looked at skill acquisition including service users’ responses to choice opportunities (Salmento & Bambara, 2000). Finally, one study (Bradshaw et al.,
focused on the impact of training on student suspensions, office discipline referrals and academic achievement.

Of the eight studies examining service user outcomes following training, seven reported consistently positive results. The eighth study (Bradshaw et al., 2010) observed reduced student suspensions and discipline referrals. However, there was no significant difference between intervention and comparison schools, in terms of academic achievement.

Organisational outcomes. Three studies addressed the impact of staff training at an organisational level. The two studies, which evaluated training in school-wide Positive Behaviour Support, evaluated the implementation of this intervention at a whole school level (Bradshaw et al., 2010; Bradshaw et al., 2008). Both groups reported superior Positive Behaviour Support implementation in those schools who had received their training. Furthermore, Bradshaw et al. (2010) observed a significant reduction in suspensions and discipline referrals for schools who received their training package. A mindfulness-based Positive Behaviour Support study (Singh et al., 2015) examined the cost-effectiveness of the programme and the impact of training on staff turn-over. The authors reported a reduction of 87.75% ($133,380.00) in the organisation’s financial burden post-training. In addition, there were no staff resignations in the 40 weeks following training, compared to four resignations in the 40 weeks prior to the training.

Applied Behaviour Analysis (ABA) Practices

Behavioural assessment strategies. Seventeen studies met the criteria for inclusion in this subcategory of ABA (Appendix A).
**Characteristics of staff training.** Ten of the studies in this category focused on staff training in preference assessments. Preference assessments trained within these studies included paired-stimulus assessment alone (n=3); paired stimulus & multiple stimuli without replacement (MSWO) (n=3); free operant & paired stimulus & MSWO (n=2); free operant & MSWO (n=1), and single stimulus, paired stimulus & MSWO (n=1). Five of the studies provided training in conducting functional analyses and two studies examined staff training for implementing components of a functional behaviour assessment.

Eight of the 17 (Lambert et al., 2013) studies in this category used video modelling in their training package, with one of these combining it with a Behaviour Skills Training approach (Bishop & Kenzer, 2012) and another using it as part of a training comparison study (Moore & Fisher, 2007). Four studies implemented Behaviour Skills Training and a further two studies compared different training packages (Graff & Karsten, 2012; Roscoe et al., 2006). The remaining studies employed individualised training packages, with one of these also incorporating a pyramidal training (i.e., train-the-trainer) approach (Pence, St Peter, & Tetreault, 2012).

**Method of assessing staff learning.** All studies in this category assessed staff learning in some format. Twelve studies evaluated the impact of training on staff skill acquisition; one focused on changes to staff knowledge and four examined both knowledge and skill acquisition.

**Staff learning.** All studies showed that training positively affected staff knowledge and/or skill acquisition. However, although one study found that group-based classroom instruction was sufficient to teach seven of 11 staff to carry out a
paired-stimulus preference assessment, the remaining four trainees required in-vivo feedback to reach mastery (Bishop & Kenzer, 2012). Finally, the three comparison studies produced the following findings: (1) complete video modelling was the most effective method for training staff in functional analysis (Moore & Fisher, 2007); (2) performance-specific instruction was more successful than contingent money in promoting correct preference assessment implementation (Roscoe et al., 2006), and (3) enhanced written instructions were superior to standard written instructions when training staff to conduct preference assessments (Graff & Karsten, 2012).

**On-the-job performance assessment.** Fifteen of the 16 studies, which evaluated staff skill acquisition, investigated the implementation of the target training skill by participants in their typical work environment.

**Service user population.** Thirteen studies targeted staff working with children only, one study (Lambert et al., 2013) assessed staff supporting adult service users only and three studies (Borgmeier, Loman, Hara, & Rodriguez, 2015; Moore & Fisher, 2007; Roscoe & Fisher, 2008) did not specify whether the service users were children, adults or a combination of both.

**Service user outcomes assessed and corresponding findings.** Service user outcomes were not assessed in any of the included studies in this category.

**Organisational outcomes.** One study (Loman & Horner, 2014) examined the impact of training at an organisational level, by recording the time taken by staff to complete a functional behaviour assessment. The average duration for completion was almost 2 hours (M=119.40 minutes, range=65-275 minutes).

**Antecedent-based intervention studies.** Eight studies met the criteria for inclusion in this subcategory of ABA-based interventions (Appendix A).
**Characteristics of staff training.** Four studies in this category involved training staff to carry out a prompting procedure. The remaining four studies focused on training staff in one of the following: (1) behaviour specific praise statements; (2) writing and using social stories; (3) environmental organisation, and (4) systematic prompting procedures and following a task analysis to conduct a literacy lesson.

Three of the 8 studies, involved video modelling, combining it with the following additional strategies: (1) workshop and in-person coaching (Brock & Carter, 2015); (2) verbal instruction, modelling and role-play (Collins, Higbee, & Salzberg, 2009), and (3) feedback (Giannakakos, Vladescu, Kisamore, & Reeve, 2015). The remaining five studies employed individualised training programmes using a variety of different approaches, including instruction, rehearsal and feedback.

**Method of assessing staff learning.** All eight studies evaluated the ability of staff participants to implement the skills they learned during training.

**Staff learning.** All studies in this category found that staff training resulted in positive changes to staff skills. Brock and Carter (2015) also found that whilst video modelling and coaching were effective in training staff to use a time delay prompting procedure, coaching alone was superior to video modelling alone.

**On-the-job performance assessment.** All eight studies examined the ability of staff to demonstrate their newly acquired skills in their typical work environment.

**Service user population.** Seven of the eight studies trained staff who supported children, with one study (Collins et al., 2009) providing training for staff working with adults.
Service user outcomes assessed and corresponding findings. Three studies in this category addressed service user outcomes. Duchaine, Jolivete, and Fredrick (2011) assessed student on-task behaviour, reporting inconclusive findings. Quilty (2007) demonstrated an immediate reduction in inappropriate service user behaviour following staff training, and this decrease was observed to maintain over time. Finally, Browder and colleagues (2007) observed an increase in service user’s independent responding after training staff to follow a task analysis and provide systematic prompting.

Organisational outcomes. No studies in this category examined the impact of training at an organisational level.

Consequence-based intervention studies. Two studies met the criteria for inclusion in this subcategory of Applied Behaviour Analysis-based interventions (Appendix A).

Characteristics of staff training. One study focused on training staff to implement the ‘Good Behaviour Game’, a group-based contingency, within a classroom environment (Maggin et al., 2012). Behaviour Skills Training was used in this instance. The second study employed instructions, trainer prompting, self-monitoring and accuracy feedback to train staff to correctly implement a token economy system (Petscher & Bailey, 2006).

Method of assessing staff learning. Both studies examined the fidelity with which staff displayed skills targeted during training. Petscher and Bailey (2006) also evaluated changes in knowledge pertaining to these skills.

Staff learning. Both studies demonstrated substantial improvements in implementation fidelity. Maggin et al. (2012) showed that this effect was maintained
over time and also resulted in increased interaction between the trainees and their students.

**On-the-job performance assessment.** Both studies examined the ability of staff to demonstrate their newly acquired skills in their typical work environment.

**Service user population.** Each study in this category investigated staff training for people working with children.

**Service user outcomes assessed and corresponding findings.** Maggin et al. (2012) assessed changes to the children’s levels of aggression and found that the rate of aggressive behaviour reduced following staff training. Petscher and Bailey (2006) did not evaluate the impact of staff training on service user outcomes.

**Organisational outcomes.** Neither study investigated the impact of training at an organisational level.

“**Mixed” interventions.** Fifty studies met the criteria for inclusion in this subcategory (Appendix A).

**Characteristics of staff training.** Eleven of the studies trained staff in discrete trial teaching (DTT) alone. Four additional studies looked at one of the following areas: (1) DTT and backward chaining; (2) DTT & prompting & reinforcement; (3) DTT & pivotal response training (PRT) & incidental teaching and, (4) DTT & activity-based intervention. A further 11 studies sought to train staff in prompting procedures, combined with one or more other approaches: (1) prompting and curriculum assessment (n=1); (2) prompting & praise (n=3); (3) prompting & preference assessment (n=1); (4) prompting, preference assessment & incidental teaching (n=1); (5) prompting & appropriate trainer (teacher) behaviour (n=2); (6) prompting, reinforcement & error correction (n=2), and (6) prompting & manual
signing (n=1). Of the remaining studies, 10 provided training in principles of behaviour and associated topics, four trained staff in Pivotal Response Training alone, three focused on the Picture Exchange Communication System (PECS), three worked on training staff in alternative and augmentative forms of communication, other than PECS, two focused on vocal mand training, one provided training to enhance staff-service user interactions and one study trained staff in behaviour support plan implementation.

Nineteen of the 50 studies employed a video component as part of the training package (e.g., video modelling, video feedback, and video instructions). Sixteen studies used Behaviour Skills Training, with two of these incorporating a pyramidal/train-the-trainer approach also. The remaining studies adopted individualised training protocols, using one or more of the following elements: (1) instruction; (2) role-play; (3) feedback, and (4) on-the-job training.

In addition, four studies compared training protocols. Granpeesheh et al. (2010) contrasted an eLearning approach and didactic instruction, while Haberlin, Beauchamp, Agnew, and O'Brien (2012) compared direct staff training with pyramidal training. Macurik, O'Kane, Malanga, and Reid (2008) examined video versus live training and Nosik and colleagues (2013) evaluated the differential impact of computer based instruction and Behaviour Skills Training (Nosik et al., 2013).

**Method of assessing staff learning.** Thirty-five studies evaluated the fidelity with which staff implemented the target skills learned through training; six examined changes to knowledge pertaining to skills targeted in training and, nine studies assessed staff performance across both of these areas.
**Staff learning.** Forty-five studies demonstrated that training had a positive impact on staff knowledge and/or skill acquisition. Fourteen of the 45 studies also reported that the effect maintained over time. Two studies (Hall, Grundon, Pope, & Romero, 2010; Suhrheinrich, 2011) found that while the workshop component alone was insufficient in training staff to implement the target skills in their typical work environments, the addition of a performance feedback component effectively addressed this problem. Finally, three studies (Barnes et al., 2011; Da Fonte & Capizzi, 2015; Dowey, Toogood, Hastings, & Nash, 2007) did not find consistent changes in targeted staff knowledge or skills following training.

The results of studies that compared training packages are as follows: (1) didactic instruction was slightly superior to eLearning in terms of knowledge acquisition among staff (Granpeesheh et al., 2010); (2) pyramidal training was more effective than consultant-led, direct staff training for staff skill acquisition and maintenance over time (Haberlin et al., 2012); (3) both video and live training were successful in improving trainee knowledge and on-the-job performance (Macurik et al., 2008), and (4) trainees who received Behaviour Skills Training outperformed the computer-based training group immediately following training and at a 6-week follow-up (Nosik et al., 2013).

**On-the-job performance assessment.** Forty-three of the 44 studies, which examined staff skill acquisition as a result of training, reported evaluating the ability of staff participants to implement the skills targeted during training in their typical work environments.

**Service user population.** Thirty-one studies provided training to staff were working with children only, eight studies examined staff supporting adult service
users only and five focused on those assisting both adults and children. Six studies did not specify whether the service users were children, adults or a combination of both.

**Service user outcomes assessed and corresponding findings.** The majority of studies did not evaluate the impact of training on service user outcomes (n=30). Seventeen studies examined the effect of training on service user acquisition of new skills and adaptive behaviour. One study (Bingham, Spooner, & Browder, 2007) assessed the impact on service user engagement with an alternative communication system and the display of challenging behaviour, one examined the effect of training on service user stereotypy (Dib & Sturmey, 2007), and one evaluated the impact on service user challenging behaviour (Zoder-Martell et al., 2014).

There was an increase in adaptive behaviour and a reduction in inappropriate behaviour across the majority of studies. However, Wood et al. (2007) reported that their service user required a full physical prompt to complete the picture exchange protocol throughout the intervention. Zoder-Martell et al. (2014) had difficulty evaluating the impact of training on service user challenging behaviour, as the levels of engagement in this type of behaviour were low throughout the study.

**Organisational outcomes.** Macurik et al. (2008) and Robinson (2011) evaluated the effect of staff training at an organisational level. Macurik and colleagues (Macurik et al., 2008) examined whether live training or video training was a more efficient use of the behaviour analyst’s (trainer) time and resources. There was no definitive conclusion to this question as video training required less direct contact time with staff but the preparation of videos was found to be time consuming. Robinson (2011) investigated the efficiency of staff training in Pivotal
Response Training implementation and found that the degree of change in treatment implementation was substantial given the duration of the training programme.

“Other” Practices

Fourteen studies met the criteria for inclusion in this category (Appendix A).

Characteristics of staff training. Seven studies examined staff training in active support (AS). Two studies provided training in cognitive behaviour therapy (CBT) and the remaining five studies trained staff in practices to support service users with mental health issues (Brookman-Frazee et al., 2012); aggressive behaviour (Campbell & Hogg, 2008; Perkins & Leadbetter, 2002); physical health problems (Marks et al., 2013), and communication difficulties (Smidt, Balandin, Reed, & Sigafoos, 2007).

A range of individualised training packages were used across the 14 studies. Three studies (Jones et al., 2001; Marks et al., 2013; Smith, Felce, Jones, & Lowe, 2002) also adopted a train-the-trainer model as part of their approach.

Method of assessing staff learning. Six of the 14 studies examined the acquisition of target skills by staff trainees; two investigated the impact of training on the knowledge, and skills of staff and one assessed changes in staff knowledge only. Five studies did not evaluate the impact of training on staff knowledge or skills. Of the two studies which examined both knowledge and skill acquisition, only one assessed changes in skills directly (Smidt et al., 2007), the other relied on staff self-reports of how they would respond to challenging behaviour (Perkins & Leadbetter, 2002).

Staff learning. There was evidence of staff learning across all nine studies that evaluated this outcome. However, four of these studies found that the effects did
not maintain over time or were inconsistent across trainees (Jones et al., 2001; Perkins & Leadbetter, 2002; Smidt et al., 2007; Totsika, Toogood, Hastings, & McCarthy, 2010).

**On-the-job performance assessment.** Seven of the eight studies, which assessed staff skill acquisition, also investigated the ability of staff trainees to demonstrate their newly acquired skills in the typical work environment. Chou et al. (2011) did not directly examine the impact of training on staff skills. However, they did demonstrate that those service users, whose staff received training, had substantially better outcomes than service users, whose staff did not receive training.

**Service user population.** Two studies involved staff working with children, ten looked at staff training for those supporting adults and two studies did not specify if service users were children or adults.

**Service user outcomes assessed and corresponding findings.** Four studies did not assess service user outcome; two examined the impact on challenging behaviour; two focused on changes to service user engagement; five looked at the effect on both challenging behaviour and service user engagement, and one study addressed the health behaviours of service users.

All 10 studies, which investigated the impact of training on service user outcomes, demonstrated measurable changes post-training. However, three of these studies reported inconsistent results for the service users involved. Smith et al. (2002) reported that changes to service user engagement were contingent on a service user’s adaptive behaviour and levels of challenging behaviour pre-training. They found that service users with higher levels of adaptive skills and lower levels of challenging behaviour were significantly more likely to show increased engagement
than those service users with poorer adaptive skills and more severe challenging behaviour. On the other hand, Totsika et al. (2010) observed that their active support training programme only produced improvements for service users whose levels of aggression were highest pre-training. Finally, Chou et al. (2011) showed that residents, whose staff received training, displayed increased levels of choice and adaptive behaviour and decreased depression in post-test and follow-up probes. However, there was no impact on the frequency of family contact, community inclusion or challenging behaviour. Furthermore, levels of depression also decreased for those residents, whose staff did not receive training.

**Organisational Outcomes.** One study (Willner et al., 2013) examined the cost of the training programme and the impact on relevant resource use. They reported that while their training programme was almost twice the cost of treatment as usual, it did not produce significant changes in service users’ self-rated anger (although it did positively impact service users’ anger coping skills and staff-rated anger).

**Discussion**

The current systematic review examined the literature, which addresses staff training in practices used to treat skill deficits and challenging behaviour for individuals with intellectual and developmental disabilities. The review comprised 107 studies, published between the years 2000 and 2015. Studies were categorised according to the practices in which staff were trained. Results showed that while the range of practices targeted for training were generally broad, the selection was nonetheless limited. The application of many established evidence-based practices was not addressed. For example, functional communication training is an
Chapter 5

empirically supported treatment for individuals with intellectual disabilities (Kurtz et al., 2011). However, the current review did not uncover any studies, which provided training in the implementation of this practice. Although there were a limited number of studies, which assessed training in the Picture Exchange Communication System (PECS), Augmentative and Alternative Communication (ACC) and vocal requesting, none involved explicitly replacing an inappropriate behaviour, with a functionally equivalent communication behaviour.

In addition, non-contingent reinforcement, a function-based approach to challenging behaviour, has been shown to be an evidence-based intervention for individuals with intellectual disabilities (Carr et al., 2009). However, none of the articles reviewed investigated staff training in this practice. The following strategies, were also notably absent from the current review: (1) shaping; (2) discrimination training; (3) reinforcement schedules; (4) differential reinforcement techniques, and (5) extinction implementation (including planned ignoring and escape-extinction). Therefore, despite rigorous and consistent scientific support for these behavioural interventions, the published literature does not appear to concurrently evaluate training protocols to successfully disseminate these practices to frontline staff. Such findings provide a potential explanation for the current disconnect between theory and effective practice in the field of intellectual disability (Campbell, 2010; Hile & Walbran, 1991; Rapp et al., 2010; Swain et al., 2010).

In spite of this significant limitation in the published research, it must be emphasised that over 30 interventions used to treat skill deficits and challenging behaviour, were targeted by studies in the current review. Findings reveal positive gains for both staff learning and client behaviour immediately following training across the majority of studies that assessed these outcomes. All of the Applied
Behaviour Analytic-based intervention studies (i.e., behavioural assessment practices; antecedent-based practices; consequence-based practices and, mixed interventions) and the majority of articles from the remaining two categories (i.e., Positive Behaviour Support and ‘Other’) assessed staff learning, either through knowledge gains, skill acquisition or a combination of both. A significant number of Applied Behaviour Analytic studies focused on skill development (n=56) or skill and knowledge gains (n=14) among staff, with only seven studies addressing knowledge acquisition alone. This may be considered a particularly positive result, as research has demonstrated that knowledge gains alone are typically insufficient in producing desired changes in staff-service user interactions following training (Embry & Biglan, 2008; Grey, Hastings, & McClean, 2007; Hall et al., 2010; Suhrheinrich, 2011). In addition, although largely evident, the degree to which staff learning and service user outcomes were maintained over time, was variable across studies. This highlights the importance of in-service training and on-going supervision to ensure continued implementation fidelity and consistency following initial training (McCabe et al., 2015; Rapp et al., 2010; Rapp et al., 2008; Reed & Henley, 2015; Swain et al., 2010).

Although a substantial portion of the training protocols reviewed had a positive impact, only eight studies in total contrasted the effects of different packages on staff and service user outcomes. In the future, comparison studies are necessary to evaluate which protocols are most effective and efficient for training staff in particular strategies. Furthermore, these comparison studies must evaluate the impact at an organisational level. Resource investment by intellectual disability services might be more forthcoming if research can demonstrate a tangible benefit (e.g., reduced staff turn-over, decrease in sick-leave, resources savings) through the
implementation of high quality, empirically supported staff training. Therefore, if future investigations adopt a practical, applied and evidence-based approach to the evaluation and promotion of staff training, people with intellectual and developmental disabilities might be more likely to access superior, empirically supported programmes that facilitate the timely attainment of their personal goals, increased independence and improved quality of life.
Chapter 6

An Examination of the Adaptive Functioning and Behavioural Profiles of Service Users, in addition to the Training Needs of Staff Providing Direct Care within a Large Intellectual Disability Service Provider Organisation
Abstract

Individuals with intellectual disabilities have the right to the highest standard of treatment and support but research has shown that access to such standards is not always attainable. The goal of the current study was to conduct a training needs analysis within the host organisation to obtain a broad, comprehensive analysis of the diagnostic, adaptive functioning and behavioural profiles of a representative sample of service users, as well as the training needs of staff providing direct care to these individuals. The information yielded from this analysis would be utilised in designing a subsequent staff training programme to address deficient skills in selecting and applying evidence-based practice for persons with intellectual disability. Sixty-two frontline staff participated in Phase I, providing information on the adaptive functioning, behavioural profiles and quality of life of their key service users. Ten staff participated in Phase II, which involved practical observations of staff supporting service users with skill acquisition in the typical work environment. Finally, 15 service users whose keys staff had participated in Phase I, were identified as having behavioural intervention plans in place to support their management of challenging behaviour. The standard of these plans was evaluated in Phase III. Results demonstrated a disconnect between the training received by staff and the adaptive and behavioural needs of the service users. In addition, communication, social skills, and daily living skills were identified as priority areas for service users within the host organisation.
The National Working Group, which was established by the Health Service Executive (HSE) in 2007, conducted a comprehensive review of day facilities for people with disabilities in the Republic of Ireland. The focus of this group was to examine the level of progress made by intellectual disability service providers, in adopting best practice, improving outcomes for their service users and achieving cost-effective service provision. The published report (Health Service Executive, 2012a) strongly recommended a thorough overhaul of day services in the Republic of Ireland to ensure the implementation of individualised supports that are flexible and responsive to each service user’s needs. In addition, these supports and their outcomes must be quantifiable to monitor and ensure the highest quality of service provision possible (Health Service Executive, 2012a).

As discussed in Chapter 1, people with intellectual disabilities experience significant limitations, in terms of their intellectual functioning and adaptive behaviour (American Association on Intellectual and Developmental Disabilities, 2017). In addition, this group are at an increased risk for engagement in challenging behaviour, including aggression, self-injury and property destruction (Emerson et al., 2001; National Disability Authority, 2003; National Institute for Health and Care Excellence, 2015). The comorbid presence of challenging behaviour magnifies the difficulties faced by people with intellectual disabilities (Emerson, 2001; Schalock et al., 1994) and complicates the delivery of services and supports by organisations and frontline staff (Emerson, 2001).

To further exacerbate this situation, many people working in the intellectual disability sector do not have qualifications directly related to their role (Campbell, 2010). The Joint Committee on Social Care Professionals (2002) examined the qualifications of staff employed in residential settings for individuals with
intellectual disabilities in the Republic of Ireland. According to their findings, 36% of staff did not have a qualification in social care. Considering that research has shown that compromised skills sets among intellectual disability staff can have substantial repercussions for service users (e.g., Finn & Sturmey, 2009; Jahr, 1998; Schepis et al., 2001), this anomaly poses a serious issue for the intellectual disability sector. However, research has reported that staff with varying backgrounds and educational qualifications can be effectively trained to carry out procedures in line with evidence-based practice (Allen & Tynan, 2000; Dench, 2005; Grey & Hastings, 2005). As such, a well-designed, comprehensive, in-house training programme could overcome the challenges posed by hiring staff without the requisite qualifications.

Van Oorsouw and colleagues (van Oorsouw et al., 2009) conducted a meta-analysis to identify and examine effective structural components of staff training. They concluded that the combined use of in-service training and on-the-job coaching produces optimal results. The authors recommended delivering target information to staff in a number of different ways (e.g., didactic instruction, modelling, and role-play) to enhance staff performance. These recommendations are consistent with Behaviour Skills Training, which employs a combination of instruction, modelling, rehearsal, and feedback (Ward-Horner & Sturmey, 2012).

Behaviour Skills Training has been shown to be a robust, efficacious method for teaching new skills. It has been successfully applied to train a wide range of groups, including people with intellectual disabilities, undergraduate students, frontline staff and other professionals and there is repeated support for the generalisation of the newly learned skills to novel environments, without additional training (e.g., Fetherston & Sturmey, 2014; Homlitas et al., 2014; Iwata et al., 2000;
Maggin et al., 2012; Sarokoff & Sturmey, 2008). The importance of implementing high quality, effective training is further highlighted by Swain et al. (2010). While investigating the sustainability of practices promoted by the ‘National Implementing Evidence Based Practices Project’ within forty-nine organisations in North America, the authors found that organisations, which maintained evidence-based practices had consistently scored higher on staff training measures.

Although frontline staff are considered one of the most valuable assets within the intellectual disability sector (Finn & Sturmey, 2009; Jahr, 1998), the demands associated with working with people with an intellectual disability are often particularly stressful (Maslach, 1999). Staff in intellectual disability settings are at an increased risk for illness, absenteeism and employee turnover (Rose, 1995), and situations, in which stress is not successfully managed, may result in staff burnout over time (Devereux, Hastings, & Noone, 2009). This can be costly for service provider organisations, in terms of financial impact and impaired service delivery (Seaward, 2004). According to work stress theories, perceived work-related demands are pivotal in the development of stress and burnout (Devereux, Hastings, & Noone, 2009) and the perception of these demands can be impacted by the psychological, social and organisational resources available to staff (Devereux, Hastings, Noone, et al., 2009; Lazarus, 1999; Lazarus & Folkman, 1984). Although this topic has represented a significant research focus within psychology (e.g., Devereux, Hastings, Noone, et al., 2009; Gibson et al., 2009; Hastings, 2002; Kozak et al., 2013; Maslach, 1999; Smyth et al., 2015), more research is required to fully comprehend this process.

The goal of the current study was to conduct a training needs analysis within the host organisation to identify the precise needs of staff in effectively supporting
service-users. The information yielded from this analysis would be utilised in
designing a subsequent staff training programme to address deficient skills in
selecting and applying evidence-based practice for persons with intellectual
disability. The study had three key objectives. The first objective was to develop an
understanding of the service user profile within the RehabCare organisation. It was
hypothesised that the results of this training needs analysis, would highlight areas of
need for service users, in terms of acquiring adaptive behavior skills, achieving a
suitable quality of life and managing challenging behaviour. The second objective
was to examine the training provided to the frontline staff, in the context of their role
as direct-support workers for people with intellectual disabilities.

The final objective of this study was to investigate the factors associated with
burnout among frontline staff in the host organisation. Based on prior research, it
was predicted that service user engagement would be positively correlated with staff
burnout (Hastings, 2002; Kozak et al., 2013; Smyth et al., 2015) and negatively
correlated with ‘wishful thinking’ (Devereux, Hastings, Noone, et al., 2009; Griffith
et al., 2014; Mitchell & Hastings, 2001), job satisfaction (Dyer & Quine, 1998;
Kozak et al., 2013), and perceived supervisory support (Gibson et al., 2009; Ito et al.,
1999).

Phase I – Administration of Measures

Method

Participants. Participants were 62 frontline staff working with adults with
intellectual disabilities. Table 7 shows the characteristics of participants according to
gender and length of service with the host organisation. As part of this study, each
participant was required to share information about their key service user. A key
service user was an individual whom the participant had known and observed in one
or more environments for at least three months. Table 8 summarises the characteristics of service users according to age (upper panel), gender (upper panel), primary placement (lower panel), and duration of placement with the host organisation (lower panel).

Table 7

*Participant Characteristics According to Gender and Length of Service with the Organisation*

<table>
<thead>
<tr>
<th>Participant Gender</th>
<th>Number of Participants</th>
<th>Length of Service with the Organisation</th>
<th>Number of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>10</td>
<td>&lt; 1 year</td>
<td>4</td>
</tr>
<tr>
<td>Male</td>
<td>10</td>
<td>1-5 years</td>
<td>22</td>
</tr>
<tr>
<td>Male</td>
<td>10</td>
<td>6-10 years</td>
<td>19</td>
</tr>
<tr>
<td>Male</td>
<td>10</td>
<td>11+ years</td>
<td>17</td>
</tr>
<tr>
<td>Female</td>
<td>52</td>
<td>&lt; 1 year</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>52</td>
<td>1-5 years</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>52</td>
<td>6-10 years</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>52</td>
<td>11+ years</td>
<td></td>
</tr>
</tbody>
</table>
Table 8

*Service user Characteristics According to Age and Gender (Upper Panel) and Type of Placement and Length of Time with the Host Organisation (Lower Panel)*

<table>
<thead>
<tr>
<th>Age</th>
<th>Number of Service Users</th>
<th>Gender</th>
<th>Number of Service Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>18-25 years</td>
<td>12</td>
<td>Male</td>
<td>33</td>
</tr>
<tr>
<td>26-35 years</td>
<td>15</td>
<td>Female</td>
<td>27</td>
</tr>
<tr>
<td>36-46 years</td>
<td>11</td>
<td>Missing data</td>
<td>2</td>
</tr>
<tr>
<td>46-55 years</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>56-65 years</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>65+ years</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type of Placement</th>
<th>Number of Service Users</th>
<th>Duration with Organisation</th>
<th>Number of Service Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residential</td>
<td>17</td>
<td>&lt; 1 year</td>
<td>1</td>
</tr>
<tr>
<td>Resource Centre</td>
<td>17</td>
<td>1-2 years</td>
<td>6</td>
</tr>
<tr>
<td>Assisted Living</td>
<td>4</td>
<td>3-5 years</td>
<td>8</td>
</tr>
<tr>
<td>Independent Living</td>
<td>3</td>
<td>6-10 years</td>
<td>14</td>
</tr>
<tr>
<td>Respite</td>
<td>2</td>
<td>11-15 years</td>
<td>15</td>
</tr>
<tr>
<td>Residential &amp; Resource</td>
<td>11</td>
<td>16-20 years</td>
<td>9</td>
</tr>
<tr>
<td>Resource &amp; Independent Living</td>
<td>3</td>
<td>20+ years</td>
<td>7</td>
</tr>
<tr>
<td>Resource &amp; Respite</td>
<td>2</td>
<td>Missing data</td>
<td>2</td>
</tr>
<tr>
<td>Resource &amp; Assisted Living</td>
<td>1</td>
<td>Missing data</td>
<td>2</td>
</tr>
<tr>
<td>Missing data</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Setting.** Participants in the current study were recruited from RehabCare; an intellectual disability service provider, in the Republic of Ireland. This organisation provides a range of health and social care services to more than 3,000 people with an intellectual disability across 88 geographical locations throughout the country. Over 1,000 staff are employed to deliver these services, across a variety of community-based applied settings, including resource centres and residential settings. Sessions were hosted by RehabCare, in either a conference room or a classroom setting. The rooms were large enough to seat up to 10 participants and each participant had access to a standard laptop.
Measures.

Demographic questionnaire. The questionnaire relating to service user demographics focused on the age, gender, and diagnosis of individual service users, as well as the nature of their placement and the length of time that they have been in receipt of services from RehabCare. The staff demographic questionnaire addressed gender, length of service with RehabCare, and the nature of preparatory and ongoing training in behaviour management strategies and methods to support service users in the learning of new skills, which is currently available to staff through the host organisation (questionnaire available from the author on request).

Life Skills Assessment (Hanley et al., 2007). The life skills assessment was developed to evaluate and promote prosocial skills among children aged 3-5 years. The assessment is a 19-item measure that assesses adaptive functioning in the following areas: (1) instruction following; (2) functional communication; (3) coping and tolerance skills, and (4) friendship and leisure skills. Each item is rated on a four-point Likert scale, with “does not respond” scoring a 1 and “responds independently most of the time” scoring a 4. It has been reported that implementation of a programme based on this assessment results in significant reductions in challenging behaviour and substantial increases in age-appropriate life skills among children, including those diagnosed with developmental delays (Hanley, Fahmie, & Heal, 2014; Hanley et al., 2007). The Life Skills Assessment was adapted for the purposes of the current study, to examine the life skill repertoires of adults with intellectual disabilities (available from the author on request).

Behaviour Problems Inventory – Short Form (BPI-S; Rojahn et al., 2012a). The BPI-S is a 30-item respondent-based rating scale, which evaluates the frequency
and severity of challenging behaviour displayed by individuals with an intellectual disability. The measure is divided into three subscales that assess self-injurious behaviour, stereotypic behaviour, and aggressive/destructive behaviour. The BPI-S uses two Likert-scale rating scales per item; a five-point frequency scale and a four-point severity scale. Rojahn et al. (2012b) evaluated the psychometric properties of this measure using archival data from 1122 cases across the USA, Wales, England, the Netherlands, and Romania. Results indicated good internal consistency and confirmatory and discriminant validity. These findings were replicated by Mascitelli and colleagues (Mascitelli et al., 2015).

Knowledge of Behaviour Management Principles Questionnaire (KBMQ; Blair & Eldridge, 1997). The KBMQ is a 30-item questionnaire based on four dimensions of behaviour management principles: (1) identifying and assessing behaviour; (2) identifying environmental influences on behaviour; (3) identifying behavioural procedures used to establish and maintain desired behaviour, and (4) identifying behavioural procedures to reduce inappropriate behaviours. The measure is a multiple-choice instrument, with four possible responses for each item. Correct answers are scored as 1 and incorrect answers as 0. The reliability and validity of this measure for use with adults diagnosed with a psychiatric disorder and living in long-term residential settings have been assessed and found to be adequate (Blair & Eldridge, 1997). According to Blair and Eldridge (1997), the KMBQ has good internal consistency, with a Cronbach’s alpha coefficient of 0.87.

Maslach Burnout Inventory – Human Services Survey (MBI-HSS; Maslach et al., 1996). The MBI-HSS is a 22-item respondent-based rating scale, which evaluates the risk of burnout among individuals working in the human services sector. The measure is divided into three subscales that emotional
exhaustion, client depersonalization, and personal achievement. The MBI-HSS uses a six-point Likert-scale rating for each item. The reliability and validity of this measure for use with individuals working in the human services sector have been assessed and found to be adequate (Gold, 1984; Iwanicki & Schwab, 1981; Maslach et al., 1996).

**Minnesota Satisfaction Questionnaire – Short Form (MSQ-SF; Weiss, Dawis, England, & Lofquist, 1967).** The MSQ-SF is a 20-item respondent-based measure with three scales: (1) intrinsic satisfaction; (2) extrinsic satisfaction, and (3) general satisfaction. The MSQ-SF uses a five-point Likert-scale rating for each item, with “very dissatisfied” scoring a 1 and “very satisfied” scoring a 5. Several studies have found the MSQ-SF to a reliable and valid instrument for the measurement of job satisfaction (Dunham, Smith, & Blackburn, 1977; Fields, 2002; Gillet & Schwab, 1975).

**Personal Outcomes Scale (POS; Van Loon, Van Hove, Schalock, & Claes, 2009).** The POS is an 8-item instrument that measures the quality of life of an individual with an intellectual disability. The POS is divided into eight subscales, which correspond to the eight quality of life core domains; personal development, self-determination, interpersonal relations, social inclusion, civil rights, emotional well-being, physical well-being, and material well-being. Each item is scored on a three-point Likert scale. The measure is administered by a professional who has substantial experience working in the field of intellectual disabilities and training in the administration and scoring of behavioural assessments. The administration of the POS can occur under two conditions; self-report by the individual with the intellectual disability or through direct observation. Direct observation involves a respondent who has known the individual being assessed for at least three months.
and has had the opportunity to observe this individual in one or more environments
during this time. The respondent can be a parent, relative or professional who has
lived or worked with the individual in question. The POS has been evaluated and
shown to be a valid and reliable instrument for the measurement of quality of life for
people with intellectual disabilities (Carbó-Carreté, Olmos, & Giné, 2015).

**The Perceptions of Supervisory Support Scale (PSS; Fukui et al., 2014).**
The PSS is a 19-item respondent-based, which evaluates frontline staffs’ perceptions
of the formal and informal supervisory support that they receive in the workplace.
The PSS uses a six-point Likert-scale rating for each item. The psychometric
properties of this measure, including reliability and validity, have be adequately
demonstrated (Fukui et al., 2014).

**The Shortened Ways of Coping Questionnaire (SWC-R; Hatton &
Emerson, 1995).** The SWC-R is a 14-item respondent-based measure used to assess
the coping style of individuals working as frontline staff in the human services
sector. The measure is divided into two subscales; practical coping and wishful
thinking. Each item is scored on a four-point Likert scale, with “not used” scoring a
1 and “used a great deal” scoring a 4. The psychometric properties of the SWC-R
have been assessed and found to be adequate (Hatton & Emerson, 1995).

Rating Form (Vineland-II; Sparrow, Cicchetti, & Balla, 2005).** The Vineland-II is
a multi-item measure of adaptive behaviour for ages birth through 90 years. The
scales can be administered by interview or through individual completion of a rating
form. The Parent/Caregiver Rating Form assesses adaptive functioning across the
domains of communication, daily living skills, socialisation, and motor skills. There
is also a section that examines engagement in challenging behaviour. The Vineland-
II has been evaluated and shown to be a valid and reliable instrument for the measurement of adaptive functioning and engagement in challenging behaviour for a wide range of populations, including people with intellectual disabilities (McCauley et al., 2012; Sparrow et al., 2005).

**Design.** A cross-sectional design was used to examine the training needs of frontline staff employed in the intellectual disability sector, as well as the profile of service users who were receiving services from this sector.

**Procedure.** A range of standardised and non-standardised measures were selected for completion by participants, as part of the training needs analysis. The goal of this analysis was to obtain a broad and comprehensive understanding of the diagnostic, adaptive functioning and behavioural profiles of service users, as well as the training needs of staff within the host organisation. Participants completed the majority of measures via a “Google Forms” platform on a laptop. The Vineland-II was completed through paper and pencil format.

The sessions were a maximum of 3 hours in duration and involved 1-7 participants at one time independently completing the battery of measures. During this time, the experimenter also conducted an interview, which was based on the Personal Outcomes Scale, with each participant. The setting for the interviews ensured that participant confidentiality was maintained. Consent was obtained from each participant prior to starting the sessions. Prior consent was also sought from the key service users. Specialised consent forms were designed to facilitate the service users’ understanding that their key worker would share information about them with the research team (available from the author on request). Guardian consent was obtained for those service users who were unable to provide consent (available from the author on request).
Data analysis.

**Missing data.** Missing and incomplete data is a pervasive problem encountered by researchers, including those conducting psychological research (Azar, 2002; Downey & King, 1998). While data may be incomplete for any number of reasons, it is particularly important to determine whether the data is missing at random or not (Little, 1988). If a variable is missing completely at random, the missingness is independent of all other variables in the data set (Little, 1988). On the other hand, if data is not missing at random, the pattern of incomplete data is dependent on other variables in the data set.

Pairwise deletion is cited as one of the most frequently used strategies to address the issue of missing data (Peugh & Enders, 2004). This technique maximises statistical power by utilising all data available for each analysis conducted. However, to use this approach, the pattern of incomplete data must be shown to be occurring completely at random (Gelman & Hill, 2006). If data are missing completely at random, the removal of cases with missing data will not “bias your inferences” (p. 530) (Gelman & Hill, 2006).

Analysis of missing data across the staff-related measures showed that 1.6% of the data was missing across each variable (Table 9). Visual inspection of the data subsequently revealed that one participant (case) was missing all scores across all measures. Therefore, this participant was excluded from all analyses, which involved staff-related measures only. Analysis of the missing data across the service user-related measures showed that 4.8% -31.7% of data was missing across each measure. Table 10 provides a summary of missing and valid cases across each measure.
As there is no recognised standard in the literature for an acceptable percentage of incomplete data, it was decided to exclude measures missing at least 30% of cases or greater. Therefore, frequency of self-stimulatory behaviour (as measured by the Behaviour Problems Inventory – Short Form) and the overall adaptive behaviour score on the Vineland-II were removed from all further analyses. All measures retained for future analyses provided complete data for at least 50 participants.

Visual inspection of the data revealed that two participants (cases) were missing all data for all measures, one participant omitted responses on all items for the Behaviour Problems Inventory – Short Form and the Life Skills Assessment (Hanley et al., 2007) and one participant omitted responses on all items for the Personal Outcomes Scale and the Vineland-II. These participants were excluded from all analyses, which involved service user-related measures only.

Finally, Little’s Missing Completely at Random (MCAR) test obtained for the remaining Vineland-II data (i.e., the overall adaptive behaviour scores were excluded) indicated that the data were indeed missing completely at random (i.e., no identifiable pattern exists in the missing data); $\chi^2 (20) = 13.97, p = .832$. As a result, pairwise deletion was used to address the issue of missing data in analyses involving the Vineland-II data set.
Table 9

*Summary of Missing and Valid Cases across Staff-Related Measures*

<table>
<thead>
<tr>
<th>Measure</th>
<th>Number of Missing Cases</th>
<th>Percentage of Missing Cases</th>
<th>Number of Valid Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shortened Ways of Coping Questionnaire (SWC-R)</td>
<td>1</td>
<td>1.6%</td>
<td>62</td>
</tr>
<tr>
<td>Perceptions of Supervisory Support Scale (PSS)</td>
<td>1</td>
<td>1.6%</td>
<td>62</td>
</tr>
<tr>
<td>Maslach Burnout Inventory (MBI-HSS)</td>
<td>1</td>
<td>1.6%</td>
<td>62</td>
</tr>
<tr>
<td>Minnesota Satisfaction Questionnaire (MSQ-SF)</td>
<td>1</td>
<td>1.6%</td>
<td>62</td>
</tr>
<tr>
<td>Knowledge of Behaviour Management Principles Questionnaire (KBMQ)</td>
<td>1</td>
<td>1.6%</td>
<td>62</td>
</tr>
<tr>
<td>Staff Demographic Information</td>
<td>1</td>
<td>1.6%</td>
<td>62</td>
</tr>
</tbody>
</table>

Table 10

*Summary of Missing and Valid Cases across Service User-Related Measures*

<table>
<thead>
<tr>
<th>Measure</th>
<th>Number of Missing Cases</th>
<th>Percentage of Missing Cases</th>
<th>Number of Valid Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vineland-II - Overall adaptive behaviour</td>
<td>20</td>
<td>31.7%</td>
<td>43</td>
</tr>
<tr>
<td>BPI-S - Frequency of self-stimulatory behaviour</td>
<td>19</td>
<td>30.2%</td>
<td>44</td>
</tr>
<tr>
<td>Vineland-II - Social Skills</td>
<td>13</td>
<td>20.6%</td>
<td>50</td>
</tr>
<tr>
<td>Vineland-II - Communication skills</td>
<td>12</td>
<td>19%</td>
<td>51</td>
</tr>
<tr>
<td>Vineland-II - Living skills</td>
<td>11</td>
<td>17.5%</td>
<td>52</td>
</tr>
<tr>
<td>Vineland-II - Overall maladaptive behaviour</td>
<td>3</td>
<td>4.8%</td>
<td>60</td>
</tr>
<tr>
<td>Vineland-II - Externalised maladaptive behaviour</td>
<td>3</td>
<td>4.8%</td>
<td>60</td>
</tr>
<tr>
<td>POS - Well-being</td>
<td>3</td>
<td>4.8%</td>
<td>60</td>
</tr>
<tr>
<td>POS - Social participation</td>
<td>3</td>
<td>4.8%</td>
<td>60</td>
</tr>
<tr>
<td>POS - Independence</td>
<td>3</td>
<td>4.8%</td>
<td>60</td>
</tr>
<tr>
<td>LSA - Leisure skills</td>
<td>3</td>
<td>4.8%</td>
<td>60</td>
</tr>
<tr>
<td>LSA - Tolerance</td>
<td>3</td>
<td>4.8%</td>
<td>60</td>
</tr>
<tr>
<td>LSA - Functional communication</td>
<td>3</td>
<td>4.8%</td>
<td>60</td>
</tr>
<tr>
<td>LSA – Instruction following</td>
<td>3</td>
<td>4.8%</td>
<td>60</td>
</tr>
<tr>
<td>BPI-S – Frequency of aggression</td>
<td>3</td>
<td>4.8%</td>
<td>60</td>
</tr>
<tr>
<td>BPI-S – Severity of aggression</td>
<td>3</td>
<td>4.8%</td>
<td>60</td>
</tr>
<tr>
<td>BPI-S – Frequency of self-injury</td>
<td>3</td>
<td>4.8%</td>
<td>60</td>
</tr>
<tr>
<td>BPI-S – Severity of self-injury</td>
<td>3</td>
<td>4.8%</td>
<td>60</td>
</tr>
</tbody>
</table>

LSA - Life Skills Assessment (Hanley et al., 2007)
Those cases that had been identified as missing data across all staff-related measures (N=1) or service-user measures (N=2) were excluded from analyses examining relationships between staff and service user measures. Frequency of self-stimulatory behaviour (as measured by the Behaviour Problems Inventory – Short Form) and the overall adaptive behaviour score on the Vineland-II were also excluded because at least 30% of cases were missing data.

Little’s Missing Completely at Random (MCAR) test obtained for the remaining overall data set, including the Vineland-II scores, indicated that the data was indeed missing completely at random (i.e., no identifiable pattern exists in the missing data); $\chi^2(173) = 156.47, p = .811$. Little’s Missing Completely at Random (MCAR) test obtained for the remaining overall data set, excluding the Vineland-II scores, also indicated that the data was indeed missing completely at random (i.e., no identifiable pattern exists in the missing data); $\chi^2(28) = 34.62, p = .181$. As a result, pairwise deletion was used to address the issue of missing data in analyses involving the Vineland-II data set.

**Likert scale data.** There is much controversy within the statistical literature as to whether or not it is appropriate to analyse Likert scale data using parametric statistical procedures (e.g., Jamieson, 2004; Lubke & Muthén, 2004). It is argued that because Likert scale data are based on ranked categories, the intervals between scale items are not equal. Therefore, it is recommended that nonparametric statistical tests should be used to analyse this type of data (Jamieson, 2004). As a result, the Spearman correlation, which is designed to quantify relationships for ordinal data,
was used to quantify all relationships involving Likert scale-based measures in this study.

**Dichotomous and interval data.** The point-biserial correlation is used to quantify the relationship between a variable measured on an interval or ratio scale and a dichotomous variable (i.e., has only two different values) (Gravetter & Wallnau, 2004). The continuous variable must be normally distributed and display homogeneity of variance for each category of the dichotomous variable to infer valid assumptions from this test (Gravetter & Wallnau, 2004). In this study, the Knowledge of Behaviour Management Principles Questionnaire scores were characterised as interval data so the relationship between this variable and a number of dichotomous variables was investigated (i.e., whether participants received training in behaviour management or skills teaching strategies prior to working with service users and whether participants received ongoing training in behaviour management or skills teaching strategies). However, visual analysis of the frequency and z-score distributions for the Knowledge Behaviour Management Principles Questionnaire scores across each category of the dichotomous variables revealed that the assumptions of normality and equality of variance had been violated. Therefore, the point-biserial correlation results were biased and could not be interpreted. As a result, a bootstrapped independent $t$-test was conducted to quantify the relationship between scores on the Knowledge of Behaviour Management Principles Questionnaire and the dichotomous variables (Field, 2016).

**Results**

**Service User-Related Measures**
Adaptive functioning. Figures 2, 3, and 4 present the percentage of service users scoring as “low”, “moderately low” or “adequate” across the Vineland-II domains of communication, daily living skills, and social skills. According to participant reports, a total of 34 (56.7%), 33 (55%), and 29 (48.3%) service users presented at the lowest level of functioning across the areas of receptive, expressive, and written communication, respectively (Figure 2). Similarly, a total of 40 (66.7%), 33 (55%), and 40 (66.7%) service users presented at the lowest level functioning for personal, domestic, and community daily living skills, respectively (Figure 3). Finally, 38 (63.3%), 33 (43.3%), and 27 (45%) service users were reported as displaying low levels of interpersonal, leisure, and coping skills, respectively (Figure 4).

Figure 2. The percentage of service users scoring “low”, “moderately low” or “adequate” across the communication domains, as measured by the Vineland-II.
Figure 3. The percentage of service users scoring “low”, “moderately low” or “adequate” across the daily living skills domains, as measured by the Vineland-II.

Figure 4. The percentage of service users scoring “low”, “moderately low” or “adequate” across the social skills domains, as measured by the Vineland-II.
**Challenging behavior.** Figure 5 presents the percentage of service users displaying “clinically significant”, “elevated” or “average” levels of externalised challenging behavior, as measured by the Vineland-II. According to participant reports, 54 (90%) service users displayed either “clinically significant” (38.3%) or “elevated” (51.7%) levels of externalised challenging behaviour (Figure 5).

![Figure 5. The percentage of service users displaying “clinically significant”, “elevated” or “average” levels of externalised challenging behaviour, according to participant reports on the Vineland-II.](image)

**Adaptive functioning and challenging behaviour.** Service users, with higher scores on the Vineland-II measure of externalised challenging behaviour, scored lower on the Vineland-II measure of social skills, $r_s = -.45$, $n = 49$, $p = .001$.

Service users, who engaged in self-injurious behaviour more frequently (as measured by the Behaviour Problems Inventory – Short Form) scored lower on the Vineland-II communication, $r_s = -.35$, $n = 50$, $p = .012$, and daily living skills
domains, $r_s = -.28, n = 51, p = .047$. Higher frequency of self-injurious behaviour was also associated with lower scores on instruction following, $r_s = -.53, n = 59, p < .001$, functional communication, $r_s = -.43, n = 59, p = .001$, tolerance, $r_s = -.43, n = 59, p = .001$, and leisure skills, $r_s = -.49, n = 59, p < .001$, as measured by the Life Skills Assessment (Hanley et al., 2007).

Service users, with more severe self-injurious behaviour (Behavior Problems Inventory – Short Form) scored lower on the Vineland-II domains of communication, $r_s = -.31, n = 50, p = .028$, and social skills, $r_s = -.30, n = 49, p = .036$. More severe displays of self-injurious behaviour was also associated with lower scores on instruction following, $r_s = -.53, n = 59, p < .001$, functional communication, $r_s = -.44, n = 59, p < .001$, tolerance, $r_s = -.42, n = 59, p = .001$, and leisure skills, $r_s = -.49, n = 59, p < .001$, as measured by the Life Skills Assessment (Hanley et al., 2007).

Service users, with elevated frequencies of aggression (Behavior Problems Inventory – Short Form) also scored lower on instruction following, $r_s = -.47, n = 59, p < .001$, functional communication, $r_s = -.54, n = 59, p < .001$, tolerance, $r_s = -.48, n = 59, p < .001$, and leisure skills, $r_s = -.64, n = 59, p < .001$, as measured by the Life Skills Assessment (Hanley et al., 2007).

Service users, with more severe displays of aggression (Behavior Problems Inventory – Short Form) scored lower on instruction following, $r_s = -.55, n = 59, p < .001$, functional communication, $r_s = -.62, n = 59, p < .001$, tolerance, $r_s = -.52, n = 59, p < .001$, and leisure skills, $r_s = -.67, n = 59, p < .001$, as measured by the Life Skills Assessment (Hanley et al., 2007).
Table 11 summarises all Spearman correlations between the service user-related measures of challenging behaviour and adaptive functioning.

Table 11

*Spearman Correlations Between Service User-Related Measures of Challenging Behaviour and Adaptive Functioning

<table>
<thead>
<tr>
<th>Measure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Vineland-II - communication</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Vineland-II - daily living skills</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Vineland-II - social skills</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>4. Vineland-II - challenging behaviour</td>
<td>-.28</td>
<td>-.14</td>
<td>-.45**</td>
<td></td>
</tr>
<tr>
<td>5. BPI-S – self-injury frequency</td>
<td>-.35*</td>
<td>-.28*</td>
<td>-.26</td>
<td></td>
</tr>
<tr>
<td>6. BPI-S - self-injury severity</td>
<td>-.31*</td>
<td>-.26</td>
<td>-.30*</td>
<td></td>
</tr>
<tr>
<td>7. BPI-S – aggression frequency</td>
<td>-.13</td>
<td>-.07</td>
<td>.03</td>
<td></td>
</tr>
<tr>
<td>8. BPI-S – aggression severity</td>
<td>-.21</td>
<td>-.13</td>
<td>-.06</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. LSA - instruction following</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. LSA - functional communication</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. LSA - tolerance</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>4. LSA - leisure skills</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>5. Vineland-II - challenging behaviour</td>
<td>-.10</td>
<td>.01</td>
<td>-.10</td>
<td>.05</td>
</tr>
<tr>
<td>6. BPI-S - self-injury frequency</td>
<td>-.53**</td>
<td>-.43**</td>
<td>-.43**</td>
<td>-.49**</td>
</tr>
<tr>
<td>7. BPI-S - self-injury severity</td>
<td>-.53**</td>
<td>-.44**</td>
<td>-.42**</td>
<td>-.49**</td>
</tr>
<tr>
<td>8. BPI-S - aggression frequency</td>
<td>-.47**</td>
<td>-.54**</td>
<td>-.48**</td>
<td>-.64**</td>
</tr>
<tr>
<td>9. BPI-S - aggression severity</td>
<td>-.55**</td>
<td>-.62**</td>
<td>-.52**</td>
<td>-.67**</td>
</tr>
</tbody>
</table>

*p<.05, two-tailed. **p<.01, two-tailed. LSA - Life Skills Assessment (Hanley et al., 2007)

Challenging behaviour and quality of life. Service users, with higher frequencies of self-injurious behaviour (Behavior Problems Inventory – Short Form) scored lower on the Personal Outcomes Scale domains of independence, $r_s = -.28$, $n = 59$, $p = .033$, social participation, $r_s = -.32$, $n = 59$, $p = .015$, and overall quality of life, $r_s = -.36$, $n = 59$, $p = .005$.

Service users, with higher severity of self-injurious behaviour (Behavior Problems Inventory – Short Form) scored lower on the Personal Outcomes Scale domains of independence, $r_s = -.32$, $n = 59$, $p = .014$, social participation, $r_s = -.34$, $n = 59$, $p = .015$. 108
= 59, \( p = .009 \), well-being, \( r_s = -.32, n = 59, p = .014 \), and overall quality of life, \( r_s = -.43, n = 59, p = .001 \).

Service users, with higher frequencies of aggression (Behavior Problems Inventory – Short Form) tended to score lower on the Personal Outcomes Scale domains of independence, \( r_s = -.37, n = 59, p = .004 \), social participation, \( r_s = -.30, n = 59, p = .021 \), well-being, \( r_s = -.32, n = 59, p = .014 \), and overall quality of life, \( r_s = -.42, n = 59, p = .001 \).

Service users, with higher severity of aggression (Behavior Problems Inventory – Short Form) scored lower on the Personal Outcomes Scale domains of independence, \( r_s = -.37, n = 59, p = .004 \), social participation, \( r_s = -.31, n = 59, p = .015 \), well-being, \( r_s = -.35, n = 59, p = .007 \), and overall quality of life, \( r_s = -.44, n = 59, p = .001 \).

Table 12 summarises all Spearman correlations between the service user-related measures of challenging behaviour and quality of life.

**Table 12**

<table>
<thead>
<tr>
<th>Measure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. POS - independence</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. POS - social participation</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. POS - well-being</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. POS - overall quality of life</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>5. Vineland-II- challenging</td>
<td>.05</td>
<td>-.12</td>
<td>.11</td>
<td>-.02</td>
</tr>
<tr>
<td>behaviour</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. BPI-S - self-injury frequency</td>
<td>-.28*</td>
<td>-.32*</td>
<td>-.22</td>
<td>-.36**</td>
</tr>
<tr>
<td>7. BPI-S - self-injury severity</td>
<td>-.32*</td>
<td>-.34**</td>
<td>-.32*</td>
<td>-.43**</td>
</tr>
<tr>
<td>8. BPI-S - aggression frequency</td>
<td>-.37**</td>
<td>-.30*</td>
<td>-.32*</td>
<td>-.42**</td>
</tr>
<tr>
<td>9. BPI-S - aggression severity</td>
<td>-.37**</td>
<td>-.31*</td>
<td>-.35**</td>
<td>-.44**</td>
</tr>
</tbody>
</table>

*\( p < .05 \), two-tailed.  **\( p < .01 \), two-tailed.
Adaptive functioning and quality of life. Service users, with higher levels of daily living skills (Vineland-II) scored higher on the Personal Outcomes Scale domains of independence, $r_s = .37, n = 51, p = .008$, and social participation, $r_s = .28, n = 51, p = .045$.

Service users, with higher instruction following scores (Life Skills Assessment; Hanley et al., 2007)) tended to score higher on the Personal Outcomes Scale domains of independence, $r_s = .30, n = 59, p = .022$, social participation, $r_s = .27, n = 59, p = .042$, well-being, $r_s = .38, n = 59, p = .003$, and overall quality of life, $r_s = .40, n = 59, p = .002$.

Service users, with higher functional communication scores (Life Skills Assessment; Hanley et al., 2007)) scored higher on the Personal Outcomes Scale domains of independence, $r_s = .34, n = 59, p = .009$, social participation, $r_s = .27, n = 59, p = .039$, well-being, $r_s = .40, n = 59, p = .002$, and overall quality of life, $r_s = .41, n = 59, p = .001$.

Service users reported as displaying higher levels of tolerance (Life Skills Assessment; Hanley et al., 2007)) scored higher on the Personal Outcomes Scales domains of well-being, $r_s = .36, n = 59, p = .006$, and overall quality of life, $r_s = .35, n = 59, p = .006$.

Finally, service users scoring higher in leisure skills (Life Skills Assessment; Hanley et al., 2007)) also scored higher on the Personal Outcomes Scales domains of independence, $r_s = .33, n = 59, p = .010$, social participation, $r_s = .32, n = 59, p = .012$, well-being, $r_s = .39, n = 59, p = .002$, and overall quality of life, $r_s = .44, n = 59, p = .001$. 

110
Table 13 summarises the Spearman correlations between the service user-related measures of adaptive functioning and quality of life.

Table 13

Spearman Correlations Between Service User-Related Measures of Adaptive Functioning and Quality of Life

<table>
<thead>
<tr>
<th>Measure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Vineland-II - communication</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2. Vineland-II - daily living skills</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3. Vineland-II - social skills</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4. POS - independence</td>
<td>.18</td>
<td>.37**</td>
<td>.15</td>
<td></td>
</tr>
<tr>
<td>5. POS - social participation</td>
<td>.20</td>
<td>.28*</td>
<td>.23</td>
<td></td>
</tr>
<tr>
<td>6. POS - well-being</td>
<td>-.08</td>
<td>.12</td>
<td>.01</td>
<td></td>
</tr>
<tr>
<td>7. POS - overall quality of life</td>
<td>.15</td>
<td>.28</td>
<td>.18</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ***LSA - instruction following</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2. ***LSA - functional communication</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3. ***LSA - tolerance</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4. ***LSA - leisure skills</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5. POS - independence</td>
<td>.30*</td>
<td>.34**</td>
<td>.21</td>
<td>.33**</td>
</tr>
<tr>
<td>6. POS - social participation</td>
<td>.27*</td>
<td>.27*</td>
<td>.23</td>
<td>.32*</td>
</tr>
<tr>
<td>7. POS - well-being</td>
<td>.38**</td>
<td>.40**</td>
<td>.36**</td>
<td>.39**</td>
</tr>
<tr>
<td>8. POS —overall quality of life</td>
<td>.40**</td>
<td>.41**</td>
<td>.35**</td>
<td>.44**</td>
</tr>
</tbody>
</table>

*p<.05, two-tailed. **p<.01, two-tailed. ***Life Skills Assessment (Hanley et al., 2007)

Adaptive functioning. Service users who scored higher on the Vineland-II communication domain also scored higher on Vineland-II daily living skills, $r_s = .73$, $n = 46$, $p < .001$, and socials skills domains, $r_s = .70$, $n = 46$, $p < .001$, as well as instruction following (Life Skills Assessment; Hanley et al., 2007), $r_s = .43$, $n = 50$, $p = .002$, functional communication (Life Skills Assessment; Hanley et al., 2007)), $r_s = .28$, $n = 50$, $p = .048$, and tolerance (Life Skills Assessment; Hanley et al., 2007), $r_s = .36$, $n = 50$, $p = .010$.

Service users with higher scores on the Vineland-II daily living skills domain also had higher scores on the Vineland-II social skills domain, $r_s = .68$, $n = 44$, $p <$
.001 and, instruction following (Life Skills Assessment; Hanley et al., 2007), $r_s = .40$, $n = 51, p = .005$.

Higher scores on the Vineland-II social skills domain were correlated with higher scores on instruction following (Life Skills Assessment; Hanley et al., 2007)), $r_s = .29, n = 49, p = .041$.

Finally, service users scoring higher on one area of the Life Skills Assessment (Hanley et al., 2007) also scored higher on all other areas of the Life Skills Assessment:

Higher scores for instruction following were accompanied by higher scores on functional communication, $r_s = .82, n = 59, p < .001$, tolerance, $r_s = .70, n = 59, p < .001$, and leisure skills, $r_s = .60, n = 59, p < .001$.

Higher scores for functional communication were associated with higher scores on tolerance, $r_s = .70, n = 59, p < .001$, and leisure skills, $r_s = .69, n = 59, p < .001$.

Service users scoring higher on tolerance also scored higher on leisure skills, $r_s = .66, n = 59, p < .001$.

Table 14 provides a summary of all the Spearman correlations for service user-related measures of adaptive behaviour.
Table 14

*Spearman Correlations Between Service User-Related Measures of Adaptive Behaviour*

<table>
<thead>
<tr>
<th>Measure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Vineland-II - communication</td>
<td></td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Vineland-II - daily living skills</td>
<td>.73**</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Vineland-II - social skills</td>
<td>.70**</td>
<td>.68**</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. LSA - instruction following</td>
<td>.43**</td>
<td>.40**</td>
<td>.29*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. LSA - functional communication</td>
<td>.28*</td>
<td>.25</td>
<td>.17</td>
<td>.82**</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>6. LSA - tolerance</td>
<td>.36**</td>
<td>.24</td>
<td>.16</td>
<td>.70**</td>
<td>.70**</td>
<td>-</td>
</tr>
<tr>
<td>7. LSA - leisure skills</td>
<td>.18</td>
<td>.19</td>
<td>.07</td>
<td>.60**</td>
<td>.69**</td>
<td>.66**</td>
</tr>
</tbody>
</table>

*p<.05, two-tailed. **p<.01, two-tailed. LSA - Life Skills Assessment (Hanley et al., 2007)*

**Staff-Related Measures**

**Current staff training programme.** Figure 6 presents the percentage of participants who received behaviour management and skills teaching training prior to working with service users and the percentage of participants receiving ongoing behaviour management and skills teaching training. A total of 19 (30.6%) participants received training in strategies to manage challenging behaviour and to teach service users new skills prior to starting work with service users. Less than half of the 19 participants (nine and seven, respectively) reported that these preparatory training programmes adequately equipped them for working directly with service users. Some 72.6% (N=45) of participants reported receiving ongoing behaviour management training, but only 38.7% (N=24) of those surveyed reported receiving ongoing training in how to teach new skills. The training methods experienced (e.g., instruction, role-play, discussion) varied substantially across participants, with no more than four participants reporting exposure to the same
training methods. A total of four participants (6.5%) were employed with the host organisation for less than a year, 22 (35.5%) for 1-5 years, 19 (30.6%) for 6-10 years, and 17 (27.4%) for more than 11 years.

Figure 6. The percentage of staff receiving behaviour management and skills teaching training before working with service users and the percentage of staff receiving ongoing behaviour management and skills teaching training.

The variances in Knowledge of Behaviour Management Principles Questionnaire (KBMQ) scores were equal for participants who did and did not receive training in behaviour management strategies prior to starting to work with service users, $F(1, 60) = .34, p = .56$. With equal variances assumed, on average, the Knowledge of Behaviour Management Principles Questionnaire scores for participants who received training in behaviour management strategies prior to starting to work with service users ($M = 18.84, SE = 1.16$) did not differ significantly from the Knowledge of Behaviour Management Principles Questionnaire scores for participants who did not receive training in behaviour management strategies prior to starting.
starting to work with service users ($M = 20.74, SE = 0.70$); $t(60) = -1.46, p = .15, BCa 95\% CI [-4.5, 0.70]$.

The variances in Knowledge of Behaviour Management Principles Questionnaire scores were not equal for participant who did and did not receive training in skills teaching strategies prior to starting to work with service users, $F(1, 60) = 5.5, p = .02$. With equal variances not assumed, the Knowledge of Behaviour Management Principles Questionnaire scores for participants who received training in skills teaching strategies prior to starting to work with service users ($M = 19, SE = 1.35$) did not differ significantly from the Knowledge of Behaviour Management Principles Questionnaire scores for participants who did not receive training in skills teaching strategies prior to starting to work with service users ($M = 20.67, SE = 0.63$); $t(26.18) = -1.12, p = .27, BCa 95\% CI [-4.87, 1.15]$.

The variances in Knowledge of Behaviour Management Principles Questionnaire scores were not equal for participants who did and did not receive ongoing training in behaviour management strategies, $F(1, 60) = 4.25, p = .04$. With equal variances not assumed, the Knowledge of Behaviour Management Principles Questionnaire scores for participants who received ongoing training in behaviour management strategies ($M = 20.16, SE = 0.76$) did not differ significantly from the Knowledge of Behaviour Management Principles Questionnaire scores for participants who did not receive ongoing training in behaviour management strategies ($M = 20.18, SE = 0.91$); $t(39.32) = -0.02, p = .99, BCa 95\% CI [-2.46, 2.62]$.

The variances in Knowledge of Behaviour Management Principles Questionnaire scores were unequal for participants who did and did not receive
ongoing training in skills teaching strategies, $F(1, 60) = 5.17, p = .03$. With equal variances not assumed, the Knowledge of Behaviour Management Principles Questionnaire scores for participants who received ongoing raining in skills teaching strategies ($M = 18.75, SE = 1.45$) did not differ significantly from the Knowledge of Behaviour Management Principles Questionnaire scores for participants who did not receive ongoing training in skills teaching strategies ($M = 21.05, SE = 0.64$); $t(37.31) = -1.75, p = .09$, BCa 95% CI [-4.99, 0.25].

**Psychological well-being.** Staff with higher scores on the Knowledge of Behaviour Management Principles Questionnaire reported lower scores on their own sense of achievement, as measured by the Maslach Burnout Inventory – Human Services Survey, $r_s = -.27, n = 62, p = .033$.

Staff with higher levels of job satisfaction (Minnesota Satisfaction Questionnaire – Short Form) reported lower levels of exhaustion and client depersonalisation (Maslach Burnout Inventory – Human Services Survey), $r_s = -.55$, $n = 62, p < .001$, and $r_s = -.42, n = 62, p = .001$, respectively.

Staff with higher levels of job satisfaction (Minnesota Satisfaction Questionnaire – Short Form) also reported higher levels of perceived supervisory support, $r_s = .51, n = 62, p < .001$, and lower levels of wishful coping, as measured by the Shortened Ways of Coping - Revised, $r_s = -.45, n = 62, p < .001$.

In addition to reporting lower levels of job satisfaction (Minnesota Satisfaction Questionnaire – Short Form), $r_s = -.55, n = 62, p < .001$, staff with higher levels of exhaustion displayed increased client depersonalisation (Maslach Burnout Inventory – Human Services Survey), $r_s = .74, n = 62, p < .001$, and reported using a
wishful coping style (Shortened Ways of Coping – Revised) more often, \( r_s = .66, n = 62, p < .001 \).

Similarly, staff reporting higher levels of client depersonalisation (Maslach Burnout Inventory – Human Services Survey) also scored higher on wishful coping (Shortened Ways of Coping – Revised), \( r_s = .67, n = 62, p < .001 \), and experienced reduced levels of perceived supervisory support (Perception of Supervisory Support Scale), \( r_s = -.35, n = 62, p = .005 \).

Finally, staff who reported frequent use of wishful thinking as a coping style (Shortened Ways of Coping – Revised), also reported lower levels of perceived supervisory support (Perception of Supervisory Support Scale), \( r_s = -.27, n = 62, p = .033 \), and lower levels of achievement (Maslach Burnout Inventory – Human Services Survey), \( r_s = -.29, n = 62, p = .023 \).

Table 15 summarises the Spearman correlations between the staff-related measures.

Table 15

<table>
<thead>
<tr>
<th>Measure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. KBMQ</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. MSQ-SF</td>
<td>-.12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. MBI-HSS - Exhaustion</td>
<td>.04</td>
<td>-.55**</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. MBI-HSS - Depersonalisation</td>
<td>.13</td>
<td>-.42**</td>
<td>.74**</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. MBI-HSS - Achievement</td>
<td>-.27*</td>
<td>.20</td>
<td>-.15</td>
<td>-.20</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. PSS</td>
<td>-.09</td>
<td>.51**</td>
<td>-.24</td>
<td>-.35**</td>
<td>.12</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>7. SWC-R - Wishful thinking coping style</td>
<td>.12</td>
<td>-.45**</td>
<td>.66**</td>
<td>.67**</td>
<td>-.29*</td>
<td>-.27*</td>
<td>-</td>
</tr>
<tr>
<td>8. SWC-R - Practical coping style</td>
<td>.16</td>
<td>-.07</td>
<td>.10</td>
<td>-.03</td>
<td>.11</td>
<td>.23</td>
<td>.12</td>
</tr>
</tbody>
</table>

*p<.05, two-tailed. **p<.01, two-tailed.
Staff and Service User Measures

Service user challenging behaviour and staff-related measures. Staff, whose service users showed higher scores on the Vineland-II externalised challenging behaviour domain, reported a higher sense of achievement (Maslach Burnout Inventory – Human Services Survey), \( r_s = .26, n = 59, p = .044 \).

Staff, who reported that their key service user engaged in self-injurious behaviour more frequently and with greater severity (Behaviour Problems Inventory – Short Form), also reported lower levels of client depersonalisation (Maslach Burnout Inventory – Human Services Survey), \( r_s = -.28, n = 59, p = .035 \) and \( r_s = -.32, n = 59, p = .012 \), respectively.

Staff, who reported that their key service user engaged in more frequent episodes of aggressive behaviour (Behaviour Problems Inventory – Short Form), also reported lower levels of client depersonalisation (Maslach Burnout Inventory – Human Services Survey), \( r_s = -.33, n = 59, p = .010 \).

Staff, who reported that their key service user engaged in more frequent episodes of self-injurious behaviour (Behaviour Problems Inventory – Short Form), also reported higher levels of perceived supervisory support (Perceptions of Supervisory Support Scale), \( r_s = .30, n = 59, p = .020 \).

Staff, who reported that their key service user engaged in aggressive behaviour more frequently and with greater severity (Behaviour Problems Inventory – Short Form), also reported higher levels of perceived supervisory support (Perceptions of Supervisory Support Scale), \( r_s = .32, n = 59, p = .012 \) and \( r_s = .34, n = 59, p = .009 \), respectively.
Staff, who reported that their key service user engaged in self-injurious behaviour more frequently and with greater severity (Behaviour Problems Inventory – Short Form), also reported higher levels of practical coping skills (Shortened Ways of Coping - Revised), $r_s = .36, n = 59, p = .006$ and $r_s = .27, n = 59, p = .039$, respectively.

Staff, who reported that their key service user engaged in more frequent episodes of aggressive behaviour (Behaviour Problems Inventory – Short Form), reported higher levels of practical coping skills (Shortened Ways of Coping - Revised), $r_s = .26, n = 59, p = .049$.

Table 16 summarises the Spearman correlations between the service-user challenging behaviour measures and staff-related measures.
Table 16

Spearman Correlations Between the Vineland-II Measure of Externalised Challenging Behaviour and Staff-Related Measures (Upper Panel); Spearman Correlations Between the Behaviour Problems Inventory – Short Form (BPI-S) Measure of Challenging Behaviour and Staff-Related Measures (Lower Panel)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Vineland-II – Challenging Behaviour</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. KBMQ</td>
<td>.02</td>
</tr>
<tr>
<td>2. MSQ-SF</td>
<td>.20</td>
</tr>
<tr>
<td>3. MBI-HSS - Exhaustion</td>
<td>-.06</td>
</tr>
<tr>
<td>4. MBI-HSS - Depersonalisation</td>
<td>-.05</td>
</tr>
<tr>
<td>5. MBI-HSS - Achievement</td>
<td>.26*</td>
</tr>
<tr>
<td>6. PSS</td>
<td>.13</td>
</tr>
<tr>
<td>7. SWC-R - Wishful thinking</td>
<td>-.15</td>
</tr>
<tr>
<td>8. SWC-R - Practical coping</td>
<td>.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measure</th>
<th>BPI – Frequency of Self-Injury</th>
<th>BPI – Severity of Self-Injury</th>
<th>BPI – Frequency of Aggression</th>
<th>BPI – Severity of Aggression</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. KBMQ</td>
<td>-.01</td>
<td>.05</td>
<td>.16</td>
<td>.13</td>
</tr>
<tr>
<td>2. MSQ-SF</td>
<td>.08</td>
<td>.02</td>
<td>.08</td>
<td>.02</td>
</tr>
<tr>
<td>3. MBI-HSS - Exhaustion</td>
<td>-.03</td>
<td>-.03</td>
<td>-.09</td>
<td>-.01</td>
</tr>
<tr>
<td>4. MBI-HSS - Depersonalisation</td>
<td>-.28*</td>
<td>-.32*</td>
<td>-.33**</td>
<td>-.25</td>
</tr>
<tr>
<td>5. MBI-HSS - Achievement</td>
<td>-.01</td>
<td>-.04</td>
<td>-.06</td>
<td>-.11</td>
</tr>
<tr>
<td>6. PSS</td>
<td>.30*</td>
<td>.22</td>
<td>.33*</td>
<td>.34**</td>
</tr>
<tr>
<td>7. SWC-R - Wishful thinking</td>
<td>-.02</td>
<td>-.001</td>
<td>-.12</td>
<td>-.02</td>
</tr>
<tr>
<td>8. SWC-R - Practical coping</td>
<td>.36**</td>
<td>.27*</td>
<td>.26*</td>
<td>.23</td>
</tr>
</tbody>
</table>

*p<.05, two-tailed. **p<.01, two-tailed.

Service user adaptive functioning and staff-related measures. Staff, whose service users performed better on leisure skills, as measured by the Life Skills Assessment (Hanley et al., 2007), reported a higher level of client depersonalisation (Maslach Burnout Inventory – Short Form), \( r_s = .29, n = 59, p = .026 \).
Staff, whose service users scored lower on the Life Skills Assessment (Hanley et al., 2007) measure of tolerance, reported a higher level of practical coping skills (Shortened Ways of Coping - Revised), $r_s = .27, n = 59, p = .043$.

Table 17 provides a summary of the Spearman correlation between service user adaptive functioning measures and staff-related measures.
Table 17

Spearman Correlations Between the Vineland-II Adaptive Functioning Domains and Staff-Related Measures (Upper Panel); Spearman Correlations Between the Life Skills Assessment (LSA) domains (Hanley et al., 2007) and Staff-Related Measures (Lower Panel)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Vineland-II - Communication</th>
<th>Vineland-II – Daily Living Skills</th>
<th>Vineland-II – Social Skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>KBMQ</td>
<td>.19</td>
<td>.03</td>
<td>.24</td>
</tr>
<tr>
<td>MSQ-SF</td>
<td>.02</td>
<td>.23</td>
<td>.13</td>
</tr>
<tr>
<td>MBI-HSS - Exhaustion</td>
<td>-.08</td>
<td>-.06</td>
<td>-.08</td>
</tr>
<tr>
<td>MBI-HSS - Depersonalisation</td>
<td>.09</td>
<td>.07</td>
<td>.07</td>
</tr>
<tr>
<td>MBI-HSS - Achievement</td>
<td>.01</td>
<td>.08</td>
<td>-.06</td>
</tr>
<tr>
<td>PSS</td>
<td>-.10</td>
<td>.16</td>
<td>.004</td>
</tr>
<tr>
<td>SWC-R - Wishful thinking coping</td>
<td>.001</td>
<td>.05</td>
<td>-.07</td>
</tr>
<tr>
<td>SWC-R - Practical coping</td>
<td>-.19</td>
<td>-.12</td>
<td>-.21</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measure</th>
<th>LSA - Instruction Following</th>
<th>LSA – Functional Communication</th>
<th>LSA - Tolerance</th>
<th>LSA – Leisure Skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>KBMQ</td>
<td>-.04</td>
<td>-.06</td>
<td>.08</td>
<td>-.04</td>
</tr>
<tr>
<td>MSQ-SF</td>
<td>-.03</td>
<td>.06</td>
<td>.13</td>
<td>.14</td>
</tr>
<tr>
<td>MBI-HSS - Exhaustion</td>
<td>-.01</td>
<td>-1</td>
<td>-.09</td>
<td>.02</td>
</tr>
<tr>
<td>MBI-HSS - Depersonalisation</td>
<td>.17</td>
<td>.17</td>
<td>.19</td>
<td>.29*</td>
</tr>
<tr>
<td>MBI-HSS - Achievement</td>
<td>.05</td>
<td>.13</td>
<td>-.10</td>
<td>-.02</td>
</tr>
<tr>
<td>PSS</td>
<td>-.08</td>
<td>-.12</td>
<td>-.12</td>
<td>-.14</td>
</tr>
<tr>
<td>SWC-R - Wishful thinking coping</td>
<td>-.04</td>
<td>-.07</td>
<td>-.07</td>
<td>.04</td>
</tr>
<tr>
<td>SWC-R - Practical coping</td>
<td>-.12</td>
<td>-.15</td>
<td>-.27*</td>
<td>-.19</td>
</tr>
</tbody>
</table>

*p<.05, two-tailed. **p<.01, two-tailed.

Service user quality of life and staff-related measures. Staff, whose service users scored lower on social participation, well-being and overall quality of life (Personal Outcomes Scale), scored higher on the Knowledge Behaviour Management Principles Questionnaire, \( r_s = -.27, n = 59, p = .040, r_s = -.27, n = 59, p = .038 \) and \( r_s = -.29, n = 59, p = .028 \), respectively.
Staff, whose service users scored lower on the Personal Outcomes Scale domain of independence, reported higher levels of job satisfaction (Minnesota Satisfaction Questionnaire – Short Form), $r_s = .35$, $n = 59$, $p = .006$.

Staff, whose service users scored lower on the Personal Outcomes Scale domain of independence, reported higher levels of exhaustion (Maslach Burnout Inventory – Human Services Survey), $r_s = -.28$, $n = 59$, $p = .035$.

Staff, whose service users scored lower on independence and overall quality of life (Personal Outcomes Scale), reported higher levels of achievement (Maslach Burnout Inventory – Human Services Survey), $r_s = .31$, $n = 59$, $p = .016$ and $r_s = .29$, $n = 59$, $p = .024$, respectively.

Table 18 summarises of the Spearman correlations between the service user quality of life measure and the staff-related measures.

Table 18
Spearman Correlation Results Between the Personal Outcomes Scale (POS) domains and Staff-Related Measures

<table>
<thead>
<tr>
<th>Variables</th>
<th>POS – Independence</th>
<th>POS – Social Participation</th>
<th>POS – Well-being</th>
<th>POS – Overall QOL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. KBMQ</td>
<td>-.23</td>
<td>-.27*</td>
<td>-.27*</td>
<td>-.29*</td>
</tr>
<tr>
<td>2. MSQ-SF</td>
<td>.35**</td>
<td>.04</td>
<td>.18</td>
<td>.19</td>
</tr>
<tr>
<td>3. MBI-HSS -</td>
<td>-.28*</td>
<td>.06</td>
<td>-.01</td>
<td>-.05</td>
</tr>
<tr>
<td>Exhaustion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. MBI-HSS -</td>
<td>.03</td>
<td>.12</td>
<td>.07</td>
<td>.11</td>
</tr>
<tr>
<td>Depersonalisation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. MBI-HSS -</td>
<td>.31*</td>
<td>.25</td>
<td>.26</td>
<td>.29*</td>
</tr>
<tr>
<td>Achievement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. PSS</td>
<td>.25</td>
<td>-.11</td>
<td>.02</td>
<td>-.01</td>
</tr>
<tr>
<td>7. SWC-R -</td>
<td>-.1</td>
<td>-.01</td>
<td>-.14</td>
<td>-.11</td>
</tr>
<tr>
<td>Wishful thinking coping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. SWC-R -</td>
<td>-.08</td>
<td>-.15</td>
<td>-.02</td>
<td>-.14</td>
</tr>
<tr>
<td>Practical coping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p<.05, two-tailed.  **p<.01, two-tailed.
Phase II – Practical Observations

Method

Participants. Participants were 10 frontline staff randomly selected from the sample of 62 participants who participated in Phase I of this study (i.e., 10 consent forms were blindly chosen from all completed consent forms). All participants were female.

Setting. Sessions were conducted in settings, where the participants typically supported their service users. These settings included supported accommodation, respite care, day services and resource centres.

Measures.

Discrete Trials Training Evaluation Form (DTTEF; Babel, Martin, Fazzio, Arnal, & Thomson, 2008). The DTTEF is a 21-item checklist that is used to evaluate the procedural fidelity, with which an individual implements discrete trial teaching. The measures is divided into sections, which focus on preparation for a discrete trial teaching session, management of antecedents, and management of consequences, following a correct and incorrect response. The DTTEF has been evaluated and shown to be a valid and reliable instrument for the measurement of an individual’s ability to implement discrete trial teaching (Babel et al., 2008; Jeanson et al., 2010).

Performance Diagnostic Checklist – Human Services (PDC-HS; Carr & Wilder, 2016; Carr, Wilder, Majdalany, Mathisen, & Strain, 2013). The PDC-HS is an instrument used to examine the environmental factors that may be adversely impacting on employee performance within the human services sector. This 20-item measure is divided into the following four sections, with 4-6 items per section: (1)
training; (2) task clarification and prompting; (3) resources, materials, and processes, and (4) performance consequences, effort, and competition. Thirteen of the items within the measure are covered through an interview format and seven of the items require direct observation of the work environment to be completed.

**Procedure.** Prescheduled observations, which focused on opportunities for staff to support their service users to acquire a new skill, were conducted within the natural service delivery environment for each service user. Prior to the onset of a session, the participant ranked their key service user’s ability across the nine skill domains of the Vineland-II; receptive language, expressive language, written language, personal care skills, domestic care skills, community living skills, social skills, leisure skills, and the ability to adapt to life events. The Vineland-II domain that ranked lowest was identified as being a priority for skill development for the service user.

Participants were then observed by the experimenter as they supported their service user to develop skills related to this domain. The experimenter evaluated the teaching opportunity using the Discrete Trials Training Evaluation Form. Observations lasted approximately 45-60 minutes. Following the practical observation component, the experimenter interviewed each participant using the Performance Diagnostic Checklist – Human Services. The seven items not covered in the interview were completed by the experimenter through direct observation of the work environment. Staff and service user consent for participation in this phase of the study were obtained during the consent process outlined in Phase I.

**Interobserver agreement.** Interobserver agreement (IOA) data were collected for 30% of the practical observation sessions. Two observers (the
experimenter and a trained research assistant) viewed the staff-service user interaction simultaneously and independently recorded the nature of the interaction using the Discrete Trials Training Evaluation Form. A point-by-point method was used to calculate IOA. The number of agreements was divided by the number of agreements plus disagreements. This figure was multiplied by 100 to establish the percentage IOA. The mean percentage of agreement across the target sessions was 100%.

Results

**Discrete Trials Teaching Evaluation Form (DTTEF).** Results from Phase II showed consistency across participants in determining the target task, gathering necessary materials, developing rapport with the service user, securing service user attention and delivering the correct instruction (Figures 7 & 8).

However, data indicated that nine of the 10 staff observed never encouraged the service user to preselect a reinforcer and only four participants delivered a reinforcer after correct or appropriate responding. While eight participants were observed to use prompts with service users, none predetermined the type of prompting and fading strategies they would employ. Furthermore, eight of the 10 participants failed to attempt prompt fading during their session. Finally, the error correction procedure, which is designed to address incorrect service user responding, was not implemented correctly by any of the participants.
Figure 7. The percentage of participants demonstrating target skills as outlined in the Discrete Trials Teaching Evaluation Form (DTTEF).

Figure 8. The percentage of participants demonstrating target skills as outlined in the Discrete Trials Teaching Evaluation Form (DTTEF).
Performance Diagnostic Checklist - Human Services (PDC-HS).

According to the Performance Diagnostic Checklist – Human Services, eight of the 10 staff interviewed were shown to require Behaviour Skills Training to address current skill deficits. Findings also indicate that all 10 staff require Behaviour Skills Training in either task clarification, appropriate prompting or a combination of both to perform the task and duties assigned to their role. Eight of the participants were also shown to require regular supervision and feedback to support them in their work with service users.

Phase III – Evaluation of Extant Behaviour Support Plans

Method

Participants. During Phase I, participants were asked if their key service user engaged in any form of challenging behaviour. Fifteen of the service users who displayed one or more typographies of challenging behaviour, were identified as having a behaviour support plan in place to manage this issue. These behaviour support plans were evaluated during Phase III.

Measures.

Behaviour Intervention Plan - Quality Evaluation Tool (BIP-QEII; Browning Wright, Mayer, & Saren, 2013). The BIP-QEII assesses the standard of a behaviour plan, with reference to 12 key areas; the identification of problem behaviour in observable and measurable terms, the identification of predictors of challenging behaviour, the analysis of the environmental factor supporting the problem behaviour, the specification of required environmental changes, the identification of the function of the challenging behaviour, the specification of a functionally equivalent replacement behaviour and how it will be established, the
identification of reinforcers, the identification of reactive strategies, the monitoring of progress, the co-ordination of team members, and specification of a system to facilitate communication between stakeholders. Each key area is scored on a three-point Likert scale, with ‘0’ indicating a “weak” score and ‘2’ indicating a “strong” score. Plans are then classified as “weak” ($\leq 12$ points), “underdeveloped” (13-16 points), “good” (17-21 points), and “superior” (22-24 points). The BIP-QEII has shown good inter-rater reliability, validity, and utility for audit purposes, within adult accommodation and day-support service settings (McVilly, Webber, Paris, & Sharp, 2013).

**Procedure.** A Board Certified Behaviour Analyst (BCBA®), within the RehabCare organisation, used the BIP-QEII to assess the quality of the behaviour support plans, which were in place for service users, whose key staff participated in Phase I. Administration and scoring followed the guidelines outlined in the Behaviour Intervention Plan Quality Evaluation Scoring Guide II.

**Results**

Figure 9 represents an analysis of the quality of the 15 behaviour support plans evaluated by the BIP-QEII. In total, 10 of the 15 support plans analysed were categorised as “weak” (66.7%), four as “good” (26.7%) and one was categorised as “superior”.
Figure 9. The percentage of behaviour plans rated according to the categories of the Behaviour Intervention Plan - Quality Evaluation Tool (BIP-QEII).

**Discussion**

The overall goal of the current study was to perform a training needs analysis within the host organisation to examine the specific needs of frontline staff in effectively supporting service users. The first objective was to develop an understanding of the service user profile within the RehabCare organisation. It was hypothesised that the results of this training needs analysis would highlight areas of need for service users, in terms of acquiring adaptive behavior skills, managing challenging behaviour, and achieving a suitable quality of life.

According to participant reports, the majority of service users were functioning at the lowest levels across all areas of the Vineland-II “communication”, “daily living skills”, and “social skills” domains. Furthermore, over 90% of service users were reported to be engaged in “elevated” (51.7%) or “clinically significant” (38.3%) levels of externalised challenging behaviour (Vineland-II). While it is
recognised that people with intellectual disabilities are at an increased risk for challenging behaviour (Kurtz et al., 2011; National Institute for Health and Care Excellence, 2015), these outcomes exceed the prevalence rates of 10-15% that are detailed in the literature (Emerson et al., 2001).

Previous research supports a relationship between restricted expressive and receptive communication and deficits in social skills and the tendency to display more demanding challenging behaviour (Emerson et al., 2001; Kiernan et al., 1997). Likewise, results from the current study found that the frequency and severity of aggressive and self-injurious behaviours displayed by service users were negatively correlated with their level of adaptive functioning, including communication, coping, and leisure skills. The frequency and severity of these challenging behaviours were also significantly related to quality of life scores; service users with higher levels of challenging behaviour typically scored lower on quality of life outcomes.

Earlier studies have also reported a significant predictive relationship between adaptive functioning and the quality of life of people with an intellectual disability (Maes et al., 2007; Schalock et al., 1994). These findings are supported by results from the current study; measures of adaptive functioning were positively correlated with service user quality of life outcomes. For instance, service users with poor communication skills, also scored lower on independence, social participation, well-being, and overall quality of life. As such, the training needs analysis identified substantial areas of need among service users, with regards to developing their adaptive repertoires and managing engagement in challenging behaviour. Furthermore, the study highlighted the significant relationship between adaptive functioning, challenging behaviour, and quality of life for service users in the host organisation.
The second objective of this study was to examine the training provided to the frontline staff, in the context of their role as direct-support workers for people with intellectual disabilities. Results showed that the majority of staff reported receiving no preparatory training from the host organisation, in either behaviour management strategies or skills teaching procedures, prior to beginning work with service users. In addition, more than half of the participants who had received preparatory training did not believe that it adequately equipped them to support service users within the organisation. Results from the current study also revealed that less than 40% of staff received ongoing training in how to teach their service users new skills, despite 93.5% of staff being employed with the host organisation for at least one year.

Analyses showed no relationship between the training delivered by RehabCare, in the areas of behaviour management and skills teaching and knowledge of behaviour management principles. This in turn, was reflected in the practical observations of staff working with service users in the typical support environment; frontline staff appeared to be unaware of evidence-based practices, such as systematic prompting and reinforcement procedures, when supporting their service users to acquire new skills. These knowledge and skill deficits may be a result of poor guidance from specialist teams; over 66% of the behaviour support plans analysed were categorised as “weak” (Behaviour Intervention Plan – Quality Evaluation Tool II), or the lack of consistency in training structure and delivery across the organisation; no more than four participants reported experiencing the same training methods. As such, the training needs analysis demonstrated substantial inconsistency, in terms of training provision to frontline staff, the methods of
dissemination employed, and the development of relevant knowledge and competencies.

The final objective of the study was to investigate the factors associated with burnout among frontline staff in the host organisation. As predicted, results from the current study showed a significant relationship between an emotionally-based ‘wishful thinking’ style and the three dimensions of burnout. ‘Wishful thinking’ was positively correlated with emotional exhaustion and client depersonalisation and negatively correlated with participants’ sense of achievement. This is consistent with recent findings reported by Griffith et al. (2014) and earlier findings reported by Devereux, Hastings, Noone, et al. (2009). In addition, a reliance on ‘wishful thinking’ was associated with lower levels of job satisfaction, which in turn was negatively correlated with emotional exhaustion and client depersonalisation.

Contrary to previous research (Kozak et al., 2013; Smyth et al., 2015), participants who reported working with more demanding service users were less likely to experience symptoms of burnout. For example, participants exposed to elevated levels of service user engagement in self-injury and aggression, reported significantly lower levels of client depersonalisation, whereas staff working with service users with better leisure skills were more likely to engage in depersonalisation. Further analyses revealed that staff, who reported that their key service users were more likely to display self-injurious and aggressive behaviours, also reported higher levels of perceived supervisory support and a reliance on practical coping styles. Consequently, it is possible that the practical coping styles and supervisory support available to staff assigned to work with challenging service users may have served as protective factors against client depersonalisation.
The cognitive-behavioural approach to stress theorises that coping style is a function of how an individual interprets an environmental threat, coupled with their perception of the resources available to them to manage that threat (Lazarus, 1999). Therefore, training which addresses at least one component of this two-step process (i.e., the resources available to an individual) may impact an individual’s coping style. Research has previously reported that the responses of frontline staff to work-related demands are a function of the standard of training received (Reed & Henley, 2015). Furthermore, results from the present study show a proclivity for participants who report higher levels of supervisory support and who demonstrated superior knowledge of behaviour management strategies, to score lower on measures of ‘wishful thinking’. Therefore, with widespread organisational training and guidance in evidence-based practices, which assist staff to manage challenging behaviour and promote service user learning, frontline staff should be better equipped with the practical skills necessary to address work-related demands as they arise.

Contrary to expectations, current analyses also revealed a negative relationship between participants’ knowledge of behaviour management approaches and service user social participation, well-being, and overall quality of life. It is possible that frontline staff with superior knowledge and expertise are assigned as key workers to service users with higher levels of dependency and need. However, with the potential for burnout and lower job satisfaction (Hensel et al., 2012; Robertson et al., 2005; Rose & Rose, 2005), the host organisation must ensure that the majority of its workforce are adequately and effectively trained to support the profiles of service users availing of its services.

One weakness in the present study relates to the measurement of service user quality of life. When measuring quality of life, both the subjective experience of an
individual and objective indicators of life circumstances should be evaluated (Verdugo, Schalock, Keith, & Stancliffe, 2005). However, a dual assessment was not possible in this study due to restrictions around accessing service users and concerns over the validity of using proxy data as a substitute for self-reports (Verdugo et al., 2005). Therefore, the Personal Outcomes Scale, which provides a researcher-led, objective evaluation of a service user’s life circumstances, as directly observed by a professional or caregiver, who lives or works with the individual (Van Loon et al., 2009) was employed in this instance. Although the omission of a subjective evaluation of service user quality of life was a limitation of the present study, research has shown that objective indicators are recommended when the objective is to appraise service programming and delivery (Verdugo et al., 2005).

A second limitation to the study was the requirement to use nonparametric statistics to analyse the relationships between variables in this study. As already mentioned, there is much controversy within the statistical literature with regards to analysing Likert scale data (Jamieson, 2004; Lubke & Muthén, 2004). According to one perspective, it is possible to find true parameter values with Likert data, under particular circumstances (e.g., assumptions of normality, equality of variance) and with specific types of analyses (e.g., factor analysis) (Lubke & Muthén, 2004). However, this position has not been clearly and consistently supported across the statistical analyses used in this study (Jamieson, 2004) and as a result, it was decided to adopt a conservative and cautious approach, by employing non-parametric tests in the current study. As a result, the likelihood of detecting significant relationships between variables was compromised (Gravetter & Wallnau, 2004). However, given that the majority of variables measured used the Likert-scale format, utilising the non-parametric alternatives increased confidence in the results.
In conclusion, the goal of the current study was to conduct a training needs analysis within the host organisation to identify the exact needs of staff in effectively supporting service users. The information yielded from this analysis would be used in the design of a staff training programme to address deficient skills in selecting and applying evidence-based practice for persons with intellectual disability. Based on the results of this training needs analysis, communication skills, social skills, and daily living skills were selected as priority areas for service users within the host organisation. Given that applied behaviour analytic research has developed a range of empirically tested interventions that target deficits in these areas (Carr et al., 2009; Grey & Hastings, 2005; Kurtz et al., 2011; National Autism Center, 2009; 2015; Wong et al., 2014), it was hypothesised that staff training in these evidence-based practices would assist frontline staff in supporting service users to achieve a quality of life commensurate with their peers. Therefore, the next study sought to identify specific evidence-based practices that promote the development of communication, social, and daily living skills among people with intellectual disabilities.
Chapter 7

An Evaluation of the Evidence-Base Supporting the Use of Specific Behavioural Interventions with Individuals with Intellectual and Developmental Disabilities:

A Series of Systematic Literature Reviews
Abstract

The quality of life of people with intellectual disabilities is significantly impacted by their levels of adaptive functioning and whether or not they engage in challenging behaviour (Schalock et al., 1994). Based on the results of the training needs analysis reported in Chapter 6, the development of communication, social, and daily living skills was identified as a priority for service users within the host organisation. Consequently, the objective of the current study was to identify evidence-based practices, which effectively address deficits in these skill areas. Using the autism-specific literature as a guide (i.e., National Autism Center, 2015; Wong et al., 2014), a series of systematic reviews were conducted to evaluate the evidence-base supporting the use of the following behavioural interventions with people with intellectual and developmental disabilities: reinforcement, systematic prompting, functional communication training, and task analysis. Results showed a substantial evidence-base supporting the use of these behavioural interventions with people with intellectual and developmental disabilities. Given these findings, the subsequent intervention employed reinforcement, systematic prompting, functional communication training, and task analysis (including chaining) within a Behavioural Skills Training programme with frontline staff.
Chapter 7

The quality of life of people with intellectual disabilities is significantly impacted by their levels of adaptive functioning and whether or not they engage in challenging behaviour (Schalock et al., 1994). However, with suitable supports that are grounded in evidence-based practice, adaptive functioning can improve, resulting in reduced instances of challenging behaviour (Grey & Hastings, 2005; Kiernan et al., 1997; Kurtz et al., 2011), increased independence, and improved quality of life outcomes (Brown et al., 2009; Dollar et al., 2012; Maes et al., 2007; National Institute for Health and Care Excellence, 2015).

Although substantial progress has been made in identifying evidence-based interventions (National Autism Center, 2009; 2015; Wong et al., 2014), Chambless and Hollon (1998) caution that “projects for which a treatment manual was not written… are of limited utility in terms of assessment of treatment efficacy” (p.11). Manualisation provides practitioners with relevant information on how to effectively apply evidence-based practice according to specified parameters (Chambless & Ollendick, 2001). Within the general field of Psychology, however, there has been strong opposition to the manualisation of treatment procedures, as it is considered by many to pose a threat to independent practice and innovation (Chambless & Ollendick, 2001). Nevertheless, research has shown that flexible treatment approaches do not produce statistically more effective therapeutic outcomes than manualised interventions and in certain cases the psychotherapeutic benefits are far superior with manualisation (Chambless & Ollendick, 2001). In addition, the scientific framework for the implementation of applied behaviour analytic practice, specifies that interventions and treatments must be conceptually systematic and described with adequate detail to allow for replication (Baer et al., 1968; 1987).
Based on the results of the training needs analysis reported in Chapter 6, communication skills, social skills, and daily living skills were selected as priority areas for service users within the host organisation. As discussed in Chapter 3, the National Autism Centre and the National Professional Development Centre independently conducted two large-scale evaluations of interventions for children, adolescents, and adults with autism (National Autism Center, 2009; 2015; Wong et al., 2014). According to the results reported by these projects, the overwhelming majority of interventions, identified as well-established, were generated from behaviour analytic research.

An analysis of these findings showed that reinforcement, systematic prompting, functional communication training, and task analysis are considered established evidence-based practices for addressing deficits in communication, social and daily living skills among people with autism (National Autism Center, 2009; 2015; Wong et al., 2014). In addition, the National Professional Development Centre, in collaboration with the University of North Carolina, created online manuals for each of the 27 evidence-based practices identified in their review, including reinforcement, systematic prompting, functional communication training, and task analysis (AFIRM Team, 2015a; 2015b; 2015c; Franzone, 2009; Wong et al., 2014). These manuals outline the procedures involved in planning for, implementing and evaluating the impact of each evidence-based practice (AFIRM Team, 2015a; 2015b; 2015c; Franzone, 2009).

Reinforcement is fundamental to the implementation of many other evidence-based strategies, including task analysis and functional communication training (Carr & Durand, 1985; Cooper et al., 2014; Franzone, 2009; Kurtz et al., 2011; Neitzel, 2009; Okinaka & Shimazaki, 2011). It describes a process whereby an event, which
occurs after the display of a behaviour, increases the likelihood of that behaviour occurring again in the future (Cooper et al., 2014). Similarly, systematic prompting is most effectively used when combined with other evidence-based practices. It involves assisting an individual to perform a target behaviour and systematically fading this assistance over time (AFIRM Team, 2015a; Nietzel & Wolery, 2009).

This approach has been used to support people with intellectual disabilities to learn a wide range of new skills. For example, Bosner and Belfiore (2001) employed least-to-most prompting methods to teach a 16 year old girl with an intellectual disability to self-administer insulin. Similarly, Cavkaytar (2012) implemented a programme, based on least-to-most prompting strategies, to successfully train three adult men with intellectual disabilities as café waiters.

Functional communication training is a method, which replaces an inappropriate communication behaviour with a functionally equivalent, socially acceptable alternative response (Kurtz et al., 2011). This approach typically employs reinforcement and systematic prompting (Cooper et al., 2014; Franzone, 2009; Skinner, 1957) and the literature repeatedly shows this intervention to be effective in reducing challenging behaviour among children and adults with intellectual disabilities (Kurtz et al., 2011). Finally, task analysis is a procedure, which breaks complex, chained behaviours (e.g., making a cup of tea) into smaller teachable steps. Each individual step is then taught systematically, usually by employing reinforcement and systematic prompting techniques, until the service user is eventually performing the larger skill independently (AFIRM Team, 2015c). Chan, Lambdin, Graham, Fragale, and Davis (2014) combined task analysis and systematic prompting procedures to teach three adults with intellectual disabilities (aged 33-57 years) to engage appropriately with an iPad, while Veazey, Valentino, Low,
McElroy, and LeBlanc (2016) used reinforcement, least-to-most prompting and task analysis techniques to teach two girls, with intellectual and developmental disabilities, female hygiene skills.

It must be noted that results from the National Standards Project (National Autism Center, 2009; 2015) and the National Professional Development Centre report (Wong et al., 2014) were specific to people with autism; a similar large-scale analysis of interventions for children and adults with intellectual disabilities does not yet exist. However, intellectual disability frequently co-occurs with autism (American Association on Intellectual and Developmental Disabilities, 2017; Matson & Shoemaker, 2009); 50-70% of individuals with autism are also diagnosed with a co-occurring intellectual disability (Matson & Shoemaker, 2009). Given this high level of co-morbidity, it was decided to use findings from the National Standards Project (National Autism Center, 2009; 2015) and the National Professional Development Centre report (Wong et al., 2014) to guide investigations into the evidence-base supporting the use of reinforcement, systematic prompting, functional communication training, and task analysis with the intellectual and developmental disability population.

**Aim of Current Study**

Personal ability is the most influential predictor of quality of life for people with intellectual and developmental disabilities (Maes et al., 2007; Schalock et al., 1994). In an empirical analysis of 16 studies, Maes et al. (2007) found that for people with an intellectual and developmental disability, a higher level of deficits in the adaptive behaviour repertoire was associated with achieving less personal outcomes than more able peers. Therefore, it seems imperative to identify evidence-based practices that effectively promote adaptive skill acquisition among this group.
Given that the results of the training needs analysis reported in Chapter 6 identified the development of communication, social and daily living skills as a priority for service users within the host organisation, the objective of the current study was to identify evidence-based practices, which effectively address deficits in these skill areas. Results from the National Standards Project (National Autism Center, 2009; 2015) and the National Development Centre report (Wong et al., 2014) were used to direct a series of systematic literature reviews, which evaluated the evidence-base supporting the use of reinforcement, systematic prompting, functional communication training, and task analysis with people with intellectual and developmental disabilities.

**Method**

**Search Procedures**

Comprehensive literature searches were carried out across the following electronic databases: PsycINFO, PsycArticles, PubMed and ERIC.

The first series of systematic searches were conducted by combining the search terms assist*, guidance, prompt* and direct teaching with the following keywords: intellectual disabil*, mental retardation and developmental disabil*.

The second series of systematic searches were conducted by combining the search terms task analysis and chain* with the following keywords: intellectual disabil*, mental retardation and developmental disabil*.

The final searches were conducted by combining the search terms functional communication training, functional communication and alternative communication with the following keywords: intellectual disabil*, mental retardation and developmental disabil*. 
Inclusion and Exclusion Criteria. Searches were limited by year of publication (1985-2016), with studies prior to 1985 being excluded from the review (Kurtz et al., 2011). In addition, only those articles, which were written in the English language, involved original research and published in peer-reviewed journals, were considered for inclusion.

Systematic prompting review. Studies included in this review detailed systematic prompting procedures, which were used to: a) teach individuals with intellectual and developmental disabilities a new skill; or b) increase the display of an existing behaviour among people with intellectual and developmental disabilities. These prompting procedures must have been implemented by another person. As such, studies that limited their investigations to self-generated, technology-based and/or video generated prompting only were excluded. Articles, which focused exclusively on participants who had a physical developmental disability only were also excluded from the review. In addition, studies that used prompting strategies to primarily reduce behaviour were also excluded.

Our search produced 243 articles. After filtering them (this process is presented in Figure 10) according to our inclusion/exclusion criteria, all 243 studies remained.
Figure 10. Flow diagram showing inclusion/exclusion of studies identified during database search process for systematic prompting procedures.

**Task analysis review.** Studies included in this review used task analysis and chaining procedures to: a) teach individuals with intellectual and developmental disabilities a new skill; or b) increase the display of an existing behaviour among people with intellectual and developmental disabilities. Articles, which focused exclusively on participants who had a physical developmental disability only were also excluded from the review. Our search produced 95 articles. After filtering them (this process is presented in Figure 11) according to our inclusion/exclusion criteria, all 95 studies remained.
Figure 11. Flow diagram showing inclusion/exclusion of studies identified during database search process for task analysis and chaining procedures.

**Functional communication training review.** Studies included in this review used functional communication training to teach individuals with intellectual and developmental disabilities a new communication response, which was functionally equivalent to an inappropriate/challenging behaviour that the individual was already displaying. Articles, which focused exclusively on participants who had a physical developmental disability only were also excluded from the review. Our search produced 71 articles. After filtering them (this process is presented in Figure 12) according to our inclusion/exclusion criteria, all 71 studies remained. Figure 12 presents a flow diagram showing the inclusion/exclusion of studies identified during the database search process.
Figure 12. Flow diagram showing inclusion/exclusion of studies identified during database search process for functional communication training.

**Data Extraction**

For each included study across all reviews, data were extracted on the following variables: (1) participant characteristics; (2) target skill for intervention; type of procedure employed, (3) additional strategies used; (4) treatment efficacy, and (5) the Scientific Merit Rating. Given that the majority of studies reviewed employed single-subject experimental design and visual analysis of results, the Percentage of Non-Overlapping Data Points (PND) was used to quantify treatment efficacy. According to this method, the number of intervention points that surpass the highest/lowest baseline point are counted and this number is subsequently represented as a percentage of the total number of intervention points. It is
recommended that if a reversal design is used, the first baseline and intervention data set should be analysed (Olive & Franco, 2008; Scruggs & Mastropieri, 1998). The PND rating system is based on a four-point scale; highly effective (90-100%), moderately effective (70-89%); minimally effective (50-69%), and ineffective (0-49%).

Among the advantages to this method are ease of calculation and interpretation. Furthermore, this method has been consistently shown to correlate well with visual analysis judgements during meta-analysis studies (Parker, Vannest, & Davis, 2011) and is a more conservative approach to assessing treatment efficacy than other methods, such as the Percentage of Data Exceeding the Median (PEM) (Mark, Matthew, Brian, & Erin, 2008). However, PND calculations are sensitive to floor or ceiling baseline data points and the number of data points recorded during intervention (Olive & Franco, 2008). Therefore it is advised to interpret any PND values that were calculated with baseline points that fell at the floor or ceiling (i.e., automatically producing a zero PND rating) with caution (Scruggs & Mastropieri, 1998) and this recommendation was followed in the current reviews.

As mentioned above, the Scientific Merit Rating for each study was also analysed. The Scientific Merit Rating Scale (SMRS) was developed by the National Standards Project to objectively evaluate the scientific methods used in the studies that they assessed (National Autism Center, 2015). The SMRS appraises each study across “five critical dimensions of experimental rigor” (p. 25): (1) research design; (2) measurement of the dependent variable; (3) measurement of the independent variable; (4) participant ascertainment, and (5) generalisation and maintenance (National Autism Center, 2015). Each of the five dimensions of scientific merit are scored on a six-point scale (0-5). A score of ‘0’ represents the weakest score.
possible and a score of ‘5’ represents the strongest score possible. The criteria for each rating on the SMRS are detailed in Appendix B. A composite score is then calculated using the individual scores across the five dimensions. The designated formula is outlined in Appendix B. SMRS composite scores of three or above signify adequate experimental rigor. Scores of two or below indicate that more scientific control is required (National Autism Center, 2015).

**Evaluation of Studies**

According to the National Standards Report’s “Strength of Evidence Classification System”, an established intervention will have a minimum of two group or four single-subject experimental design studies with at least 12 participants and no conflicting results or three group or six single-subject experimental design studies with at least 18 participants and less than 10% of studies reporting conflicting results. Furthermore, these studies should have an SMRS composite score of 3, 4 or 5 and demonstrate a beneficial effect (National Autism Center, 2015). Across all reviews in this study, the beneficial effect was quantified using the PND method described above. Only those studies, for which a PND rating could be calculated, were analysed for potential classification as an “established” intervention, using the Strength of Evidence Classification System. Table 19 presents criteria for “established” interventions, according to the Strength of Evidence Classification System.
Table 19

*Strength of Evidence Classification System – Criteria for “Established” Interventions*

<table>
<thead>
<tr>
<th>Criteria for Established Interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Studies must score 3, 4 or 5 on the SMRS</td>
</tr>
<tr>
<td>• Two group or four single-subject experimental design studies with at least 12 participants and no conflicting results or Three group or six single-subject experimental design studies with at least 18 participants and less than 10% of studies reporting conflicting results</td>
</tr>
<tr>
<td>• Beneficial effects from intervention must be demonstrated (quantified by PND rating in current study)</td>
</tr>
</tbody>
</table>

**Interrater Reliability**

For each review, the accuracy of the SMRS dimension ratings and the PND ratings was evaluated by having a second rater review randomly selected studies. A second rater reviewed 55 (22.6%) studies from the systematic prompting review. Inter-rater reliability for the SMRS dimension ratings was found to be 93.1% and inter-rater reliability for the PND ratings was found to be 91.5%. A second rater also reviewed 26 (27.4%) studies from the task analysis review. Inter-rater reliability for the SMRS dimension ratings was found to be 97.7% and inter-rater reliability for the PND ratings was found to be 95.8%. Finally, a second rater reviewed 21 (29.6%) studies from the task analysis review. Inter-rater reliability for the SMRS dimension ratings was found to be 100% and inter-rater reliability for the PND ratings was found to be 100%.

**Results**

**Systematic Prompting Review**

A total of 243 articles, published between the years 1985-2015, met the inclusion criteria (Appendix C – Table 20). The data extracted across all included
studies were summarised and are available from the author on request. Table 21 lists all studies, which scored 3 or more on the SMRS.

Table 21

*Systematic Prompting Studies Scoring at Least an SMRS Score of 3*

<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ahlgrim-Delzell, Browder, and Wood (2014)</td>
<td>Children</td>
</tr>
<tr>
<td>Ahlgrim-Delzell et al. (2016)</td>
<td>Children</td>
</tr>
<tr>
<td>Akçin (2013)</td>
<td>Children</td>
</tr>
<tr>
<td>Akmanoglu and Batu (2004)</td>
<td>Children</td>
</tr>
<tr>
<td>Akmanoglu and Tekin-Iftar (2011)</td>
<td>Children</td>
</tr>
<tr>
<td>Akmanoglu, Yanardag, and Batu (2014)</td>
<td>Children</td>
</tr>
<tr>
<td>Akmanoglu-Uludag and Batu (2005)</td>
<td>Children</td>
</tr>
<tr>
<td>Alberto, Cihak, and Gama (2005)</td>
<td>Children</td>
</tr>
<tr>
<td>Alcantara (1994)</td>
<td>Children</td>
</tr>
<tr>
<td>Allen, Burke, Howard, Wallace, and Bowen (2012)</td>
<td>Children</td>
</tr>
<tr>
<td>Ault, Gast, and Wolery (1988)</td>
<td>Children</td>
</tr>
<tr>
<td>Bannerman, Sheldon, and Sherman (1991)</td>
<td>Adults</td>
</tr>
<tr>
<td>Barbeta, Heward, and Bradley (1993)</td>
<td>Children</td>
</tr>
<tr>
<td>Batu (2008)</td>
<td>Children</td>
</tr>
<tr>
<td>Batu, Ergenekon, Erbas, and Akmanoglu (2004)</td>
<td>Children</td>
</tr>
<tr>
<td>Bell, Young, Salzberg, and West (1991)</td>
<td>Children</td>
</tr>
<tr>
<td>Bennett, Gast, Wolery, and Schuster (1986)</td>
<td>Children</td>
</tr>
<tr>
<td>Bennett, Reichow, and Wolery (2011)</td>
<td>Children</td>
</tr>
<tr>
<td>Bevill, Gast, Maguire, and Vail (2001)</td>
<td>Children</td>
</tr>
<tr>
<td>Birkan (2005)</td>
<td>Children</td>
</tr>
<tr>
<td>Browder, Hudson, and Wood (2013)</td>
<td>Children</td>
</tr>
<tr>
<td>Cavallaro and Poulson (1985)</td>
<td>Children</td>
</tr>
<tr>
<td>Cavkaytar (2012)</td>
<td>Children</td>
</tr>
<tr>
<td>Çelik and Vuran (2014)</td>
<td>Children</td>
</tr>
<tr>
<td>Chan et al. (2014)</td>
<td>Adults</td>
</tr>
<tr>
<td>Choi, O'Reilly, Sigafoos, and Lancioni (2010)</td>
<td>Children</td>
</tr>
<tr>
<td>Cihak, Alberto, Taber-Doughty, and Gama (2006)</td>
<td>Children</td>
</tr>
<tr>
<td>Cipani (1985)</td>
<td>Children</td>
</tr>
<tr>
<td>Coe, Matson, Craigie, and Gossen (1991)</td>
<td>Children</td>
</tr>
<tr>
<td>Coleman, Cherry, Moore, Yujeong, and Cihak (2015)</td>
<td>Children</td>
</tr>
<tr>
<td>Collins, Hall, and Branson (1997)</td>
<td>Children</td>
</tr>
<tr>
<td>Colyer and Collins (1996)</td>
<td>Children</td>
</tr>
<tr>
<td>Cooper and Browder (1997)</td>
<td>Adults</td>
</tr>
<tr>
<td>Copeland and Hughes (2000)</td>
<td>Children</td>
</tr>
<tr>
<td>Creech-Galloway, Collins, Knight, and Bausch (2013)</td>
<td>Children</td>
</tr>
<tr>
<td>Source</td>
<td>Group</td>
</tr>
<tr>
<td>-----------------------------------------------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Day (1987)</td>
<td>Children</td>
</tr>
<tr>
<td>Devine, Malley, Sheldon, Dattilo, and Gast (1997)</td>
<td>Adults</td>
</tr>
<tr>
<td>Dollar et al. (2012)</td>
<td>Adults</td>
</tr>
<tr>
<td>Dorminy, Luscre, and Gast (2009)</td>
<td>Children</td>
</tr>
<tr>
<td>Downing (1987)</td>
<td>Children</td>
</tr>
<tr>
<td>Duker, Dortmans, and Lodder (1993)</td>
<td>Children and adults</td>
</tr>
<tr>
<td>Duker and Jutten (1997)</td>
<td>Adults</td>
</tr>
<tr>
<td>Emmick, Cihon, and Eshleman (2010)</td>
<td>Children</td>
</tr>
<tr>
<td>Ergenekon, Tekin-Iftar, Kapan, and Akmanoglu (2014)</td>
<td>Children</td>
</tr>
<tr>
<td>Fetko, Schuster, Harley, and Collins (1999)</td>
<td>Children</td>
</tr>
<tr>
<td>Fox, Shores, Lindeman, and Strain (1986)</td>
<td>Children</td>
</tr>
<tr>
<td>Gee, Graham, Goetz, Oshima, and Yoshioka (1991)</td>
<td>Children</td>
</tr>
<tr>
<td>Gibson and Schuster (1992)</td>
<td>Children</td>
</tr>
<tr>
<td>Gilson and Carter (2016)</td>
<td>Children</td>
</tr>
<tr>
<td>Godby, Gast, and Wolery (1987)</td>
<td>Children</td>
</tr>
<tr>
<td>Gruber and Poulson (2016)</td>
<td>Children</td>
</tr>
<tr>
<td>Halasz-Dees and Cuvo (1986)</td>
<td>Children and adults</td>
</tr>
<tr>
<td>Hartzell, Liauspin, Gann, and Clem (2015)</td>
<td>Children</td>
</tr>
<tr>
<td>Heinrich, Collins, Knight, and Spriggs (2016)</td>
<td>Children</td>
</tr>
<tr>
<td>Hoogeveen, Kouwenhoven, and Smee (1989)</td>
<td>Children</td>
</tr>
<tr>
<td>Hudson and Browder (2014)</td>
<td>Children</td>
</tr>
<tr>
<td>Hudson, Browder, and Jimenez (2014)</td>
<td>Children</td>
</tr>
<tr>
<td>Ivancic and Schepis (1995)</td>
<td>Adults</td>
</tr>
<tr>
<td>James and Egel (1986)</td>
<td>Children</td>
</tr>
<tr>
<td>Kagoohara et al. (2011)</td>
<td>Children</td>
</tr>
<tr>
<td>Karl, Collins, Hager, and Ault (2013)</td>
<td>Children</td>
</tr>
<tr>
<td>Knudson et al. (2009)</td>
<td>Adults</td>
</tr>
<tr>
<td>Kouri (2005)</td>
<td>Children</td>
</tr>
<tr>
<td>Krstovska-Guerrero and Jones (2016)</td>
<td>Children</td>
</tr>
<tr>
<td>Kryzak and Jones (2015)</td>
<td>Children</td>
</tr>
<tr>
<td>Lalli, Mace, Browder, and Brown (1989)</td>
<td>Adults</td>
</tr>
<tr>
<td>Leaf, Leaf, Taubman, McEachin, and Delmolino (2014)</td>
<td>Children</td>
</tr>
<tr>
<td>Leaf et al. (2011)</td>
<td>Children</td>
</tr>
<tr>
<td>Lorah, Karnes, and Speight (2015)</td>
<td>Children</td>
</tr>
<tr>
<td>Luciano, Barnes-Holmes, and Barnes-Holmes (2002)</td>
<td>Children</td>
</tr>
<tr>
<td>MacDuff, Krantz, and McClannahan (1993)</td>
<td>Children</td>
</tr>
<tr>
<td>Maeser and Thyer (1990)</td>
<td>Children</td>
</tr>
<tr>
<td>McAdam and Cuvo (1994)</td>
<td>Children</td>
</tr>
<tr>
<td>McDonnell (1987)</td>
<td>Children</td>
</tr>
<tr>
<td>McDonnell and Ferguson (1989)</td>
<td>Children</td>
</tr>
</tbody>
</table>
McKay, Weiss, Dickson, and Ahearn (2014)  
McLaughlin and Walsh (1996)  
Mechling and Gustafson (2009a)  
Mechling and Gustafson (2009b)  
Mechling and Stephens (2009)  
Miller and Test (1989)  
Minarovic and Bambara (2007)  
Newman and Eyck (2005)  
O’Handley, Dadakhodjaeva, Radley, and Dart (2016)  
Ozen (2008)  
Parker and Schuster (2002)  
Pennington, Delano, and Scott (2014)  
Pennington, Collins, Stenhoff, Turner, and Gunselman (2014)  
Pennington, Stenhoff, Gibson, and Ballou (2012)  
Phillips and Vollmer (2012)  
Rao and Kane (2009)  
Rao and Mallow (2009)  
Roy-Wsiaki, Marion, Martin, and Yu (2010)  
Schuster and Griffen (1993)  
Seward, Schuster, Ault, Collins, and Hall (2014)  
Shimizu and McDonough (2006)  
Shimizu, Yoon, and McDonough (2010)  
Singleton, Schuster, Morse, and Collins (1999)  
Sisson, Kilwein, and Van Hasselt (1988)  
Smeets (1992)  
Smeets, Lancioni, and Striefel (1987)  
Smith, Schuster, Collins, and Kleinert (2011)  
Sowers, Verdi, Bourbeau, and Sheehan (1985)  
Spriggs, Gast, and Ayres (2007)  
Stokes, Cameron, Dorsey, and Fleming (2004)  
Strand and Morris (1986)  
Strand and Morris (1988)  
Strasberger and Ferreri (2014)  
Swain, Lane, and Gast (2015)  
Swenson-Pierce, Kohl, and Egel (1987)  
Tam, Phillips, and Mudford (2011)  
Taras, Matson, and Felps (1993)  
Tarnowski and Drabman (1987)  
Tekin-Iftar (2008)  
Tekin-Iftar, Kurt, and Acar (2008)
In total 125 of the studies in Table 21 investigated the use of systematic prompting with children (≤22 years), 14 studies assessed the impact with adult participants (>22 years), and five studies involved both children and adults.

According to the Strength of Evidence Classification System (National Autism Center, 2015), systematic prompting can be categorised as an “established” intervention for children with intellectual and developmental disabilities. A total of 25 studies using simultaneous prompting strategies, 25 studies employing either least-to-most prompting procedures (n=21) or least-to-most prompting combined with another prompting techniques (n=4), and seven studies implementing most-to-least prompting (SMRS >3), produced at least a minimally effective outcome (PND) for participants aged 22 years or younger. The participant requirement (i.e., at least four studies with 12 or more participants) was met in all cases. There were no conflicting results (Table 22). Although four studies demonstrated at least minimally effective outcomes, following the use of graduated guidance, only nine individuals participated across these four studies. Therefore graduated guidance would be considered an “emerging” intervention (National Autism Center, 2015) (Table 22).
Table 22

*Total Number of Studies for which a PND Rating could be Calculated, Total Number of Studies with at Least a Minimally Effective PND and the Status of the Target Interventions for a Population of Children with Intellectual and Developmental Disabilities*

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Number of Studies with PND Rating</th>
<th>Number of Studies with at Least a Minimally Effective Outcome (PND)</th>
<th>Established Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Least-to-most prompting</td>
<td>21</td>
<td>21</td>
<td>Yes</td>
</tr>
<tr>
<td>Least-to-most prompting + other stimulus/response prompting procedures</td>
<td>4</td>
<td>4</td>
<td>Yes</td>
</tr>
<tr>
<td>Most-to-least prompting</td>
<td>7</td>
<td>7</td>
<td>Yes</td>
</tr>
<tr>
<td>Graduated Guidance</td>
<td>4</td>
<td>4</td>
<td>No</td>
</tr>
<tr>
<td>Graduated guidance + video modelling</td>
<td>3</td>
<td>3</td>
<td>No</td>
</tr>
<tr>
<td>Simultaneous prompting</td>
<td>25</td>
<td>25</td>
<td>Yes</td>
</tr>
</tbody>
</table>

The results of the review on studies employing systematic prompting also showed that the least-to-most prompting procedure can be characterised as an “emerging” intervention for adults with intellectual and developmental disabilities. A total of four studies using least-to-most prompting (SMRS ≥3), produced at least a minimally effective outcome (PND) for adult participants (Table 23). However, there was conflicting results observed in a fifth study. Therefore, least-to-most prompting did not meet the criteria for an “established” intervention.
Table 23

Total Number of Studies for which a PND Rating could be Calculated, Total Number of Studies with at Least a Minimally Effective PND and the Status of the Target Interventions for a Population of Adults with Intellectual and Developmental Disabilities

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Number of Studies with PND Rating</th>
<th>Number of Studies with at Least a Minimally Effective Outcome (PND)</th>
<th>Established Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Least-to-most prompting</td>
<td>5</td>
<td>4</td>
<td>No</td>
</tr>
<tr>
<td>Least-to-most prompting</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Most-to-least prompting</td>
<td>2</td>
<td>2</td>
<td>No</td>
</tr>
<tr>
<td>Most-to-least prompting</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graduated Guidance</td>
<td>0</td>
<td>0</td>
<td>No</td>
</tr>
<tr>
<td>Graduated guidance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simultaneous prompting</td>
<td>2</td>
<td>2</td>
<td>No</td>
</tr>
<tr>
<td>Simultaneous prompting</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Task Analysis Review

A total of 95 articles, published between the years 1985-2015, met the inclusion criteria (Appendix C – Table 24). The data extracted across all included studies were summarised and are available from the author on request. Table 25 lists all studies, which scored 3 or more on the SMRS.
Table 25

*Task Analysis Studies Scoring at Least an SMRS Score of 3*

<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ayres, Maguire, and Mcclimon (2009)</td>
<td>Children</td>
</tr>
<tr>
<td>Browder, Lim, Lin, and Belfiore (1993)</td>
<td>Adults</td>
</tr>
<tr>
<td>Browder, Snell, and Wildonger (1988)</td>
<td>Children</td>
</tr>
<tr>
<td>Cavkaytar (2012)</td>
<td>Children</td>
</tr>
<tr>
<td>Chandler et al. (1993)</td>
<td>Children</td>
</tr>
<tr>
<td>Choi, Wong, and Chung (2012)</td>
<td>Children</td>
</tr>
<tr>
<td>Dollar et al. (2012)</td>
<td>Adults</td>
</tr>
<tr>
<td>Epps, Stern, and Horner (1990)</td>
<td>Children and adults</td>
</tr>
<tr>
<td>Ergenekon et al. (2014)</td>
<td>Children</td>
</tr>
<tr>
<td>Fallow and Rehfeldt (2008)</td>
<td>Children</td>
</tr>
<tr>
<td>Fetko et al. (1999)</td>
<td>Children</td>
</tr>
<tr>
<td>Goh and Bambara (2013)</td>
<td>Adults</td>
</tr>
<tr>
<td>Goodson, Sigafos, O'Reilly, Cannella, and Lancioni (2007)</td>
<td>Adults</td>
</tr>
<tr>
<td>Griffin, Wolery, and Schuster (1992)</td>
<td>Children</td>
</tr>
<tr>
<td>Gruber and Poulson (2016)</td>
<td>Children</td>
</tr>
<tr>
<td>Halasz-Dees and Cuvo (1986)</td>
<td>Children and adults</td>
</tr>
<tr>
<td>Hall, Schuster, Wolery, Gast, and Doyle (1992)</td>
<td>Children</td>
</tr>
<tr>
<td>Heinrich et al. (2016)</td>
<td>Children</td>
</tr>
<tr>
<td>Kagohara et al. (2011)</td>
<td>Children</td>
</tr>
<tr>
<td>Keogh et al. (1987)</td>
<td>Children and adults</td>
</tr>
<tr>
<td>Knight, Spooner, Browder, Smith, and Wood (2013)</td>
<td>Children</td>
</tr>
<tr>
<td>Kourassanis, Jones, and Fienup (2015)</td>
<td>Children</td>
</tr>
<tr>
<td>Le Grice and Blampied (1994)</td>
<td>Children</td>
</tr>
<tr>
<td>Lifshitz (1999)</td>
<td>Children and adults</td>
</tr>
<tr>
<td>Maeser and Thyer (1990)</td>
<td>Children</td>
</tr>
<tr>
<td>McAdam and Cuvo (1994)</td>
<td>Children</td>
</tr>
<tr>
<td>McLaughlin and Walsh (1996)</td>
<td>Children</td>
</tr>
<tr>
<td>McWilliams, Nietupski, and Hamre-Nietupski (1990)</td>
<td>Children</td>
</tr>
</tbody>
</table>
Fifty-two of the studies reported in Table 25 investigated the use of task analysis with children (≤22 years), seven studies assessed its impact with adult participants (>22 years) and five studies involved both children and adults.

According to the Strength of Evidence Classification System (National Autism Center, 2015), task analysis can be categorised as an “established” intervention for children with intellectual and developmental disabilities. A total of
eight studies employing task analysis, chaining, and systematic prompting (SMRS \( \geq 3 \)) produced at least a minimally effective outcome (PND) for participants aged 22 years or younger. At least four of these studies involved 12 or more participants. There were no conflicting results (Table 26). Forty studies (SMRS \( \geq 3 \)), implementing task analysis, chaining, systematic prompting, and reinforcement, demonstrated at least a minimally effective outcome for participants and at least four of these studies involved 12 or more participants. There were no conflicting results (Table 26). Table 27 details the number of studies scoring 3, 4 or 5 on the SMRS, which used total task presentation, forward chaining, and/or backward chaining as part of the intervention.

Table 26

*Total Number of Studies for which a PND Rating could be Calculated, Total Number of Studies with at Least a Minimally Effective PND and the Status of the Target Interventions for a Population of Children with Intellectual and Developmental Disabilities*

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Number of Studies with PND Rating</th>
<th>Number of Studies with at Least a Minimally Effective Outcome (PND)</th>
<th>Established Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task Analysis</td>
<td>8</td>
<td>8</td>
<td>Yes</td>
</tr>
<tr>
<td>Task analysis + chaining + systematic prompting</td>
<td>40</td>
<td>40</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 27

*Number of Studies (SMRS \( \geq 3 \)) Employing Each Chaining Procedure*

<table>
<thead>
<tr>
<th>Total Task Presentation</th>
<th>Forward Chaining</th>
<th>Backward Chaining</th>
</tr>
</thead>
<tbody>
<tr>
<td>46*</td>
<td>8*</td>
<td>2</td>
</tr>
</tbody>
</table>

* Two studies used both forward chaining and total task presentation
The results of the review showed that task analysis can also be classified as an “established” intervention for adults with intellectual and developmental disabilities (Table 28). Four studies using task analysis (SMRS $\geq 3$) chaining, and systematic prompting produced at least a minimally effective outcome (PND) for participants aged 22 years or older. At least four of these studies involved 12 or more participants. There were no conflicting results. A total of five studies (SMRS $\geq 3$), implementing task analysis, chaining, systematic prompting, and reinforcement, demonstrated at least a minimally effective outcome for participants and at least four of these studies involved 12 or more participants. There were no conflicting results. Table 29 details the number of studies scoring 3, 4 or 5 on the SMRS, which used total task presentation, forward chaining, and/or backward chaining as part of the intervention.

Table 28

*Total Number of Studies for which a PND Rating could be Calculated, Total Number of Studies with at Least a Minimally Effective PND and the Status of the Target Interventions for a Population of Adults with Intellectual and Developmental Disabilities*

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Number of Studies with PND Rating</th>
<th>Number of Studies with at Least a Minimally Effective Outcome (PND)</th>
<th>Established Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task Analysis</td>
<td>5</td>
<td>4</td>
<td>Yes</td>
</tr>
<tr>
<td>Task analysis + chaining + systematic prompting</td>
<td>5</td>
<td>5</td>
<td>Yes</td>
</tr>
<tr>
<td>Task analysis + chaining + systematic prompting + reinforcement</td>
<td>5</td>
<td>5</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Table 29

*Number of Studies (SMRS ≥3) Employing Each Chaining Procedure*

<table>
<thead>
<tr>
<th>Total Task Presentation</th>
<th>Forward Chaining</th>
<th>Backward Chaining</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

**Functional Communication Training (FCT) Review**

A total of 71 articles, published between the years 1985-2015, met the inclusion criteria (Appendix C – Table 30). The data extracted across all included studies were summarised and are available from the author on request. Table 31 lists all studies, which scored 3 or more on the SMRS.
Table 31

*Functional Communication Training Studies Scoring at Least an SMRS Score of 3*

<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brown et al. (2000)</td>
<td>Children</td>
</tr>
<tr>
<td>Byiers, Dimian, and Symons (2014)</td>
<td>Children and adults</td>
</tr>
<tr>
<td>Carr and Carlson (1993)</td>
<td>Children</td>
</tr>
<tr>
<td>Chezan, Drasgow, and Martin (2014)</td>
<td>Adults</td>
</tr>
<tr>
<td>Conklin and Mayer (2011)</td>
<td>Adults</td>
</tr>
<tr>
<td>Dalmau et al. (2011)</td>
<td>Children</td>
</tr>
<tr>
<td>Doyle, DeRosa, and Roane (2013)</td>
<td>Children</td>
</tr>
<tr>
<td>Drasgow, Halle, Ostrosky, and Harbers (1996)</td>
<td>Children</td>
</tr>
<tr>
<td>Durand and Carr (1991)</td>
<td>Children</td>
</tr>
<tr>
<td>Durand and Carr (1992)</td>
<td>Children</td>
</tr>
<tr>
<td>Falcomata, Wacker, Ringdahl, Vinquist, and Dutt (2013)</td>
<td>Children</td>
</tr>
<tr>
<td>Hetzroni and Roth (2003)</td>
<td>Children</td>
</tr>
<tr>
<td>Kahng, Iwata, DeLeon, and Worsdell (1997)</td>
<td>Adults</td>
</tr>
<tr>
<td>Kelley, Lerman, and Van Camp (2002)</td>
<td>Children</td>
</tr>
<tr>
<td>Kemp and Carr (1995)</td>
<td>Adults</td>
</tr>
<tr>
<td>Kuhn, Chirighin, and Zelenka (2010)</td>
<td>Children</td>
</tr>
<tr>
<td>Lambert, Bloom, and Irvin (2012)</td>
<td>Children</td>
</tr>
<tr>
<td>Langdon, Carr, and Owen-DeSchryver (2008)</td>
<td>Children</td>
</tr>
<tr>
<td>Lerman, Kelley, Vorndran, Kuhn, and LaRue (2002)</td>
<td>Children and adults</td>
</tr>
<tr>
<td>O'Reilly et al. (2012)</td>
<td>Children</td>
</tr>
<tr>
<td>Radstaake et al. (2013)</td>
<td>Children</td>
</tr>
<tr>
<td>Rispoli, Camargo, Machalicek, Lang, and Sigafoos (2014)</td>
<td>Children</td>
</tr>
<tr>
<td>Schmidt, Drasgow, Halle, Martin, and Bliss (2014)</td>
<td>Children</td>
</tr>
<tr>
<td>Shirley, Iwata, Kahng, Mazaleski, and Lerman (1997)</td>
<td>Adults</td>
</tr>
</tbody>
</table>
Chapter 7

Twenty-three of the studies in Table 31 investigated the impact of functional communication training (FCT) on children (≤22 years), five studies examined the effects on adult participants (>22 years), and three studies involved both children and adults.

According to the Strength of Evidence Classification System (National Autism Center, 2015), functional communication training can be characterised as an “established” intervention for children with intellectual and developmental disabilities (Table 32). Five single-subject experimental design studies (SSED) using functional communication training alone (SMRS ≥3), produced at least a minimally effective outcome for participants aged 22 years and younger. At least four of these studies involved 12 or more participants. There were no conflicting results. A total of 13 SSED studies, which implemented a combination of functional communication training and extinction (SMRS ≥3), resulted in at least a minimally effective outcome for participants aged 22 years or younger. At least four of these studies involved 12 or more participants. There were no conflicting results. Table 33 shows that access to tangibles was the most frequently targeted behavioural function across studies (SMRS ≥3) implementing functional communication training alone, while access to escape was most often the focus of studies using functional communication training and extinction.
Chapter 7

Table 32

*Total Number of Studies for which a PND Rating could be Calculated, Total Number of Studies with at Least a Minimally Effective PND and the Status of the Target Interventions for a Population of Children with Intellectual and Developmental Disabilities*

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Number of Studies with PND Rating</th>
<th>Number of Studies with at Least a Minimally Effective Outcome (PND)</th>
<th>Established Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FCT</td>
<td>6*</td>
<td>5</td>
<td>Yes</td>
</tr>
<tr>
<td>FCT + other</td>
<td>4</td>
<td>3</td>
<td>No</td>
</tr>
<tr>
<td>FCT + extinction (EXT)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FCT + EXT</td>
<td>14*</td>
<td>13</td>
<td>Yes</td>
</tr>
<tr>
<td>FCT + punishment (PUN)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FCT + PUN</td>
<td>1</td>
<td>1</td>
<td>No</td>
</tr>
<tr>
<td>FCT + PUN + EXT + other</td>
<td>0</td>
<td>0</td>
<td>No</td>
</tr>
</tbody>
</table>

* Two studies initially implemented FCT alone for all participants but subsequently implemented FCT + EXT for at least one participant (but not all participants)

Table 33

*Breakdown of Behavioural Function According to Treatment Package*

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Number of Studies: Tangible Function</th>
<th>Number of Studies: Attention Function</th>
<th>Number of Studies: Escape Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCT</td>
<td>6</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>FCT + other</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>FCT + extinction (EXT)</td>
<td>4</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>FCT + PUN</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>FCT + PUN + EXT + other</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

The results of the review also showed that functional communication training combined with extinction can be characterised as an “emerging” intervention for adults with intellectual and developmental disabilities (National Autism Center, 2015). A total of three studies (n=9) using functional communication training and
extinction (SMRS ≥3), produced at least a minimally effective outcome (PND) for adult participants (see Table 34). Access to tangibles was the most frequently targeted behavioural function in studies (SMRS ≥3) employing functional communication training only. However, there were no clear distinctions across the other categories (i.e., FCT + extinction, FCT + other) (Table 35).

Table 34

*Total Number of Studies for which a PND Rating could be Calculated, Total Number of Studies with at Least a Minimally Effective PND and the Status of the Target Interventions for a Population of Adults with Intellectual and Developmental Disabilities*

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Number of Studies with PND</th>
<th>Number of Studies with at Least a Minimally Effective Outcome (PND)</th>
<th>Established Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FCT</td>
<td>2</td>
<td>2</td>
<td>No</td>
</tr>
<tr>
<td>FCT + extinction</td>
<td>3</td>
<td>3</td>
<td>No</td>
</tr>
<tr>
<td>FCT + other</td>
<td>2</td>
<td>2</td>
<td>No</td>
</tr>
</tbody>
</table>

Table 35

*Breakdown of Behavioural Function According to Treatment Package*

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Number of Studies: Tangible Function</th>
<th>Number Studies: Attention Function</th>
<th>Number of Studies: Escape Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCT</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>FCT + extinction</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>FCT + other</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

**Evaluation of the Use of Reinforcement**

Tables 36, 37, and 38 (Appendix D) list all studies from the systematic prompting, task analysis, and functional communication training reviews, which used reinforcement and scored 3 or more on the SMRS.
Results show that 123 out of 146 studies included in the review of systematic prompting implemented reinforcement as part of the intervention package and 122 of these studies used positive reinforcement. Furthermore, 92 of the 93 studies for which a PND rating could be calculated, demonstrated at least a minimally effective outcome. Forty-seven of the 63 articles included in the task analysis review employed reinforcement. All of these studies employed positive reinforcement. Data showed that 100% of studies for which a PND rating could be calculated, produced at least a minimally effective outcome. Finally, all studies in the functional communication training review implemented either positive or negative reinforcement and 92.3% demonstrated at least a minimally effective outcome (Table 39).

Table 39

<table>
<thead>
<tr>
<th></th>
<th>Percentage of Studies Using Reinforcement</th>
<th>Percentage of Studies Using Positive Reinforcement</th>
<th>Percentage of Studies with PND that Produced at Least a Minimally Effective Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Systematic prompting</td>
<td>86%</td>
<td>85.3%</td>
<td>98.9%</td>
</tr>
<tr>
<td>Task analysis</td>
<td>74.6%</td>
<td>74.6%</td>
<td>100%</td>
</tr>
<tr>
<td>FCT</td>
<td>100%</td>
<td>87.1%</td>
<td>92.3%</td>
</tr>
</tbody>
</table>

Discussion

The objective of the current study was to identify evidence-based practices, which effectively address deficits in communication, social, and daily living skills for this population. Using the autism-specific literature as a guide (i.e., National Autism Center, 2015; Wong et al., 2014), a series of systematic reviews were
conducted to evaluate the evidence-base supporting the use of the following 
behavioural interventions with people with intellectual and developmental 
disabilities: reinforcement, systematic prompting, functional communication training, 
and task analysis.

According to the National Standards Report and the National Professional 
Development Centre review, the evidence-based behavioural interventions listed 
above are effective treatments for increasing a variety of adaptive behaviour skills 
(e.g., communication, academic skills, interpersonal interactions, self-regulation and 
personal responsibility) among people with autism. Data extracted from the current 
series of systematic literature reviews support this position for individuals with 
intellectual and developmental disabilities.

The evidence-base for systematic prompting included studies focused on 
increasing phonic skills (Ahlgrim-Delzell et al., 2014), pedestrian safety (Batu et al., 
2004), self-monitoring skills for performing tasks independently (Copeland & 
Hughes, 2000), and conversational skills (Downing, 1987). Likewise, the studies 
supporting the efficacy of task analysis tended to target functional living skills, such 
as learning to use a photocopier (Browder et al., 1993), improving personal hygiene 
(Stokes et al., 2004), and carrying out shopping tasks independently (Taylor & 
O'Reilly, 2000)). Finally, studies from the functional communication training review 
taxtured service users to appropriately express their needs and wants, without reverting 
to challenging behaviour (e.g., Derby et al., 1997; Falcomata et al., 2013; Hagopian, 

A separate literature review on reinforcement was deemed unnecessary as 
experts in the field of Applied Behaviour Analysis recognise reinforcement as “the 
most important and widely applied principle of behaviour analysis… the
fundamental building block for the selection of operant behaviour” (Cooper et al., 2014, p. 255). This is reflected in the frequency of its implementation across studies included in the three systematic literature reviews, with reinforcement procedures being employed in 74.6%, 86% and 100% of studies in the task analysis, systematic prompting and functional communication training reviews, respectively. In addition, over 90% of studies with PND data tested interventions, which produced at least a minimally effective outcome. Therefore, it can be argued that when combined with systematic prompting, task analysis or functional communication training, reinforcement represents an important component of evidence-based practice. This is consistent with findings from both the National Standards Project and the National Professional Development Centre review, which demonstrate that reinforcement is an established evidence-based practice for people with autism (National Autism Center, 2015; Wong et al., 2014).

For the systematic prompting, functional communication training and task analysis reviews, the Strength of Evidence Classification System (National Autism Center, 2015) was employed to evaluate the experimental rigor of included studies. Systematic prompting procedures (i.e., least-to-most prompting, most-to-least prompting and simultaneous prompting), functional communication training (used alone or in combination with extinction), and task analysis (combined with chaining, systematic prompting, and/or reinforcement) were identified as “established” interventions for people with intellectual and developmental disabilities, aged 22 years old or younger. Four high quality studies also demonstrated the efficacy of graduated guidance for this population. However, the number of participants across these studies was insufficient to meet criteria for the “established” classification. For adults (>22 years) with intellectual and developmental disabilities, task analysis
(combined with chaining, systematic prompting, and/or reinforcement) was shown to be an “established” intervention. Least-to-most prompting and functional communication training combined with extinction were characterised as “emerging” treatments for this adult population.

It is clear that the number of “established” interventions identified from the current literature reviews was substantially greater for people aged 22 years and younger. However, this does not imply that the behavioural interventions reviewed are ineffective for adults with intellectual and developmental disabilities. A more plausible explanation for the discrepancy is that the majority of studies in each review tended to focus on participants aged 22 years and younger. For example, from the 143 experimentally rigorous studies (SMRS >3) that were identified for the systematic prompting review, 125 focused exclusively on children (<22 years) and 14 on adults. Similarly, only eight studies (SMRS >3) in the functional communication training review involved an adult population.

This issue may have been further compounded by the exclusion of some high-quality studies (SMRS >3) due to PND ratings being unattainable. However, an objective method of evaluating intervention outcomes was required and the PND approach is generally recommended because of its ease of interpretation for behavioural researchers (Olive & Franco, 2008). Nonetheless, it is clear that the research base for adults is restricted, making it difficult to establish the evidence-based status of these practices. The National Standards Project reported similar difficulties in the literature for adults with autism (National Autism Center, 2015). However, despite limited research, the available studies are supportive of the use of systematic prompting, task analysis, and functional communication training with adults with intellectual and developmental disabilities (raw data available from the
author on request). Furthermore, with some additional research, these interventions may also be characterised as “established” interventions for adults with intellectual and developmental disabilities.

The current study not only demonstrated the efficacy of the four behavioural interventions reviewed, it also showed their effectiveness when used in combination. For example, it is recognised that functional communication training often relies on reinforcement (Fisher, Kuhn, & Thompson, 1998; Skinner, 1957) and systematic prompting strategies (e.g., Brown et al., 2000; Carr & Carlson, 1993) to achieve success. In their seminal study, Carr and Durand (1985) employed verbal prompts and positive reinforcement, as part of a functional communication training package, to successful replace instances of challenging behaviour with functionally matched, socially appropriate communication responses for two children with developmental disabilities and two children with brain damage. Similarly, Braithwaite and Richdale (2000) implemented functional communication training to teach a 7 year old boy with a dual diagnosis of autism and an intellectual disability to make appropriate vocal requests when he wanted to escape something or access a desired item or activity. Their treatment package included verbal prompts and positive and negative reinforcement.

The critical importance of reinforcement and systematic prompting strategies in the success of instructional procedures, such as task analysis, was also demonstrated. Data from included studies show that a multi-element treatment package, which includes task analysis, chaining, systematic prompting, and reinforcement, meets criteria for classification as an “established” intervention for both children and adults with intellectual and developmental disabilities. In his study, Cavkaytar (2012) effectively taught waiter skills to three young adults (18-22
years) with intellectual disabilities. The target skills were task analysed into 125 individual steps, divided across five specific waiter duties. Using a combination of task analysis, total task presentation, positive reinforcement, and a least-to-most prompting procedure, all three participants successfully acquired the target waiter skills. Likewise, Shimizu et al. (2010) employed a combination of task analysis, forward chaining, least-to-most prompting, and positive reinforcement to effectively teach pre-schoolers with developmental disabilities to use a computer mouse.

Given the weight of supporting evidence for the behavioural interventions reviewed, the degree of compatibility between these interventions and the needs of service users within the host organisation, it was decided to proceed with a staff training package that would focus on reinforcement, systematic prompting, functional communication training, and task analysis. As previously mentioned, the National Professional Development Centre, in collaboration with the University of North Carolina, have created online manuals that focus on reinforcement, systematic prompting, functional communication training, and task analysis (AFIRM Team, 2015a; 2015b; 2015c; Franzone, 2009; Wong et al., 2014).

These manuals outline the procedures involved in planning for, implementing and evaluating the impact of each of these evidence-based practices (AFIRM Team, 2015a; 2015b; 2015c; Franzone, 2009) and were the primary resource in the development of content for the staff training programme that would be tested in the next study. However, in order to achieve functional dissemination and change among frontline staff, a multifaceted training approach would be required (van Oorsouw et al., 2009; Ward-Horner & Sturmey, 2012). The objective for the subsequent study was the identification and evaluation of an effective training model,
in order to disseminate the selected evidence-based practices to staff supporting people with intellectual and developmental disabilities.
Chapter 8

Dissemination of Evidence-Based Practice to Frontline Staff Working in the Field of Intellectual Disability: A Pilot Study
Abstract

The competency of frontline staff, working in the intellectual disability sector, significantly impacts the behaviour and quality of life of the service users they support (Jahr, 1998). Behaviour Skills Training has been shown to be a robust training method (e.g., Parsons et al., 2012; Sarokoff & Sturmey, 2008) and employs the structural components identified as necessary for the successful dissemination of training material (Parsons et al., 2012; van Oorsouw et al., 2009). Therefore, a randomised control trial experimental design was selected to evaluate the effectiveness of a Behaviour Skills Training package in disseminating four behavioural practices, which were identified as having an adequate evidence base for use with people with intellectual disabilities. Randomised control trials can be costly (Haskins & Feldman, 2016) and it is prudent to conduct a pilot study to assess feasibility, before proceeding with a larger-scale study. Seventeen frontline staff participated in a pilot study employing a pre- and post-intervention design. Results showed that the training package could be delivered within the time-period allocated by the host organisation. In addition, a significant improvement in behavioural knowledge and targeted skills was observed among participants. The results of the pilot study were used to inform the subsequent randomised control trial in relation to training content, mastery criterion of target skills, and the assessment of service user outcomes. The degree of change in participants’ behavioural knowledge was also used to guide sample size calculation for the subsequent trial.
Chapter 8

The goal of the training needs analysis, outlined in Chapter 6, was to identify the needs of staff with respect to effectively supporting service users within the host organisation. According to participant reports, the majority of service users were functioning at the lowest levels across all areas of the Vineland-II “communication”, “daily living skills” and “social skills” domains. Furthermore, over 90% of service users were reported to be engaged in elevated (51.7%) or clinically significant (38.3%) levels of externalised challenging behaviour (Vineland-II). However, results showed that the majority of staff reported receiving no preparatory training from the host organisation, in either behaviour management strategies or skills teaching procedures, prior to starting work with service users. Furthermore, less than 40% of staff reported that they received ongoing training in how to teach service users new skills. As such, over half of the participants were not receiving ongoing training despite 93.5% being employed with the RehabCare organisation for at least one year.

Furthermore, the results of Chapter 6 showed no correlation between training delivered by the host organisation and knowledge of behaviour management principles. This disconnect was reflected in the practical observations of staff working with service users in the typical support environment; they did not demonstrate any evidence of utilising systematic prompting and reinforcement procedures, when supporting service users to develop new skills. The results of the training needs analysis demonstrated the exigency for support among service users and the inconsistencies and discrepancies associated with the extant staff training protocol within the organisation. Therefore, an empirically supported evidence-based training package to address service user deficits and challenges, was deemed necessary for the RehabCare organisation.
As discussed above, the training needs analysis highlighted communication, social and daily living skills as priority areas for service users in the RehabCare organisation. Although significant progress has been made in identifying a comprehensive list of evidence-based practices for people with autism (National Autism Center, 2009; 2015; Wong et al., 2014), similar progress has not yet been achieved for the intellectual disability population. Reinforcement, systematic prompting, functional communication training, and task analysis procedures were highlighted by the National Standards Project (National Autism Center, 2009; 2015) and the National Professional Development Centre report (Wong et al., 2014) as effective practices for targeting the deficits in communication, social and daily living skills among individuals with autism. As a result, a series of systematic literature reviews were conducted to evaluate the evidence-base supporting the use of these practices with individuals with an intellectual disability.

The results of the systematic reviews reported in Chapter 7 provided empirical support for the use of reinforcement, systematic prompting, functional communication training, and task analysis in interventions designed for people with intellectual disabilities. Using the online manuals, created by the National Professional Development Centre (AFIRM Team, 2015a; 2015b; 2015c; Franzone, 2009; Wong et al., 2014), the content for the evidence-based staff training package was developed for the RehabCare organisation. However, according to the literature there is a significant gap between the emerging knowledge pertaining to evidence-based practice for people with intellectual disabilities and the competencies and skills displayed by staff in applied settings (Campbell, 2010; Hile & Walbran, 1991; Rapp et al., 2010; Swain et al., 2010). This theory-practice gap can arise as a result of a number of factors, including insufficient knowledge (Hastings, 1996), and
inadequate training (Campbell, 2007; Jahr, 1998). Results from the systematic review, conducted as part of Chapter 5, demonstrated that training in the application of many established evidence-based practices, such as functional communication training, was not evaluated in the literature. Therefore, a method for effectively disseminating these practices must be identified.

In 2009, van Oorsouw and colleagues conducted a meta-analysis to investigate the structural components necessary for effective staff training (van Oorsouw et al., 2009). The authors concluded that optimal results are achieved through the combined use of in-service training and on-the-job coaching, with target information being delivered to staff in a number of different ways. According to Reid et al. (2003), staff training should be performance- and competency-based as knowledge acquisition alone will not produce improved staff-service user interactions (Embry & Biglan, 2008; Grey et al., 2007); there must also be an emphasis on skill acquisition (Parsons et al., 2012). A performance-based approach to training emphasises active display of the skills being trained by trainers and trainees, while a competency-based approach highlights the importance of persisting with training until pre-determined performance criteria are achieved (Parsons et al., 2012).

As previously discussed in Chapter 4, Behaviour Skills Training combines instruction, modelling, rehearsal, and feedback to teach new skills (Ward-Horner & Sturmey, 2012). When implemented appropriately, this training package is both performance- and competency-based (Parsons et al., 2012) and has been successfully applied to train a wide range of groups, including frontline staff, a variety of new work-related skills (e.g., Homlitas et al., 2014; Maggin et al., 2012; Parsons et al., 2012; Sarokoff & Sturmey, 2008). However, the efficacy of Behaviour Skills
Training has yet to be tested in a large group-based experimental design. This approach to staff training has also never been evaluated in the context of delivering the following combination of evidence-based practices to frontline staff, working in the intellectual disability sector: reinforcement, systematic prompting, functional communication training, and task analysis.

In order to rigorously test the Behaviour Skills Training method, a randomised control trial would be required. According to Akobeng (2005), the randomised control trial is the most reliable experimental design in determining the effectiveness of an intervention, as it reduces the impact of confounding variables to a greater extent than other research designs. However, randomised control trials can be costly, in terms of resource investment (Haskins & Feldman, 2016; Liu & Wyatt, 2011). Therefore, it is considered prudent to conduct a pilot study, in advance of the randomised control trial, to ascertain the practicalities of implementation (Tickle-Degnen, 2013). A pilot study can be defined as “a smaller version of the main study that is run in miniature to test whether the components of the main study can all work together… [and resembles] the main study in many respects, including the assessment of the primary outcome” (National Institute for Health Research, 2012).

Therefore, the goal of the current study was to conduct a pilot study to assess the feasibility of implementing an evidence-based Behaviour Skills Training package within the host organisation. The first objective was to determine if the proposed study was practical and identify potential threats to outcome validity. The second objective sought to examine the primary outcomes of the study, in the context of measurability, degree of change, and potential barriers to dissemination in the randomised control trial. It was hypothesised that the proposed training package would lead to an improvement in target behavioural knowledge, measures of
psychological well-being and the implementation of pre-determined evidence-based skills among the participating frontline staff from the RehabCare organisation. The pilot study also set out to investigate potential barriers to effective dissemination, related to perceived supervisory support among participants, as well as their attitudes toward evidence-based practice.

**Phase I – Behaviour Skills Training**

**Method**

**Participants.** Participants were 17 frontline staff, recruited from two RehabCare service sites. All participants worked with adults with intellectual disabilities. Table 40 summarises the characteristics of these participants. The upper panel provides details about the number of participants recruited from each service site, the number of participants working in day support and residential services, respectively, and the number of participants employed in the host organisation for less than a year, 1-5 years, 6-10 years and over 11 years. The lower panel characterises the participants according to gender, education and job title.
Table 40

Summary of Participant Characteristics: Upper Panel - Geographic Location, Service Type, and Length of Service with the Organisation; Lower Panel – Gender, Education, and Job Title (N=17)

<table>
<thead>
<tr>
<th>Geographic Location</th>
<th>Number of Participants</th>
<th>Service Type</th>
<th>Number of Participants</th>
<th>Length of Service with the Organisation</th>
<th>Number of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leinster</td>
<td>8</td>
<td>Day support</td>
<td>13</td>
<td>&lt; 1 year</td>
<td>1</td>
</tr>
<tr>
<td>Munster</td>
<td>9</td>
<td>Residential</td>
<td>4</td>
<td>1-5 years</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6-10 years</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11+ years</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
<th>Number of Participants</th>
<th>Education</th>
<th>Number of Participants</th>
<th>Job Title</th>
<th>Number of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>3</td>
<td>Certificate</td>
<td>2</td>
<td>Careworker</td>
<td>1</td>
</tr>
<tr>
<td>Female</td>
<td>14</td>
<td>Diploma</td>
<td>5</td>
<td>Programme facilitator</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Undergraduate degree</td>
<td>9</td>
<td>Programme supervisor</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Other</td>
<td>1</td>
<td>Team leader</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Service manager</td>
<td>3</td>
</tr>
</tbody>
</table>

Trainers. Two behaviour therapists, who were employed by the RehabCare organisation, were enlisted as trainers in the current study. Both trainers held postgraduate qualifications in Applied Behaviour Analysis and one was a board certified behaviour analyst (BCBA®).

Experimenters. The lead experimenter held an MSc in Psychology and was a board certified behaviour analyst (BCBA®). She had over 10 years’ experience working with children and adults with intellectual and developmental disabilities. The second experimenter had achieved a Postgraduate Diploma in Applied Behaviour Analysis and was pursuing an MSc in Applied Behaviour Analysis at the time of the study.

Setting. Participants were recruited from RehabCare, the organisation that participated in the training needs analysis described in Chapter 6.
**Trainer instruction.** A 1-day instructional session for trainers was hosted by the service provider organisation in a classroom setting. The room was large enough to deliver lecture-based instruction to two participants and facilitate in-vivo skill demonstration and role-play rehearsal of the target skills.

**Administration of measures.** Measures were administered in a conference room, classroom or office setting.

**Behaviour Skills Training.** Two 3-day training sessions were delivered; one in the Leinster region and one in the Munster region. The first session took place in a classroom setting and the second session was hosted in a conference room. Both rooms were large enough to seat and deliver didactic instruction via projector to 10 participants. There was also adequate space to accommodate role-play rehearsal of the target skills (i.e., sufficient space for five dyads) and manage the potential for uncontrolled observations among participants.

**Materials.**

**Trainer instruction.**

**Trainer pack.** Each trainer received an instructional pack containing a hand-out of the “Behaviour Skills Training” PowerPoint presentation, and instructions for the staff instructional sessions (available from the author on request). In addition, the trainers received a pack containing the following materials for each of the evidence-based practice modules (i.e., reinforcement, systematic prompting, task analysis, and functional communication training modules): PowerPoint presentation hand-out, training manual, mastery assessment data sheet, reflective practice, and additional hand-outs relevant to specific modules (available from the author on request).
Evidence-based practice training packs. Participants received a pack containing the following materials for each of the evidence-based practice modules: PowerPoint presentation hand-out, training manual, mastery assessment data sheet, reflective practice, and additional hand-outs relevant to specific modules (available from the author on request).

PowerPoint Presentations. A PowerPoint presentation was developed for each evidence-based practice module. Presentations included didactic instruction on planning for, implementing and evaluating the impact of each evidence-based practice. Video models of target skills were embedded within the presentations, to ensure standardisation of models across trainers and training sessions. The video models presented at least one skill from each of the four evidence-based practices: positive reinforcement (reinforcement), least-to-most prompting (systematic prompting), requesting access to a desired item or activity (functional communication training) and backward chaining and total task presentation (task analysis). The second experimenter and a postgraduate psychology student were recruited as confederates, with the second experimenter playing the part of the staff member and the student acting as the service user. Scenarios, which are typically experienced on a daily basis in adult services (e.g., folding laundry, making tea, requesting a treat), were used to demonstrate the correct implementation of each skill. A standard digital video recorder and tripod were employed to record the video models and the lead experimenter delivered a voiceover at the beginning of each segment, explaining the sequence of steps involved in implementing the target skills.

Rehearsal props. Items that would typically be available at each service site (e.g., dish cloths, cups, cutlery, hairbrush, face cloths, food) were used as props during rehearsal sessions.
Measures.

**Demographic questionnaire.** The staff demographic questionnaire addressed gender, educational level, job title, work setting and length of service within the RehabCare organisation. The questionnaire also examined whether participants had previously received training in reinforcement, systematic prompting, functional communication training, and task analysis, as well as the method of delivery for training in these practices.

**Knowledge assessments for each evidence-based practice module (available from the author on request).** The knowledge assessments were developed by the experimenter using the content of the National Professional Development Centre online manuals and associated quizzes (AFIRM Team, 2015a; 2015b; 2015c; Franzone, 2009). There were four assessments, each measuring knowledge on one of the four target evidence-based practices; reinforcement, systematic prompting, functional communication training, and task analysis. Each assessment contained 10 multiple-choice items, with each correct item receiving a score of ‘1’; the maximum score on the measure was 10.

When the primary focus of an assessment is measurement of “within-individual growth” before and after a training programme, the assessment should be evaluated in terms of its edumetric rather than its psychometric properties (Carver, 1974, p. 512). Applying a psychometric analysis on an edumetric assessment, may result in a test that is insensitive to individual gain (Carver, 1974). Therefore, as the purpose of the knowledge assessments in the current study was to evaluate potential knowledge gain following exposure to Behaviour Skills Training, an edumetric analysis was conducted. This analysis examined: (1) item difficulty; (2) edumetric
validity (i.e., the degree to which a measure captures gain in situations where it is expected), and (3) edumetric reliability (i.e., consistency of gain as reflected by the measure).

**Item difficulty.** The difficulty index of items on each of the knowledge assessments was assessed and found to be adequate. The difficulty index is defined as the percentage of participants providing a correct response on a particular item. The optimal range is 20-80%, with a lower percentage indicating higher levels of difficulty (Dixon, 1994; Taib & Yusoff, 2014). At baseline, the difficulty index for items ranged from 11.8% to 100% ($M=50.6\%, SD=27.6\%$) on the reinforcement knowledge assessment, 5.9% to 100% ($M=40.6\%, SD=31.8\%$) on the systematic prompting knowledge assessment, 0% to 58.8% ($M=30.6\%, SD=19.8\%$) on the functional communication training knowledge assessment, and 0% to 94.1% ($M=50.6\%, SD=29.9\%$) on the task analysis knowledge assessment. Based on these results, it is clear that the average level of difficulty for each knowledge assessment fell within the optimal range. At post-intervention testing, the difficulty index for items ranged from 11.8% to 100% ($M=68.2\%, SD=26.5\%$) on the reinforcement knowledge assessment, 11.8% to 94.9% on the systematic prompting assessment ($M=63\%, SD=28.2\%$), 17.6% to 100% ($M=74.7\%, SD=25\%$) on the functional communication training knowledge assessment, and 11.8%-100% ($M=78.2\%, SD=29.2\%$) on the task analysis knowledge assessment. Again, the average level of difficulty for each knowledge assessment fell within the optimal range.

**Edumetric validity.** Edumetric validity for all knowledge assessments was demonstrated; participants scored significantly higher on the knowledge assessments, following completion of the relevant evidence-based practice module (Carver, 1974; Deniz & Alsaffar, 2013) and the effect size was consistently large. Detailed analyses
can be found in the results section. Analyses also showed a reduction in item difficulty, across all knowledge assessments, at post-intervention testing.

_Edumetric reliability._ Edumetric reliability can be evaluated by administering a measure on two occasions, under equivalent treatment conditions. A measure is considered edumetrically reliable if it demonstrates consistency of gains for an individual across the two assessments. In the current study, edumetric reliability was demonstrated for all knowledge assessments; participants’ scores did not change significantly from post-intervention to follow-up testing. A series of sign tests were conducted and no significant differences were detected between post-intervention and follow-up scores for the reinforcement knowledge assessment \( (p = .687) \), the systematic prompting knowledge assessment \( (p = 1.00) \), the functional communication training knowledge assessment \( (p = .687) \), and the task analysis knowledge assessment \( (p = 1.00) \). The fact that there was no significant change at post-intervention and follow-up testing indicates that the knowledge assessments were consistent in detecting knowledge gain after training.

_Mastery assessments for each evidence-based practice module (available from the author on request)._ The mastery assessments were task analyses, developed by the experimenter, to evaluate participants’ procedural fidelity while implementing skills related to each of the four evidence-based practice modules. These assessments examined participants’ fidelity in the context of a service user confederate responding correctly and not responding to a particular situation, during rehearsal sessions. The task analyses were based on the National Professional Development Centre guidelines on appropriate implementation of the four target evidence-based practice modules (AFIRM Team, 2015a; 2015b; 2015c; Franzone, 2009).
Each task analysis was divided into 2-3 sections, depending on the target skill. The sections listed steps, which participants were expected to perform. Trainers recorded whether or not a step had been performed correctly by indicating “Y” for yes and “N” for no. The percentage of steps correctly performed was calculated for each section on the task analysis. The average percentage of steps performed correctly for each target skill was then determined.

**Training Acceptability Survey for each evidence-based practice module** *(Underwood, O’Meara, & Harvey, 2002).* The training acceptability survey is a six-item respondent-based measure that examines training, in relation to its relevance, planning, opportunities for participation, resource use, areas of difficulty, providing motivation to learn more, and overall usefulness. Responses are measured on a four-point Likert scale, with “not true” scoring a 1 and “completely true” scoring a 4. The psychometric properties of this measure had not previously been assessed. However, in the current study the Cronbach alpha coefficient was 0.89 for the reinforcement module, 0.93 for the systematic prompting module, 0.89 for the functional communication training module, and 0.92 for the task analysis module.

**Attitudes to Evidence-Based Practice Questionnaire (McKenna, Ashton, & Keeney, 2004).** The Attitudes to Evidence-Based Practice Questionnaire is a 26-item respondent-based rating scale that examines barriers to the use of evidence-based practice in care settings. This scale uses a five-point Likert-scale rating for each item. A panel of experts determined the content validity of this measure and face validity was established through a pilot study, involving a sample of 40 randomly selected participants. Finally, the questionnaire had a reliability coefficient (internal consistency) of 0.74 (McKenna et al., 2004).
Maslach Burnout Inventory – Human Services Survey (MBI-HSS; Maslach et al., 1996). This was identical to the measure employed in the study reported in Chapter 6.

Minnesota Satisfaction Questionnaire – Short Form (MSQ-SF; Weiss et al., 1967). This was identical to the measure employed in the study reported in Chapter 6.

Short Version of the Occupational Self-Efficacy Scale (Rigotti, Schyns, & Mohr, 2008). The Short Version of the Occupational Self-Efficacy Scale is a six-item respondent-based instrument that measures an individual’s occupational self-efficacy. This scale uses a six-point Likert scale, with ‘not true at all’ scoring a 1 and ‘completely true’ scoring a 6. The internal consistency, structural and construct validity was assessed among a sample of 1,535 participants and found to be adequate. The measure was shown to have a reliability coefficient (internal consistency) ranging from 0.85 to 0.90 and confirmatory factor analyses supported its factorial validity. In addition, content validity was established by replicating previously observed correlations between occupational self-efficacy and two job-related attitudes; job satisfaction and commitment (Rigotti et al., 2008).

Test of Knowledge (Denne, Thomas, Hastings, & Hughes, 2015). The Test of Knowledge evaluates knowledge of behavioural principles and applications. It contains 20 short answer and multiple-choice questions. Each item answered correctly is scored as a ‘1’; the maximum score on the measure is 20. The construct validity has been found to be adequate based on assessment by a team of senior behaviour analysts (Denne et al., 2015).
The Perceptions of Supervisory Support Scale (PSS; Fukui et al., 2014). This was identical to the measure employed in the study reported in Chapter 6.

The Shortened Ways of Coping Questionnaire (SWC-R; Hatton & Emerson, 1995). This was identical to the measure employed in the study reported in Chapter 6.

Design. A within-subjects, pre-post-test design was used to evaluate the impact of an evidence-based Behaviour Skills Training programme on participant knowledge of each target evidence-based practice, as well as participants’ psychological well-being. Observations to criterion, a type of event recording, was also employed during training to measure the number of practice opportunities required by participants to reach a pre-determined level of proficiency/mastery in the implementation of skills related to each target evidence-based practice (Cooper et al., 2014).

Procedure.

Development of evidence-based practice modules. Training content was based on the online evidence-based practice manuals created by the National Professional Development Centre (AFIRM Team, 2015a; 2015b; 2015c; Franzone, 2009) and supplemented, in some instances, to provide further explanation and information on each practice (e.g., Cooper et al., 2014; Kurtz et al., 2011; National Autism Center, 2009). Instruction manuals were developed that focused on describing, planning for, implementing and monitoring the use of each evidence-based practice. Skills targeted for video modelling and rehearsal sessions were selected according to results from the systematic literature reviews reported in Chapter 7.
Findings from these reviews showed that systematic prompting, functional communication training, and task analysis are considered “established” evidence-based practices for people with intellectual and developmental disabilities, aged 22 years and younger. Task analysis (“established”), least-to-most prompting (“emerging”) and functional communication training combined with extinction (“emerging”) are considered evidence-based practices for use with people with intellectual and developmental disabilities, aged 23 years or older. Reinforcement procedures, when combined with these practices, also represent an important component of evidence-based practice (Carr & Durand, 1985; Cooper et al., 2014; Franzone, 2009; Kurtz et al., 2011; Neitzel, 2009; Okinaka & Shimazaki, 2011). As a result, the following skills were targeted in video models and rehearsal sessions for the reinforcement, systematic prompting, and functional communication training modules, respectively: positive reinforcement; least-to-most prompting, and requesting access to items or activities. For the task analysis module, implementation of total task presentation and backward chaining were selected for depiction in video models and backward chaining, alone, was targeted for the rehearsal sessions. The rationale for covering total task presentation in the video models was the frequency with which it was successfully implemented in the literature. However, based on previous research, backward chaining was selected ahead of other chaining methods as the target skill for the rehearsal sessions.

In a study of 22 participants, Walls, Zane, and Ellis (1981) found that forward and backward chaining produced significantly less errors on vocational tasks than total task presentation. Notwithstanding the delayed progress, increased errors have also been associated with elevated levels of frustration and self-injurious behaviour among people with developmental disabilities (Green, 1996; Smith, Iwata, Goh, &
Shore, 1995; Smith, 2001). In addition, to a reduced probability of error, backward chaining is also likely to promote independence (Cooper et al., 2014). Foxx (1982) argues that backward chaining is often effective for learners with disabilities as the last response in the chain is closest to the terminal reinforcer, which is generally the natural reinforcement contingency. As a result, of being repeatedly paired with the terminal reinforcer, responding is strengthened. Therefore, although backward chaining was employed less often than other strategies in the high-quality studies (SMRS ≥3) reviewed in Chapter 7, the results were predominantly positive across a wide number of participants (n=29), making it a viable and functional skill to teach staff.

**Trainer instruction.** Trainers received a total of 4.5 hours coaching, divided across three sessions; all sessions were scheduled for the same day. The first session involved instruction on the content and structure of each module, combined with a PowerPoint presentation on Behaviour Skills Training. This session also included in-vivo modelling of the rehearsal and feedback components of the reinforcement and systematic prompting modules. The lead experimenter played the part of trainer and the second experimenter and a confederate acted as the staff trainee and service user, respectively. The first session lasted approximately 1.5 hours.

The second session focused on facilitating each trainer to rehearse the delivery of either the reinforcement or systematic prompting module. As such, if the first trainer rehearsed the reinforcement module, the second trainer focused on the systematic prompting module. As part of rehearsal, trainers were expected to summarise the material that they would deliver, outline the learning objectives for their target module and present the video models. For the rehearsal and feedback components, the second experimenter and a confederate acted the parts of the staff.
trainee and service user. Each trainer conducted a full rehearsal and feedback session with the “staff trainee” (i.e., the second experimenter) and recorded implementation fidelity using the designated mastery assessment data sheet (available from the author on request). Throughout this session, the lead experimenter also recorded procedural fidelity data on trainer performance (data sheet available from the author on request). At the end of rehearsal, the experimenter delivered feedback to each trainer, based on the procedural fidelity data which they had recorded. Trainers were expected to implement all steps across the instruction, modelling, rehearsal and feedback sections with at least 90% accuracy to achieve mastery. The second session lasted approximately 1.5 hours. Both trainers reached mastery criterion on their first attempt.

The final session functioned as an opportunity for the trainers to generalise their newly acquired competency in Behaviour Skills Training to novel material. Each trainer was required to deliver one of the remaining evidence-based practice modules (i.e., functional communication training and task analysis) using the same format as that employed in the previous session. In this instance, trainers did not receive prior instruction and modelling in the material to be delivered. Mastery criterion remained the same and both trainers achieved mastery on their first attempt. This session was 1.5 hours in duration.

**Administration of measures – Time 1 (Baseline testing).** The baseline administration of measures occurred 1 hour before the first module began. Participants provided informed consent for their participation (consent form available from the author on request) and then completed the following battery of measures through paper and pencil format: (1) demographic questionnaire; (2) knowledge assessments for each evidence-based practice module; (3) Attitudes to Evidence-
Based Practice Questionnaire (McKenna et al., 2004); (4) Maslach Burnout Inventory – Human Services Survey (MBI-HSS; Maslach et al., 1996); (5) Minnesota Satisfaction Questionnaire – Short Form (MSQ-SF; Weiss et al., 1967); (6) Short Version of the Occupational Self-Efficacy Scale (Rigotti et al., 2008); (7) Test of Knowledge (Denne et al., 2015), and (8) The Shortened Ways of Coping Questionnaire (SWC-R; Hatton & Emerson, 1995). Figure 13 (pg. 197) provides a timeline of study implementation.

**Intervention (Behaviour Skills Training).** Training was delivered across two service sites, with one trainer allocated to each group. Both groups received the same training package, which included modules on reinforcement, systematic prompting, functional communication training, and task analysis (in that order), using the Behaviour Skills Training model. Each of the four modules was divided into four sessions. The first session provided instruction and video modelling in the target practice; the second and third session covered rehearsal and feedback, and the final session facilitated a review of material covered. Training was scheduled from 9am-5pm, over three consecutive days.

**Session 1 – Instruction and modelling.** Participants were given a training pack corresponding to the evidence-based module that they were participating in (available from the author on request). The session began with the trainer providing didactic instruction, using a PowerPoint presentation. As part of this session, the trainer described the evidence-based practice, explained its importance and relevance and described how target skills should be performed. Questions and discussion were encouraged throughout. Once instruction was complete, the trainer demonstrated correct implementation of the evidence-based skill through video models and role-play, if requested by participants. This session lasted approximately 60 minutes.
Session 2 and 3 – Rehearsal and feedback (Reinforcement module: service site 1 only). The second and third sessions began with a brief review of the video model from the previous session. The trainer then read through the relevant mastery assessment data sheet, outlining each step (available from the author on request). Next, the participants formed dyads, with one of the participants in each pair acting the part of service user. These confederates were given a box containing six flash cards; ‘correct response’ was printed on three of the cards and ‘no response’ was printed on the remaining three cards. When the confederate was expected to respond, they selected a card from the box and followed its direction. On selection of a ‘correct response’ card, the service user confederate responded appropriately to the initial delivery of an instruction or prompt. On selection of a ‘no response’ card, the service user confederate did not respond to the initial delivery of an instruction or prompt. However, they could respond to the subsequent delivery of instructions and prompts. Cards were not returned to the box until all flashcards were used.

Participants were expected to implement the target skill (i.e., positive reinforcement) with the confederate service user, according to the steps outlined on the mastery assessment data sheet (available from the author on request). The trainer reviewed each dyad individually, recording participants’ performance using the designated mastery assessment data sheet. A single observation involved six trials; three trials with the service user confederate responding correctly and three trials with the service user confederate not responding. The trainer delivered positive and corrective feedback immediately following an observation. The trainer continued to review each dyad until the mastery criterion of 90% correct implementation across three consecutive observations was achieved or 100% correct implementation was shown on the first observation. The dyads remained the same for the third session.
However, participants switched roles; the participant who was observed in the previous session now acted the part of confederate. The second and third sessions were scheduled to be a maximum of 90 minutes each. Given the duration of these sessions, participants were provided with a reflective practice in their training packs. These could be completed before or after the trainer reviewed their performance.

**Session 2 and 3 – Rehearsal and feedback (Reinforcement module: service site 2; systematic prompting, functional communication training, and task analysis modules: service sites 1 and 2).** Mastery data for service site 1 showed rapid acquisition of the positive reinforcement skill. Seven of the eight participants acquired the target skill with the minimum amount of rehearsal necessary (i.e., reached mastery within three observations), indicating a degree of redundancy in the number of trials per observation. Therefore, following a review of the Behaviour Skills Training literature (e.g., Bishop & Kenzer, 2012; Dib & Sturmey, 2007; Lambert et al., 2013; Sarokoff & Sturmey, 2004; Suhrheinrich, 2011) and participant responses to the reinforcement module’s training acceptability survey, a decision was made to reduce trials from six to two per observation. This change was implemented for the systematic prompting, functional communication training, and task analysis modules for the first service site and across all four evidence-based practice modules for the second service site.

The second and third sessions continued to begin with a brief review of the video model from the previous session, an outline of the steps on the mastery assessment data sheet and designating participants into dyads. However, under the new procedure, service user confederates were now given a box containing two flash cards; ‘correct response’ was printed on one card and ‘no response’ was printed on the other card. As before, when the confederate was expected to respond, they
selected a card from the box and followed its direction (as outlined above).
Flashcards were not returned to the box until all were used. The procedure continued as outlined above, with the exception that a single observation now involved observing participants respond to one instance of the service user confederate correctly responding and one instance of the service user not responding. As an observation only involved two trials, it was decided that participants were required to demonstrate at least 90% correct implementation across three observations, regardless of whether or not they displayed 100% correct implementation on the first observation. As the number of trials per observation was decreased the sessions lasted approximately 60 minutes.

Session 4 – Review and summary. The final session facilitated the review of instructional material, video models and a question and answer session for participants. During the functional communication training module, trainers also used this session to support participants to develop communication plans for the service users they were working directly with. This was achieved by using a functional communication training plan template, already employed by the host organisation (available from the author on request). The review and summary sessions lasted a maximum of 60 minutes for each module.

Administration of measures - Time 2 (Post-intervention testing).
Immediately following each evidence-based training module, participants completed the knowledge assessment that was related to the module that had just concluded. They also completed a training acceptability survey (available from the author on request). On completion of the fourth module, participants were also requested to complete the Attitudes to Evidence-Based Practice Questionnaire (McKenna et al., 2004) and the Test of Knowledge (Denne et al., 2015).
Administration of measures – Time 3 (Follow-up testing). At 4-weeks post-intervention, participants were allocated 1 hour to complete the following measures in their typical work environment: (1) demographic questionnaire; (2) Knowledge assessments for each evidence-based practice module; (3) Attitudes to Evidence-Based Practice Questionnaire (McKenna et al., 2004); (4) Maslach Burnout Inventory – Human Services Survey (MBI-HSS; Maslach et al., 1996); (5) Minnesota Satisfaction Questionnaire – Short Form (MSQ-SF; Weiss et al., 1967); (6) Short Version of the Occupational Self-Efficacy Scale (Rigotti et al., 2008); (7) Test of Knowledge (Denne et al., 2015); (8) The Perceptions of Supervisory Support Scale (PSS; Fukui et al., 2014), and (9) The Shortened Ways of Coping Questionnaire (SWC-R; Hatton & Emerson, 1995).
Figure 13. Flow diagram showing the timeline of study implementation.

**Procedural fidelity.** Procedural fidelity data (available from the author on request) were collected for 100% of sessions during this study. Results indicated that the training package was being implemented with 100% fidelity by the first trainer and 99.2% by the second trainer. On two separate occasions during the feedback sessions, the second trainer miscalculated participants’ average correct performance, thereby incorrectly concluding that the participants had reached mastery criterion. The trainer was alerted to this by the experimenters during the session.

**Interobserver agreement.** Interobserver agreement (IOA) data were collected for 98.4% of reinforcement, 97.6% of systematic prompting, 96.6% of
functional communication training, and 98.8% of task analysis rehearsal sessions.
The trainer and one of the experimenters independently recorded the accuracy of the participants’ delivery of each of the target skills during role-play scenarios. A point-by-point method was used to calculate IOA. The number of agreements was divided by the number of agreements and disagreements. This figure was multiplied by 100 to establish the percentage IOA. The mean percentage of agreement ranged from 94.4% to 99.2%; reinforcement module (97.6%), systematic prompting module (99.2%), functional communication training module (97.5%), and task analysis module (94.4%).

A second interrater agreement analysis was conducted using the Kappa statistic (Cohen, 1960). According to Cohen (1960), the statistic can be interpreted as follows: < 0 indicates “no agreement”, 0.01-0.20 indicates “none to slight” agreement, 0.21-0.40 indicates “fair” agreement, 0.41-0.60 indicates “moderate” agreement, 0.61-0.80 indicates “substantial” agreement, and 0.81-1.00 indicates “almost perfect” agreement. In the current study, agreement among raters was characterised as “moderate” to “almost perfect”; reinforcement module (Kappa = 0.84, p<.001); systematic prompting module (Kappa = 0.92, p<.001); functional communication training module (Kappa = 0.78, p<.001), and the task analysis module (Kappa = 0.42, p<.001). Although the Kappa statistic recorded for the task analysis module was lower than values calculated for the other three modules, examination of the data using the point-to-point method (see above) indicated that raters agreed on 94.4% of items. Therefore, the reduced Kappa statistic was not a concern in this instance.

Data analysis.
**Missing data.** As previously discussed, incomplete data is a common problem encountered by researchers (Azar, 2002; Downey & King, 1998). Analysis of data in the present study showed no missing data across the measures administered at Time 1 (baseline) and Time 2 (post-intervention testing). However, it was not possible to gather data from the first service site and from three participants in the second service site at Time 3 (follow-up). Therefore, it was decided to analyse Time 1 (baseline) and Time 2 (post-intervention testing) data for all participants and separately examine the Time 1, Time 2 and Time 3 (follow-up) data for the six participants with complete data sets. Given the small sample number of participants with complete data sets, the assumptions of normality and equality of variances for a repeated-measures ANOVA were violated (Gravetter & Wallnau, 2004). As a result, the equivalent non-parametric test, the Friedman’s test, was conducted.

**Clustering.** Participants in this study were sampled from two separate service sites. As a result, there was a possibility that responses from individuals within one service site (cluster) would be more similar to one another than to responses from individuals in the second service site (Rutterford, Copas, & Eldridge, 2015). If this were true, response independence would be compromised. However, a series of Mann-Whitney tests (Appendix E) showed that the Time 1 (baseline) and Time 2 (post-intervention testing) knowledge scores for participants from the first service site did not differ significantly from the Time 1 and Time 2 scores for participants from the second service site. In addition, participants’ difference scores (i.e., the difference between the baseline and post-intervention scores for each participant) did not differ significantly between service sites.

A subsequent analysis of the interclass correlation (ICC), which controlled for the time of measurement, also showed minimal-to-no clustering of knowledge
scores within participants’ respective service sites. The ICC computed for responses on the reinforcement assessment was 5.4% (i.e., minimal) and no ICC could be calculated for the remaining knowledge assessments due to the absence of clustering within service sites. These results supported the assumption of response independence; participants’ scores did not appear to be clustered according to service site. The data satisfied the assumptions for implementing related-samples t-tests, including the independence of participant responses within each condition and the normal distribution of difference scores (Gravetter & Wallnau, 2004).

Likert data. Likert scale data are based on ranked categories, the intervals between scale items are not equal. Therefore, it is recommended that nonparametric statistical tests should be used to analyse this type of data (Jamieson, 2004).

Results

Participant training history. Figure 14 represents the number of participants who had previously received training across the four target evidence-based practices. Data indicate that the majority of participants had not received training previously; three participants had received training in reinforcement procedures, two in systematic prompting and functional communication training protocols, and one participant had been trained in task analysis implementation. None of these participants had received training delivered through a Behaviour Skills Training format.
Figure 14. The number of participants who received previous training in each of the four evidence-based practices (N=17).

**Knowledge acquisition.**

**Baseline (Time 1) and post-intervention (Time 2) testing.** Behaviour Skills Training resulted in an increase ($M = 18.2$, $SD = 17.4$) in the number of items answered correctly on the knowledge assessment for the reinforcement module. This increase was statistically significant, $t(16) = 4.32$, $p = .001$, $d = 1.05$.

Behaviour Skills Training resulted in an increase ($M = 34.1$, $SD = 20.01$) in the number of items answered correctly on the knowledge assessment for the systematic prompting module. This increase was statistically significant, $t(16) = 7.03$, $p < .001$, $d = 1.7$.

Behaviour Skills Training resulted in an increase ($M = 44.1$, $SD = 17.34$) in the number of items answered correctly on the knowledge assessment for the functional communication training module. This increase was statistically significant, $t(16) = 10.49$, $p < .001$, $d = 2.54$. 
Behaviour Skills Training resulted in an increase ($M = 27.6, SD = 19.53$) in the number of items answered correctly on the knowledge assessment for the task analysis module. This increase was statistically significant, $t(16) = 5.84, p < .001, d = 1.41$

Behaviour Skills Training resulted in an increase ($M = 10.3, SD = 8.56$) in the number of items answered correctly on the Test of Knowledge (Denne et al., 2015). This increase was statistically significant, $t(16) = 4.96, p < .001, d = 1.2$.

*Baseline (Time 1), post-intervention (Time 2) and follow-up (Time 3)*

**Testing.** A non-parametric Friedman’s test of differences among repeated measures was conducted and rendered a Chi-square value of 8.44, which represented a significant difference between the distributions of baseline ($M = 50, Mdn = 50$), post-intervention ($M = 70, Mdn = 70$) and follow-up ($M = 70.3, Mdn = 70$) scores on the reinforcement knowledge assessment ($p = .015$). Post-hoc analysis with a sign test was conducted with a Bonferroni correction applied, resulting in a significance level set at $p < .017$. There were no significant differences detected between baseline and post-intervention scores ($p = .063$), baseline and follow-up scores ($p = .031$) or post-intervention and follow-up scores ($p = .687$).

A non-parametric Friedman’s test of differences among repeated measures was conducted and rendered a Chi-square value of 9.82, which represented a significant difference between the distributions of baseline ($M = 31.7, Mdn = 35$), post-intervention ($M = 56.7, Mdn = 60$) and follow-up ($M = 53.3, Mdn = 55$) scores on the systematic prompting knowledge assessment ($p = .007$). Post-hoc analysis with a sign test was conducted with a Bonferroni correction applied, resulting in a significance level set at $p < .017$. There were no significant differences detected.
between baseline and post-intervention scores ($p = .031$), baseline and follow-up scores ($p = .031$) or post-intervention and follow-up scores ($p = 1.00$).

A non-parametric Friedman’s test of differences among repeated measures was conducted and rendered a Chi-square value of 9.82, which represented a significant difference between the distributions of baseline ($M = 48.3$, $Mdn = 50$), post-intervention ($M = 78.3$, $Mdn = 80$) and follow-up ($M = 70.3$, $Mdn = 80$) scores on the functional communication training knowledge assessment ($p = .007$). Post-hoc analysis with a sign test was conducted with a Bonferroni correction applied, resulting in a significance level set at $p < .017$. There were no significant differences detected between baseline and post-intervention scores ($p = .031$), baseline and follow-up scores ($p = .031$) or post-intervention and follow-up scores ($p = 1.00$).

A non-parametric Friedman’s test of differences among repeated measures was conducted and rendered a Chi-square value of 6.1, which represented a significant difference between the distributions of baseline ($M = 50$, $Mdn = 40$), post-intervention ($M = 80$, $Mdn = 80$) and follow-up ($M = 83.3$, $Mdn = 90$) scores on the task analysis knowledge assessment ($p = .047$). Post-hoc analysis with a sign test was conducted with a Bonferroni correction applied, resulting in a significance level set at $p < .017$. There were no significant differences detected between baseline and post-intervention scores ($p = .063$), baseline and follow-up scores ($p = .219$) or post-intervention and follow-up scores ($p = 1.00$).

A non-parametric Friedman’s test of differences among repeated measures was conducted and rendered a Chi-square value of 3.74, which represented no significant difference between the distributions of baseline ($M = 15$, $Mdn = 15$), post-
intervention \((M = 23.3, Mdn = 25)\) and follow-up \((M = 27.5, Mdn = 22.5)\) scores on
the Test of Knowledge \((p = .154)\).

**Target skill acquisition.** Figure 15 represents the mean number of
observations taken by participants to reach the mastery criterion during the rehearsal
and feedback component of each evidence-based practice module. The data for the
reinforcement module are presented according to service site because there were
differences in the data collection protocol used between the two service sites. The
mean number of observations to criterion was 2.4 (range = 1-4 observations) and 3
(range = 3 observations) for service site 1 and 2, respectively. The average number
of observations required for participants was 3.3 (range = 3-5 observations), 4.5
(range = 3-7 observations) and 3.3 (range = 3-5 observations) for the systematic
prompting, functional communication training and task analysis modules,
respectively. Table 41 displays the mean percentage of steps completed correctly on
the first observation and the final observation for each module. The data show that
following rehearsal and feedback, participants were achieving, on average,
procedural fidelity scores exceeding 96%, across all skills.
Figure 15. Mean number of observations taken by participants to reach mastery during rehearsal and feedback for each evidence-based practice module (N=17).

Table 41

<table>
<thead>
<tr>
<th>Evidence-Based Practice Modules</th>
<th>Mean Percentage of Steps Correctly Implemented on First Observation</th>
<th>Mean Percentage of Steps Correctly Implemented on Final Observation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reinforcement – Service 1</td>
<td>93.8% (85-100%)</td>
<td>98.8% (90-100%)</td>
</tr>
<tr>
<td>Reinforcement – Service 2</td>
<td>95% (90-100%)</td>
<td>96.3% (90-100%)</td>
</tr>
<tr>
<td>Prompting</td>
<td>90.8% (57-100%)</td>
<td>98.2% (90-100%)</td>
</tr>
<tr>
<td>Functional Communication Training</td>
<td>84.9% (63.5-100%)</td>
<td>99.5% (91.5-100%)</td>
</tr>
<tr>
<td>Task Analysis</td>
<td>94.7% (86-100%)</td>
<td>98.8% (93-100%)</td>
</tr>
</tbody>
</table>

Psychological well-being.

Maslach Burnout Inventory (MBI-HSS). The majority of the participants in this study scored low on the dimensions of exhaustion (low score = ≤17; n=14) and on depersonalisation (low score = ≤5; n=10), as measured by the Maslach Burnout Inventory during baseline testing (Time 1). In addition, a large portion of these
participants (n=11) reported experiencing moderate levels of achievement (i.e., scoring between 34 and 39, inclusive). These baseline data patterns are generally reflected in the results for the six participants, who provided both baseline (Time 1) and follow-up (Time 3) data. Consistent with the overall sample, all six participants reported moderate levels of achievement at baseline. However, at follow-up testing (Time 3), three participants scored low on levels of achievement (i.e., scoring 33 or below), two remained at a moderate level and one participant scored high (i.e., scoring 40 or above). Table 42 summarises these results.

Table 42

| Number of Participants from the Overall Sample Scoring Low, Moderate or High on the MBI-HSS Dimensions (N=17) and the Number of Participants, with Both Baseline and Follow-Up Data, who Scored Low, Moderate or High on the MBI-HSS Dimensions (N=6) |
|-------------------------------------------------|-------------------|-------------------|-------------------|
| Exhaustion (MBI)                                | Baseline Testing (N=17) | Baseline Testing (N=6) | Follow-Up Testing (N=6) |
| Low                                             | n=14              | n=6               | n=6               |
| Moderate                                        | n=2               | n=0               | n=0               |
| High                                            | n=1               | n=0               | n=0               |
| Depersonalisation (MBI)                         |                   |                   |                   |
| Low                                             | n=10              | n=6               | n=5               |
| Moderate                                        | n=6               | n=0               | n=1               |
| High                                            | n=1               | n=0               | n=0               |
| Achievement (MBI)                               |                   |                   |                   |
| Low                                             | n=2               | n=0               | n=3               |
| Moderate                                        | n=11              | n=6               | n=2               |
| High                                            | n=4               | n=0               | n=1               |

The results for the six participants who provided baseline (Time 1) and follow-up (Time 3) data were rank-ordered by the magnitude of the change in raw scores on each dimension of the Maslach Burnout Inventory, and a sign test was used to evaluate the data. The results did not show a significant change in exhaustion.
scores \( (p = .69) \), client depersonalisation scores \( (p = .63) \) or achievement scores \( (p = .63) \) between baseline (Time 1) and follow-up testing (Time 3).

**Minnesota Satisfaction Questionnaire (MSQ-SF).** Table 43 summarises the potential range of scores, the actual range of scores, and the mean and median scores for the overall sample at baseline (Time 1). This table also displays data for the six participants who completed this measure at follow-up testing (Time 3), as well as baseline testing (Time 1). A comparison of the potential range of scores with the actual range indicates that participants scored relatively high on this measure of job satisfaction.

The results for the six participants who provided baseline (Time 1) and follow-up (Time 3) data were rank-ordered by the magnitude of the change in raw scores on the Minnesota Satisfaction Questionnaire, and a sign test was used to evaluate the data. The results did not show a significant change in scores between baseline (Time 1) and follow-up testing (Time 3) \( (p = 1.00) \).

Table 43

| Potential Score Range, Actual Score Range and the Mean and Median Scores for the Overall Sample \( (N=17) \) and the Participants with Both Baseline and Follow-Up Data \( (N=6) \), on the Minnesota Satisfaction Questionnaire (MSQ-SF) |
|---|---|---|
| Potential score range | Baseline Testing \( (N=17) \) 20-100 | Baseline Testing \( (N=6) \) 20-100 | Follow-Up Testing \( (N=6) \) 20-100 |
| Actual score range | 60-94 | 63-89 | 73-85 |
| Mean score | 75.2 | 78.8 | 80.5 |
| Median score | 75 | 80.5 | 83 |

**The Short Version of the Occupational Self-Efficacy Scale.** Table 44 summarises the potential range of scores, the actual range of scores, and the mean and median scores for the overall sample at baseline (Time 1). This table also
Chapter 8

displays data for the six participants who completed this measure at follow-up testing (Time 3), as well as baseline testing (Time 1). A comparison of the potential range of scores with the actual range indicates that participants tended to score relatively high on this measure of occupational self-efficacy.

The results for the six participants who provided baseline (Time 1) and follow-up (Time 3) data were rank-ordered by the magnitude of the change in raw scores on the Short Version of the Occupational Self-Efficacy Scale, and a sign test was used to evaluate the data. The results did not show a significant change in scores between baseline (Time 1) and follow-up testing (Time 3) ($p = 1.00$).

Table 44

_Potential Score Range, Actual Score Range and the Mean and Median Scores for the Overall Sample (N=17) and the Participants with Both Baseline and Follow-Up Data (N=6), on the Short Version of the Occupational Self-Efficacy Scale_

<table>
<thead>
<tr>
<th>Potential score range</th>
<th>Baseline Testing (N=17) 6-36</th>
<th>Baseline Testing (N=6) 6-36</th>
<th>Follow-Up Testing (N=6) 6-36</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual score range</td>
<td>17-33</td>
<td>24-33</td>
<td>26-32</td>
</tr>
<tr>
<td>Mean score</td>
<td>26.5</td>
<td>28.8</td>
<td>29.3</td>
</tr>
<tr>
<td>Median score</td>
<td>27</td>
<td>28</td>
<td>29</td>
</tr>
</tbody>
</table>

*The Shortened Ways of Coping Questionnaire (SWC-R).* Table 45 summarises the potential range of scores, the actual range of scores, and the mean and median scores for the overall sample at baseline (Time 1). This table also displays data for the six participants who completed this measure at follow-up testing (Time 3), as well as baseline testing (Time 1). A review of the mean, median and...
range data shows that participants tended to report using a practical coping style, rather than the emotionally-based, wishful thinking style.

The results for the six participants who provided baseline (Time 1) and follow-up (Time 3) data were rank-ordered by the magnitude of the change in raw scores on the Shortened Ways of Coping Questionnaire, and a sign test was used to evaluate the data. The results did not show a significant change in ‘Wishful Thinking’ scores ($p = .063$) or ‘Practical Coping’ scores ($p = 1.00$) between baseline (Time 1) and follow-up testing (Time 3).

Table 45

<table>
<thead>
<tr>
<th>Potential Score Range, Actual Score Range and the Mean and Median Scores for the Overall Sample (N=17) and the Participants with Both Baseline and Follow-Up Data (N=6), on the Shortened Ways of Coping Questionnaire (SWC-R)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Wishful Thinking</strong></td>
</tr>
<tr>
<td>Potential score range</td>
</tr>
<tr>
<td>Baseline Testing (N=17)</td>
</tr>
<tr>
<td>Actual score range</td>
</tr>
<tr>
<td>Mean score</td>
</tr>
<tr>
<td>Median score</td>
</tr>
<tr>
<td>Follow-Up Testing (N=6)</td>
</tr>
<tr>
<td>Actual score range</td>
</tr>
<tr>
<td>Mean score</td>
</tr>
<tr>
<td>Median score</td>
</tr>
<tr>
<td><strong>Practical Coping</strong></td>
</tr>
<tr>
<td>Potential score range</td>
</tr>
<tr>
<td>Baseline Testing (N=17)</td>
</tr>
<tr>
<td>Actual score range</td>
</tr>
<tr>
<td>Mean score</td>
</tr>
<tr>
<td>Median score</td>
</tr>
<tr>
<td>Follow-Up Testing (N=6)</td>
</tr>
<tr>
<td>Actual score range</td>
</tr>
<tr>
<td>Mean score</td>
</tr>
<tr>
<td>Median score</td>
</tr>
</tbody>
</table>

Perception of Supervisory Support Scale (PSS). The Perception of Supervisory Support Scale was administered at the 4-week follow-up (Time 3) to increase the likelihood that participants would evaluate the supervisory support they received in the context of implementing evidence-based practices in the workplace. Table 46 provides a summary of the potential score range, actual score range and the mean and median scores for participants who completed the Perception of
Supervisory Support Scale. According to the results, participants perceived the supervisory support they received relatively positively; the potential score range on the scale was 0-114 and the mean score among participants was 78.5. In addition, there was no significant relationship between perceived supervisory support and follow-up scores on the reinforcement assessment, $r_s = -.42$, $n = 6$, $p = .407$, the systematic prompting assessment, $r_s = -.54$, $n = 6$, $p = .272$, the functional communication training assessment, $r_s = .02$, $n = 6$, $p = .978$, the task analysis assessment, $r_s = -.19$, $n = 6$, $p = .720$ or the Test of Knowledge, $r_s = -.52$, $n = 6$, $p = .288$.

Table 46

<table>
<thead>
<tr>
<th>Potential Score Range, Actual Score Range and the Mean and Median Scores for the Participants with Data on the Perceptions of Supervisory Support Scale (N=6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Follow-Up Testing (N=6)</td>
</tr>
<tr>
<td>Potential score range</td>
</tr>
<tr>
<td>Actual score range</td>
</tr>
<tr>
<td>Mean score</td>
</tr>
<tr>
<td>Median score</td>
</tr>
</tbody>
</table>

Social validity of the Behaviour Skills Training package. Table 47 (Appendix F) provides a summary of participant responding to each item on the training acceptability survey across each evidence-based practice module. The data show that participants evaluated the behaviour skill training package positively, in terms of its relevancy, organisation and potential for interaction and engagement. Table 48 summarises the potential score range, actual score range and the mean and median scores for participants. The potential score range on the training acceptability survey was 8-32 and the median rating score was at least 29 across all
modules. The functional communication training and task analysis modules recorded the highest mean and median scores.

Table 48

*Potential Score Range, Actual Score Range and the Mean and Median Scores for Participants’ Evaluations of the Four Evidence-Based Practice Modules (N=17)*

<table>
<thead>
<tr>
<th></th>
<th>Reinforcement</th>
<th>Prompting</th>
<th>Functional Communication Training</th>
<th>Task Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potential score range</td>
<td>8-32</td>
<td>8-32</td>
<td>8-32</td>
<td>8-32</td>
</tr>
<tr>
<td>Actual score range</td>
<td>22-32</td>
<td>21-32</td>
<td>21-32</td>
<td>24-32</td>
</tr>
<tr>
<td>Mean score</td>
<td>28.8</td>
<td>28.8</td>
<td>29.6</td>
<td>29.4</td>
</tr>
<tr>
<td>Median score</td>
<td>29</td>
<td>30</td>
<td>32</td>
<td>31</td>
</tr>
</tbody>
</table>

**Attitudes to Evidence-Based Practice Questionnaire.** Table 49 (Appendix G) summarises the percentage of participants selecting a specific response on each item of the Attitudes to Evidence-Based Practice Questionnaire at baseline (Time 1) and post-intervention testing (Time 2). Baseline results show that participants reported relatively positive attitudes towards evidence-based practice. All participants “agreed” or “strongly agreed” that implementing evidence-based practices would benefit them professionally and over 80% reported benefits to changing their approach according to recommendations from high quality research. However, over 50% of the participants reported a lack of confidence in their ability to evaluate available research, almost 50% reported experiencing difficulty when implementing change in the workplace and 82.4% reported that they would feel more confident if a professional skilled in research was available to provide them with relevant information.

Post-intervention (Time 2) data revealed changes to participants’ pattern of responding. A total of 82.3% of participants “agreed” or “strongly agreed” that
implementing evidence-based practices would benefit them professionally. While this represented a reduction of 17.6% from baseline levels (Time 1), it must be noted that no participants disagreed with this statement at post-intervention testing (Time 2). In addition, over 75% of participants reported benefits to changing their approach to approximate research recommendations, which represented a small reduction from baseline testing. Post-intervention, 76.4% of participants reported a lack of confidence in their ability to evaluate research. However, there was a reduction in the percentage of participants reporting that it would be difficult to implement change in the workplace (23.5%). Finally, 82% of participants reported that they would “feel more confident if a professional skilled in research” was available to provide them with relevant information. The follow-up data (Time 3) was not reported in this instance as it only represented six of the 17 participants, thereby rendering comparisons across time ineffective.

**Phase II – Skill Generalisation**

**Method**

**Participants.** Participants were the same 17 frontline staff who participated in Phase I.

**Trainers.** Trainers were the same two trainers who delivered training during Phase I.

**Experimenters.** The lead and second experimenters from Phase I also acted as experimenters in Phase II.

**Service users.** Eleven service users whose direct support staff had participated in the training sessions during Phase I participated in Phase II. To be considered a direct support staff (or “keyworker”) for a service user, staff must have known the
service user and had the opportunity to observe them in one or more environments for 3 months or more. Specialised consent forms were designed to facilitate service user understanding that their keyworker(s) would be observed by a trainer and experimenter while they worked with them (available from the author on request). Guardian consent was obtained for those service users unable to provide assent (consent form available from the author on request). All service users were at least 18 years of age and had an intellectual disability.

**Setting.** Sessions were conducted in settings, where the participants typically supported their service users. These settings included residential care and day support services.

**Measures.**

*Mastery assessment for functional communication training in the natural environment (available from the author on request).* The mastery assessment, used during the functional communication training rehearsal session, was adapted to record participants’ performance when implementing functional communication training with service users in the typical work environment.

**Design.** Observations to criterion was employed during the practical observations to measure the number of opportunities required by participants to reach a pre-determined level of proficiency/mastery in the implementation of the target functional communication skill with a service user in the typical work environment (Cooper et al., 2014).

**Procedure.** During the 4-week follow-up period, trainers observed the cohort of participants who they had trained, implementing functional communication training in their typical work environment. Participants were expected to implement
the practice with a key service user and received advanced notice of the observation. Immediately before the observation, the trainer and participant reviewed the content of the functional communication training plan that was developed during the final session of the functional communication training module. At this time, the trainer provided the participant with the mastery assessment data sheet and allowed them to read through it.

The trainer notified the participant when the recording period would begin and assessed the fidelity with which the participant implemented the steps outlined in the functional communication training mastery assessment (available from the author on request). An observation lasted a maximum of 10 minutes and based on this observation the trainer provided immediate positive and corrective feedback to the participant. To achieve mastery on the target skill, participants were required to display at least 90% accuracy in its implementation across one observation. An observation could require participants to implement relevant procedures if the service user responded: (1) appropriately; (2) did not respond, or (3) responded appropriately and did not respond within the 10 minutes allocated. Observations and feedback continued until mastery was achieved or a total of three observations had occurred. Feedback from trainers on the implementation of the other target evidence-based practices was not permitted during the 4-week follow-up period.

**Interobserver agreement.** Interobserver agreement was documented for 75% of sessions. The trainer and one of the experimenters independently recorded the accuracy of the participants’ delivery of functional communication training to service users. A point-by-point method was employed and the mean percentage of agreement was 98%. The Kappa statistic (Cohen, 1960) was also calculated, in order
to evaluate interrater reliability. Results indicated “substantial” agreement among raters (Kappa = 0.76, p < .001) (Cohen, 1960).

**Results**

Data showed that 12 participants (70.6%) displayed mastery of functional communication training with a service user in their typical work setting on the first observation. One participant did not reach mastery after three observations and was identified to receive additional training after the 4-week follow-up period had elapsed. Three participants had no opportunity to display the target skill due to an absence of motivation on the part of their service user to request. One participant was unavailable for observation.

**Discussion**

The overall goal of the pilot study was to evaluate the practicalities of implementation and “build the foundation for the planned intervention study” (Tickle-Degnen, 2013, p. 171). The first objective was to test the feasibility of the experimental design, methodology and logistics of delivery, ahead of full-scale implementation in the randomised control trial. The pilot study demonstrated that the evidence-based Behaviour Skills Training package could be successfully delivered within the period of time allocated by the host organisation. Results also showed that the experimental design and associated methodology were relatively sensitive to the effects of the training intervention.

However, limitations were observed during the pilot study and addressed, in order to minimise any threats to the successful implementation and validity of the randomised control trial (Tickle-Degnen, 2013). Alterations were made to the training content and the criterion required to achieve mastery in the application of the
target skills during rehearsal sessions. The failure to examine the impact of the indirect effects of the training intervention on service users was also addressed.

While results showed significant increases in levels of target knowledge and skill acquisition among participants in the pilot study, training content was supplemented to reduce the likelihood that participants would be selecting “the correct response without any real understanding” (Gajjar, Sharma, Kumar, & Rana, 2014, p. 18) and to maximise learning outcomes during the randomised control trial. According to van Oorsouw et al. (2009), the acquisition of novel information can be optimised by targeting multiple learning modalities. Therefore, in an attempt to target both the auditory and visual modalities, the PowerPoint presentations for each module were visually augmented with additional information from the training manuals for the purpose of the randomised control trial (available from the author on request).

A second alteration to content focused on the assessment of skill development. During the evaluation of skill generalisation (Phase II), it was noted that half of the participants implemented functional communication training targeting escape-motivated behaviour. While this skill had been discussed during the instruction component of training, it had not been modelled or rehearsed as part of the training package. Therefore, in order to enhance the social and external validity of the training programme, it was decided to include functional communication training to target escape-motivated behaviour in the video modelling and rehearsal sessions during the randomised control trial.

The change to the mastery criterion, which was mentioned above, was introduced during the roll-out of the pilot study. Appropriate implementation of
positive reinforcement procedures was rapidly acquired by participants in the first group exposed to the training protocol; seven of the eight participants acquired the skill with minimal rehearsal, achieving mastery within three observations. This pattern of skill acquisition highlighted a degree of redundancy in the number of trials per observation. In addition, analysis of the training acceptability survey for the reinforcement module revealed that participants had reported that the number of trials per observation was excessive. Therefore, following a review of the Behaviour Skills Training literature (e.g., Bishop & Kenzer, 2012; Dib & Sturmey, 2007; Lambert et al., 2013; Sarokoff & Sturmey, 2004; Suhrheinrich, 2011) the number of trials per observation was reduced from six to two. However, the overall mastery criterion continued as 90% correct implementation of the target skill over three consecutive observations. This methodological change was implemented for the systematic prompting, functional communication training, and task analysis modules for the first service site and for all modules for the second service site and was employed during the randomised control trial.

Finally, the impact of the training intervention on service user outcomes was not evaluated during the pilot study. According to Jahr (1998), the competency of frontline staff, employed in the intellectual disability sector, significantly influences the behaviour and quality of life of the service users they support. Therefore, it was considered necessary to include an assessment of service user ability to achieve a comprehensive and socially valid evaluation of this evidence-based Behaviour Skills Training package during the randomised control trial.

The second objective of the pilot study was to examine the primary outcomes of the study (i.e., knowledge acquisition, skill development, and psychological well-being), in terms of measurability, degree of change, and potential barriers to
dissemination in the randomised control trial. Firstly, results from the pilot study demonstrated that the measures employed provided adequate data to effectively analyse the impact of the training intervention on participants’ target behavioural knowledge. Across all participants, scores on the Test of Knowledge and the knowledge assessments for each training module improved significantly from baseline (Time 1) to post-intervention testing (Time 2). This increase represented a large effect size for all knowledge acquisition measures.

However, for the six participants with follow-up (Time 3) data, post hoc tests did not reveal any significant differences between baseline (Time 1) and post-intervention (Time 2) scores, baseline (Time 1) and follow-up (Time 3) scores, and post-intervention (Time 2) and follow-up (Time 3) scores. Nonetheless, analyses showed that the increase in scores from baseline to post-intervention testing and from baseline to follow-up testing, typically approached significance. Furthermore, there was little evidence of differences between post-intervention and follow-up scores. This pattern suggests that there was a degree of knowledge acquisition and subsequent maintenance over the 4-week follow-up period. The failure to isolate significant effects was most likely due to inadequate statistical power; a direct result of the poor response rate at follow-up. Restricted statistical power may also account for the absence of change in staff-related measures of psychological well-being (i.e., burnout, occupational self-efficacy, job satisfaction, and coping style). However, no firm conclusion can be drawn at this point.

As the purpose of the knowledge assessments in the current study was to measure participants’ knowledge about specific evidence-based subject matter pre- and post-intervention, an edumetric rather than a psychometric evaluation was conducted. According to Carver (1974), measures can be evaluated in terms of either
their psychometric properties or their edumetric properties. A psychometric analysis investigates the degree to which a measure reflects “the stable between-individual differences that traditionally have been of interest to psychological testing”, while a edumetric analysis examines “the extent to which it [the measure] reflects the within-individual growth” (Carver, 1974, p. 512). As a result, an edumetric analysis is more suited to situations, in which knowledge is being assessed before and after training, as a psychometric evaluation may result in a test that is insensitive to individual gain (Carver, 1974).

For example, when assessing knowledge gains pre-and post-training, Cronbach alpha coefficients may render any gains as unreliable, due to low alpha coefficients pre-training (indicating no knowledge on the subject matter and a tendency to answer randomly) and undefined alpha coefficients post-training (indicating that participants are providing correct answers for the majority of items). However, in edumetric terms, adequate validity and reliability would be demonstrated if results showed substantial within-individual gains from pre-training to post-training and consistent scores on the measure through repeated testing under equivalent conditions. Therefore, it is argued that gain scores on a psychometrically developed measure may prove inappropriate, and the use of an edumetrically developed measure is prudent when gain scores are the primary focus (Carver, 1974).

In the context of the current study, analyses found the knowledge assessments to be edumetrically valid and reliable. In addition, items were generally pitched at acceptable levels of difficulty. Items, which were characterised as more or less difficult than optimal levels (i.e., percentage of participants answering correctly was less than 20% or greater than 80%), were retained in the assessments for the purpose of the randomised control trial, as reference to the relevant literature (AFIRM Team,
Chapter 8

2015a; 2015b; 2015c; Cooper et al., 2014; Franzone, 2009; National Autism Center, 2009; 2015; Neitzel, 2009; Nietzel & Wolery, 2009; Wong et al., 2014) indicated that omission of these items may compromise the content validity of the measures (Parmenter & Wardle, 2000). In addition, the average difficulty level of items on each knowledge assessment fell within the optimal range. It must be noted that follow-up data were not included in the evaluation of item difficulty or edumetric validity; the response rate was extremely limited, thereby necessitating steps to be taken by the host organisation to minimise this threat to validity during the randomised control trial.

With regards to effectively evaluating skill development, findings from the rehearsal sessions were positive. Results showed that participants mastered implementation of the target evidence-based practice skills relatively quickly. On average, five observations were required for participants to achieve mastery in the implementation of the target skill in the functional communication training module and this was the maximum number of observations required across the four training modules. In addition, participants readily transferred the functional communication training procedures to the typical work environment. A total of 12 participants (70.6%) achieved fidelity in the application of functional communication training with a service user on the first observation. Three of the participants did not have an opportunity to display the skill and only one participant did not achieve mastery and was identified as requiring additional support.

Successful generalisation of the functional communication training skills may be due to the fact that participants consistently rated all training modules and associated practices very favourably. According to Callahan et al. (2008) and Callahan et al. (2016) professionals, including frontline staff, are more likely to
implement practices, which they perceive as socially valid. In addition, participants were asked to complete the Perceptions of Supervisory Support Scale and the Attitudes to Evidence-Based Practice Questionnaire to help identify potential barriers to the dissemination of the targeted evidence-based practices. However, barriers to dissemination did not appear to be an issue in the pilot study, given the participant performance across the knowledge assessments, rehearsal sessions and practical observations in the real-world setting.

The pilot study aimed to test the feasibility of implementing the evidence-based Behaviour skills Training package and associated measures, within a subsequent randomised control trial. Despite changes to training content, the criterion required to achieve mastery in the application of the target skills, and the assessment of service user outcomes, the method of dissemination and training content remained largely unaltered for the randomised control trial. The degree of change in participants’ behavioural knowledge acquisition was used to guide sample size calculation for the randomised control trial and the objective of the subsequent study was the large-scale evaluation of the Behaviour Skills Training package, trialed in the pilot study, as a method of disseminating the four target evidence-practices to frontline staff in the RehabCare organisation.
Chapter 9

Dissemination of Evidence-Based Practice to Frontline Staff Working in the Field of Intellectual Disability: A Randomised Control Trial Study
Abstract

Although there is strength of evidence and empirical support for Behaviour Skills Training, this intervention has yet to be evaluated in a large group-based experimental design. Given that randomised control trials are the most reliable experimental design in determining the effectiveness of an intervention (Akobeng, 2005), the goal of the current study was to conduct a clustered randomised control trial to comprehensively evaluate if Behaviour Skills Training was an effective method for disseminating evidence-based practice to frontline staff working with adults with intellectual disabilities. One hundred and four frontline staff were recruited from twelve service sites within the RehabCare organisation. A total of 54 participants were assigned to the Behaviour Skills Training condition. This condition combined instruction, modelling, rehearsal, and feedback within a Behaviour Skills Training intervention (Ward-Horner & Sturmey, 2012) and was used to coach participants in reinforcement, systematic prompting, functional communication training, and task analysis. Fifty participants were assigned to the wait list control condition. Results showed that participants who received Behaviour Skills Training demonstrated statistically significant improvements on four knowledge assessments and the Test of Knowledge at the 4-week follow-up. In contrast, participants in the wait list control condition showed either no change or a statistically significant decrease in their knowledge scores over the study period. No statistically significant changes to psychological well-being were observed for either experimental group and no salient barriers to dissemination were reported by participants.
Although Behaviour Skills Training has proven effective in training a broad selection of professionals, including frontline staff, in a diverse range of work-related skills (e.g., Homlitas et al., 2014; Maggin et al., 2012; Parsons et al., 2012; Sarokoff & Sturmey, 2008), the efficacy of this training method has yet to be tested in a large group-based experimental design. Furthermore, until now the effectiveness of this training approach has never been evaluated in the context of training staff working in the intellectual disability sector, in the following combination of evidence-based practices: reinforcement, systematic prompting, functional communication training, and task analysis.

The randomised control trial is considered the most reliable experimental design for analysing and determining the impact of an intervention (Akobeng, 2005). However, this type of design can be costly, in terms of time commitment and financial investment (Haskins & Feldman, 2016; Liu & Wyatt, 2011). Therefore, it is recommended that a pilot study be carried out, prior to a randomised control trial, to determine the practicalities of implementation (Tickle-Degnen, 2013). As such, the purpose of the study outlined in Chapter 8, was to assess the viability of implementing a randomised control trial, which would analyse the impact of an evidence-based Behaviour Skills Training intervention on frontline staff, working in the intellectual disability sector. Results showed that the experimental design and associated methodology were robust in measuring outcomes related to staff participants, and the intervention, itself, could be successfully delivered within the allocated time-period.

However, limitations were observed during the pilot study and procedural adjustments were developed qualitatively, through reflection on data trends and feedback from participants’ training acceptability questionnaires. While significant
levels of target knowledge and skill acquisition were observed, training content was supplemented, following the pilot study, to maximise learning outcomes during the randomised control trial. In addition, the mastery criterion of target skills was altered and this procedural change was outlined in detail in Chapter 8.

With results from the pilot study supporting the practicalities of implementation, the goal of the current study was to conduct a randomised control trial to comprehensively evaluate if Behaviour Skills Training was an effective method for disseminating evidence-based practice to frontline staff, working with adults with intellectual disabilities. In a traditional randomised control trial, participants sampled from a larger population, are randomly assigned to one of two or more experimental conditions and monitored for a predetermined period. With the exception of the intervention being evaluated, participants across all experimental groups are treated identically. On completion of the study, participant groups are compared using preselected outcomes. If the experimental groups have been treated identically, apart from exposure to the target intervention, any difference in outcomes can be confidently attributed to the impact of the intervention (Akobeng, 2005).

Cluster randomisation, whereby groups rather than individuals are randomised (Donner & Klar, 2004; Rutterford et al., 2015), was employed in the current study. Recruited service sites were randomly assigned to a “Behaviour Skills Training” condition or to a wait list control condition. The principle advantages of adopting this approach were as follows: (1) the intervention was “naturally applicable to the cluster” (Wears, 2002, p. 330); (2) training discrete service sites within an intellectual disability provider was more practical and efficient than training individual staff members (Wears, 2002), and (3) active engagement from participants was considered more likely (Donner & Klar, 2004).
The first objective of the study was to examine whether the Behaviour Skills Training intervention would positively impact participants’ knowledge of the target evidence-based practices. It was hypothesised that when compared to baseline (Time 1) assessments, participants in the Behaviour Skills Training condition would score significantly higher on knowledge assessments at Time 3 (follow-up testing). However, this outcome was not expected for participants in the wait list control condition. The second objective was to identify potential barriers to the effective dissemination of evidence-based practice, by examining participants’ perception of the supervisory support they receive, as well as their attitudes towards evidence-based practice.

The final objective was to examine participants’ psychological well-being over the course of the study. According to Maslach (1999), the demands associated with working with people with an intellectual disability are often particularly stressful for frontline staff. Prolonged exposure to stress is associated with poorer psychological outcomes, including lower levels of job satisfaction and a reliance on wishful thinking. Work-stress theories propose that an individual’s perception of work-related demands, plays an important role in the development of stress and burnout. However, the trajectory of this relationship can be altered depending on the nature of the organisational resources available to an individual, including effective staff training (Devereux, Hastings, & Noone, 2009; Lazarus, 1999; Lazarus & Folkman, 1984). Therefore, it was hypothesised that exposure to the Behaviour Skills Training package would lead to an improvement in measures of burnout, job satisfaction, self-efficacy, and the use of effective coping strategies. However, it was hypothesised that a similar change would not be observed for participants in the wait list control condition. Finally, given the time constraints governing the current
programme of research, participants in the wait list control condition were designated to receive the Behaviour Skills Training intervention outside the time-frame allocated for the current study.

Method

Service Sites

Twelve service sites (i.e., clusters) were successfully recruited for the randomised control trial. All 12 service sites were recruited from RehabCare, the host organisation that participated in the training needs analysis and pilot studies, which were outlined in Chapters 6 and 8, respectively. Six of the 12 service sites were randomly assigned to the “Behaviour Skills Training” intervention condition and six service sites were randomly assigned to the wait list control condition. Within each experimental condition four service sites provided residential support to service users, while two offered support through day-based services.

Participants

One hundred and four frontline staff were recruited from the twelve service sites targeted for the clustered randomised control trial. A total of 54 participants were assigned to the intervention condition (i.e., Behaviour Skills Training) and 50 participants were assigned to the control condition (i.e., wait list control).

Eighty-two participants were female ($n=41$ intervention; $n=41$ control) and 77 participants listed their job title as careworker ($n=27$ intervention; $n=27$ control) or programme facilitator ($n=10$ intervention; $n=13$ control). Forty-three participants had achieved an undergraduate degree ($n=23$ intervention; $n=20$ control) and 18 had earned a postgraduate qualification ($n=9$ intervention; $n=9$ control). Overall, 65 participants had been employed with the RehabCare organisation for 5 years or less
(n=30 intervention; n=35 control), 20 participants (n=14 intervention; n=6 control) reported an employment period of 5-10 years, and 19 participants (n=10 intervention; n=9 control) reported an employment period of 10+ years. Table 50 characterises participants according to service site and length of service with the host organisation. Table 51 characterises participants according to gender, education and job title for the intervention and control groups, respectively.

There was no significant relationship between experimental allocation and gender ($\chi^2 (1, n = 104) = 0.57, p = .449$), level of education (i.e., pre-undergraduate degree & other; undergraduate degree; postgraduate qualification) ($\chi^2 (2, n = 102) = 0.08, p = .963$) or length of service within the host organisation ($\chi^2 (3, n = 104) = 4.91, p = .179$). As such, the participants in the two experimental conditions were comparable.
### Table 50

**Participant Characteristics According to Service Site and Length of Service with the Host Organisation (N=104)**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Service Site</th>
<th>Service Setting</th>
<th>Number of Participants</th>
<th>Length of Service</th>
<th>Number of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention 1</td>
<td>Day</td>
<td>7</td>
<td>&lt; 1 year</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1-5 years</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5-10 years</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10+ years</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Intervention 3</td>
<td>Residential</td>
<td>7</td>
<td>&lt; 1 year</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1-5 years</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5-10 years</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10+ years</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Intervention 5</td>
<td>Residential</td>
<td>9</td>
<td>&lt; 1 year</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1-5 years</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5-10 years</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10+ years</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Intervention 7</td>
<td>Residential</td>
<td>9</td>
<td>&lt; 1 year</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1-5 years</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5-10 years</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10+ years</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Intervention 9</td>
<td>Residential</td>
<td>9</td>
<td>&lt; 1 year</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1-5 years</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5-10 years</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10+ years</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Intervention 11</td>
<td>Day</td>
<td>13</td>
<td>&lt; 1 year</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1-5 years</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5-10 years</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10+ years</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Control 2</td>
<td>Day</td>
<td>10</td>
<td>&lt; 1 year</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1-5 years</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5-10 years</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10+ years</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Control 4</td>
<td>Residential</td>
<td>5</td>
<td>&lt; 1 year</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1-5 years</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5-10 years</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10+ years</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Control 6</td>
<td>Residential</td>
<td>11</td>
<td>&lt; 1 year</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1-5 years</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5-10 years</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10+ years</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Control 8</td>
<td>Residential</td>
<td>9</td>
<td>&lt; 1 year</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1-5 years</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5-10 years</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10+ years</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Control 10</td>
<td>Residential</td>
<td>8</td>
<td>&lt; 1 year</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1-5 years</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5-10 years</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10+ years</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Control 12</td>
<td>Day</td>
<td>7</td>
<td>&lt; 1 year</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1-5 years</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>
Table 51

Summary of Participant Characteristics: Upper Panel - Gender, Education, and Job Title for the Intervention Group (N=54); Lower Panel - Gender, Education, and Job Title for the Control Group (N=50)

<table>
<thead>
<tr>
<th>Gender</th>
<th>Number of Participants</th>
<th>Education</th>
<th>Number of Participants</th>
<th>Job Title</th>
<th>Number of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>13 (24.1%)</td>
<td>Leaving cert. Certificate</td>
<td>3 (5.6%)</td>
<td>Careworker</td>
<td>27 (50%)</td>
</tr>
<tr>
<td>Female</td>
<td>41 (75.9%)</td>
<td></td>
<td>4 (7.4%)</td>
<td>Supervisor</td>
<td>10 (18.5%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diploma</td>
<td>13 (24.1%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Undergraduate degree</td>
<td>23 (42.6%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Postgraduate qualification</td>
<td>9 (16.7%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Other</td>
<td>2 (3.7%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>9 (18%)</td>
<td>Leaving cert. Certificate</td>
<td>1 (2%)</td>
<td>Careworker</td>
<td>27 (54%)</td>
</tr>
<tr>
<td>Female</td>
<td>41 (82%)</td>
<td></td>
<td>5 (10%)</td>
<td>Supervisor</td>
<td>13 (26%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diploma</td>
<td>11 (22%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Undergraduate degree</td>
<td>20 (40%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Postgraduate degree</td>
<td>9 (18%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Other</td>
<td>2 (4%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Missing</td>
<td>2 (4%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Experimenters

The two individuals who acted as experimenters in the pilot study (Chapter 8) remained in these roles for the duration of the randomised control trial.

Trainers
**Intervention condition only.** Four behaviour therapists, recruited from the host organisation, acted as trainers for the Behaviour Skills Training condition. Three of the four trainers held postgraduate qualifications in Applied Behaviour Analysis and two were Board Certified Behaviour Analysts (BCBA®). The fourth trainer held a Masters level qualification in Analysis and Intervention in Intellectual and Developmental Disabilities.

**Setting**

**Intervention condition.**

**Trainer instruction.** A 1-day instructional session for trainers, who had not participated in the pilot study reported in Chapter 8, was hosted by the service provider organisation in a classroom setting. The room was large enough to deliver lecture-based instruction to two participants and facilitate in-vivo skill demonstration and role-play rehearsal of the target skills.

**Administration of measures.** Measures were administered in a conference room, classroom or office setting.

**Behaviour Skills Training.** Six 3-day training sessions were delivered nationwide. Each session was hosted by the intellectual disability service provider, in either a conference room or classroom setting. The rooms were large enough to seat and deliver didactic instruction via projector to 10-15 participants. There was also adequate space to accommodate role-play rehearsal of the target skills (i.e., sufficient space for five-seven dyads) and manage the potential for uncontrolled observations among participants.

**Control condition.**
**Administration of measures.** All measures were administered in an office setting.

**Materials**

*Intervention condition only.*

*Trainer instruction.*

*Trainer pack.* The material provided in the trainers’ instructional packs was identical to that supplied in the pilot study reported in Chapter 8. Relevant content had been updated in accordance with recommendations from the pilot study (materials available from the author on request).

*Behaviour Skills Training.*

*Evidence-based practice training packs.* These packs contained the same material as that provided to participants in the pilot study reported in Chapter 8. This material had been supplemented according to recommendations from the pilot study (training packs available from the author on request).

*PowerPoint presentations.* The general structure and format of the PowerPoint presentations remained the same for this study. The content of all four presentations was visually augmented, in line with recommendations from the pilot study (PowerPoint presentations available from the author on request).

*Rehearsal props.* As in the pilot study, items that would typically be available at each service site (e.g., dish cloths, cups, cutlery, hairbrush, face cloths, food) were used as props during rehearsal sessions.
Measures

The battery of measures was identical to those used during the pilot study reported in Chapter 8: (1) demographic questionnaire; (2) knowledge assessments for each evidence-based practice module; (3) mastery assessments for each evidence-based practice module; (4) Training Acceptability Survey for each evidence-based practice module; (5) Attitudes to Evidence-Based Practice Questionnaire (McKenna et al., 2004); (6) Maslach Burnout Inventory – Human Services Survey (MBI-HSS; Maslach et al., 1996); (7) Minnesota Satisfaction Questionnaire – Short Form (MSQ-SF; Weiss et al., 1967); (8) Short Version of the Occupational Self-Efficacy Scale (Rigotti et al., 2008); (9) Test of Knowledge (Denne et al., 2015); (10) The Perceptions of Supervisory Support Scale (PSS; Fukui et al., 2014), and (11) The Shortened Ways of Coping Questionnaire (SWC-R; Hatton & Emerson, 1995).

Knowledge assessments for each evidence-based practice module (available from the author on request). The knowledge assessments were identical to those employed during the pilot study reported in Chapter 8 and were developed by the experimenter using the content of the National Professional Development Centre online manuals and associated quizzes (AFIRM Team, 2015a; 2015b; 2015c; Franzone, 2009). The purpose of the knowledge assessments was to assess potential knowledge gain following exposure to Behaviour Skills Training. Therefore, consistent with the approach taken in Chapter 8, an edumetric analysis rather than a psychometric analysis was conducted on the four knowledge assessments (Carver, 1974). This analysis examined: (1) edumetric validity; (i.e., the degree to which a measure captures gain in situations where it is expected), and (2) edumetric reliability (i.e., consistency of gain as reflected by the measure).
Edumetric validity. Edumetric validity for all knowledge assessments was demonstrated. Participants exposed to the Behaviour Skills Training intervention scored significantly higher on the knowledge assessments, following completion of the relevant evidence-based practice module. This pattern was not observed for participants assigned to the wait list control condition (Carver, 1974; Deniz & Alsaffar, 2013). Detailed analyses can be found in the results section.

Edumetric reliability. Edumetric reliability can be evaluated by administering a measure on two occasions, under equivalent treatment conditions. A measure is considered edumetrically reliable if it demonstrates consistency of gains for an individual across the two assessments (Carver, 1974). In the current study, edumetric reliability was demonstrated for the reinforcement, functional communication training, and task analysis knowledge assessments; scores from participants in the Behaviour Skills Training intervention did not change significantly from post-intervention (Time 2) to follow-up testing (Time 3). The fact that there was no significant change between Time 2 and Time 3 indicates that the knowledge assessments were consistent in detecting knowledge gain after training. Between Time 2 and Time 3, there was a statistically significant decrease in knowledge scores on the systematic prompting assessment. While this decrease was relatively small (7.69%), it does compromise the edumetric reliability of the measure. Detailed analyses can be found in the results section.

Design

A clustered randomised control trial, which employed a wait list control condition, was used to evaluate the differential impact of a Behaviour Skill Training
package, in terms of disseminating evidence-based practices to frontline staff and improving the psychological well-being of these staff.

**Procedure**

**Randomisation of service sites to conditions.** Twelve service sites were recruited to participate in the randomised control trial. Eight service sites provided residential care and four provided day service support. Specifically, the matching process involved randomly allocating a residential care service to the intervention condition and a subsequent residential care service to the control condition. Each of these residential care services were also matched on geographical location (i.e., they were in the same province). This was repeated for day service support sites. In total, two day service supports sites and four residential care services received intervention. The remaining sites were allocated to the control condition.

Randomisation to conditions and the matching process involved writing the name of a single service, the type of support provided (i.e., day or residential service) and its geographic location (based on regions within the host organisation) on a card and repeating until the details of all services were written on individual cards. The cards of service sites providing the same type of support, within the same geographic location, were placed in a single pile, shuffled and the first card was taken from the top of the pile. The lead experimenter performed a coin toss and noted on the chosen card whether it was ‘heads’ or ‘tails’. The next card was then taken from the pile and was assigned the opposite (e.g., if the first card was assigned ‘heads’, then the next card was automatically assigned ‘tails’). This constituted a matched pair of service sites.
This process was repeated until all service sites were matched. An individual, blind to the nature of the study and the participating service sites, independently decided that all service sites designated as ‘tails’ should be allocated to the control condition, and all service sites designated as ‘heads’ should be allocated to the intervention condition.

**Intervention condition.** Figure 16 (pg. 240) provides a timeline of study implementation for the intervention group.

*Trainer instruction.* The instruction programme provided was identical to that outlined in the pilot study reported in Chapter 8.

**Administration of measures – Time 1 (Baseline testing).** Participants completed the same battery of measures as participants in the control condition and the measures were administered according to the procedures outlined in the pilot study reported in Chapter 8.

**Intervention (Behaviour Skills Training).** Behaviour Skills Training was delivered across six service sites, with one trainer allocated to each group. All groups received the same training package and the training schedule, format and structure were identical to those employed in the pilot study reported in Chapter 8. Each of the four training modules was divided into four sessions: The first provided instruction and video modelling in the target practice; the second and third session covered rehearsal and feedback; and the final session facilitated a review of material covered. Training was scheduled from 9am-5pm, over three consecutive days.

*Session 1 – Instruction and modelling.* The procedures employed were identical to those used in the pilot study reported in Chapter 8.
Session 2 and 3 – Rehearsal and feedback. As previously outlined in Chapter 8, a procedural alteration to the administration of the rehearsal and feedback sessions was made during the pilot study. This modified procedure was maintained for the current study. As such, the second and third sessions began with a brief review of the video model from the previous session, an outline of the steps on the mastery assessment data sheet, and the designation of participants into dyads. Service user confederates were given a box containing two flash cards; ‘correct response’ was printed on one card and ‘no response’ was printed on the other card. When the confederate was expected to respond, they selected a card from the box and followed its direction (i.e., respond correctly or give no response). Cards were not returned to the box until all flashcards were used.

Participants were expected to implement the selected skill (e.g., positive reinforcement) with the confederate service user. The trainer reviewed each dyad individually, evaluating the trainee’s performance using the relevant mastery assessment data sheet (available from the author on request). Based on this assessment, the trainer provided immediate positive and corrective feedback. The trainer continued to review each dyad until the mastery criterion of 90% correct implementation across three consecutive observations was achieved. A single observation involved observing the participant respond to one instance of correct responding and one instance of incorrect responding from the service user confederate.

The dyads remained the same for Session 3. However, the participants switched roles, with the confederate from Session 2 becoming the trainee in Session 3 and vice versa. Sessions 2 and 3 were scheduled to be a maximum of 60 minutes each. Given the duration of these sessions, participants were provided with a
reflective practice in their training packs. These could be completed before or after
the trainers reviewed their performance.

**Session 4 – Review and summary.** The procedures employed were identical
to those employed in the pilot study.

**Administration of measures - Time 2 (Post-intervention testing).**
Immediately following each evidence-based training module, participants completed
the knowledge assessment that was related to the module that had just concluded.
They also completed a training acceptability survey (available from the author on
request). On completion of the fourth module, participants were also requested to
complete the Attitudes to Evidence-Based Practice Questionnaire (McKenna et al.,
2004) and the Test of Knowledge (Denne et al., 2015).

**Administration of measures – Time 3 (Follow-up testing).** Four weeks post-
intervention, participants in the intervention condition were allocated 1-hour to
complete the same battery of measures that they had completed at baseline (Time 1).
Participants also completed “The Perceptions of Supervisory Support Scale” (PSS;
Fukui et al., 2014) at this time.

**Skill generalisation.** The procedures followed were identical to those
employed in Phase II of the pilot study, which was reported in Chapter 8.

**Control condition.** Figure 16 (pg. 240) provides a timeline of study
implementation for the control group.

**Administration of measures – Time 1 (Baseline testing).** One week prior to
their matched service site receiving the Behaviour Skills Training intervention,
participants in the control condition provided informed consent for their participation
(consent form available from the author on request) and then completed the
following battery of questionnaires through paper and pencil format: (1) demographic questionnaire; (2) Knowledge assessments for each evidence-based practice module; (3) Attitudes to Evidence-Based Practice Questionnaire (McKenna et al., 2004); (4) Maslach Burnout Inventory – Human Services Survey (MBI-HSS; Maslach et al., 1996); (5) Minnesota Satisfaction Questionnaire – Short Form (MSQ-SF; Weiss et al., 1967); (6) Short Version of the Occupational Self-Efficacy Scale (Rigotti et al., 2008); (7) Test of Knowledge (Denne et al., 2015), and (8) The Shortened Ways of Coping Questionnaire (SWC-R; Hatton & Emerson, 1995). Each participant was given 1 hour for this assignment.

*Administration of measures – Time 3 (testing).* Four weeks after their matched service site had received the Behaviour Skills Training intervention, participants in the control condition were allocated 1-hour to complete the same battery of measures that they had completed at baseline (Time 1). Participants also completed “The Perceptions of Supervisory Support Scale” (PSS; Fukui et al., 2014) at this time.
Procedural Fidelity

**Intervention condition only.** Procedural fidelity data were collected for 100% of intervention sessions during this study and focused on the trainers’ accurate implementation of all components of the training package, across the four evidence-based practice modules. Results indicated that the intervention training package was being implemented with 100% fidelity by trainer 1 (service site 1), 100% by trainer 2 (service site 3), 99.4% by trainer 3 (service site 5 = 98.1%; service site 7 = 100%; service site 9 = 100%), and 100% by trainer 4 (service site 11). On two separate
occasions during the feedback sessions, trainer 3 miscalculated participants’ average correct performance, thereby incorrectly concluding that the participants had reached criterion. The trainer was alerted to this by the experimenters during session.

**Data Analysis**

A number of assumptions must be met, if a consumer is to be confident in the validity of the outcomes from a randomised control trial. Given that statistical power is compromised when random allocation is based on an individual rather than a cluster basis (Donner & Klar, 2004; Rutterford et al., 2015; Wears, 2002), a priori estimations of sample size are necessary (Wears, 2002). In addition, there is a possibility that responses from individuals within one “cluster” will be more similar to one another than to responses from individuals from a second “cluster” (Rutterford et al., 2015), thereby violating the assumption of response independence. Given that many parametric tests (e.g., ANOVA, ANCOVA) depend upon response independence (Field, 2016), analyses must be conducted to determine the appropriate method of statistical analysis.

Research standards for randomised control trials also stipulate that participant demographics remain comparable across experimental groups, as well as within experimental groups, over the course of the study (Amico, 2009). As such, retention rates should reach at least 70-80% in each experimental arm (e.g., Lyles et al., 2007; Whitlock, Polen, Green, Orleans, & Klein, 2004) and the evaluation of attrition patterns should include clarification of the “specific demographics and characteristics of those retained” (Amico, 2009, p. 1570). It is also recommended that any incomplete data should be missing completely at random (MCAR) (Kristman, Manno, & Cote, 2004).
In order to test the assumptions outlined above, relevant analyses were conducted to examine each of the following areas: (1) sample size calculation; (2) participation across study; (3) attrition across study; (4) missing data; (5) clustering of data, and (6) treatment of Likert scale data.

**Sample size calculation.** Statistical efficiency is adversely impacted with the use of a clustered randomised control trial (Donner & Klar, 2004; Rutterford et al., 2015; Wears, 2002). Traditional parametric approaches to experimental hypothesis testing (e.g., t-test, linear regression, ANOVA) make certain assumptions, including assuming the independence of scores within each treatment (Gravetter & Wallnau, 2004). However, the nature of clustered designs means that the responses from individuals within a cluster are more likely to be similar to each other than to responses from individuals from different clusters (Rutterford et al., 2015). This results in higher levels of variance (i.e., variability between participants in a cluster and variability between clusters), which must be accounted for by increasing the confidence intervals and associated p-values. In so doing, statistical power is compromised, thereby necessitating a larger sample size to negate this effect (Wears, 2002).

Given the resource investment involved in implementing a clustered randomised control trial in the host organisation, an a priori estimation of sample size was conducted. The University of California, San Francisco (UCSF) provides an online platform for ‘Calculating sample size when designing clinical research’ (Clinical and Transactional Science Institute (UCSF), 2017) and this platform includes a method for calculating sample size for clustered data. Devane, Begley, and Clarke (2004) strongly recommend that sample size should neither be too big nor too small considering the ethical implications related to resource wastage in clinical
research. Therefore, given that the pilot study revealed large effect sizes \(d > 1.0\) for knowledge acquisition (primary outcome), it was decided to base sample size for the current study on achieving a large effect size \(d \geq 0.08\). According to the online calculator a sample of 12 clusters (six clusters in each experimental arm), with at least five individuals in each cluster, was required to demonstrate a large effect size when the \(p\)-value was set at 0.05 (Clinical and Transactional Science Institute (UCSF), 2017). As such, a minimum of 60 participants were required to achieve adequate statistical power.

**Participation across the study.** A total of 104 participants were recruited for the clustered randomised control trial; 54 participants in the intervention condition and 50 participants in the control condition. However, there was evidence of attrition across both experimental conditions, over the course of the study. In addition, five participants from the control condition, who had not participated at Time 1 completed the battery of measures at Time 3. The patterns of participation and attrition observed during the study are discussed below.

**Participant characteristics – Time 1 (Baseline testing).** A total of 99 participants across 12 service sites participated in the administration of measures at Time 1. A total of 54 participants were assigned to the intervention condition and 45 participants were assigned to the control condition. There was no significant relationship between experimental allocation and gender \( (\chi^2 (1, n = 99) = 0.24, p = .627)\), level of education (i.e., pre-undergraduate degree & other, undergraduate degree, postgraduate qualification) \( (\chi^2 (2, n = 97) = 0.46, p = .793)\) or length of service within the host organisation \( (\chi^2 (3, n = 99) = 4.89, p = .180)\). Table 52 presents a summary of participant characteristics at Time 1.
Table 52

Summary of Participant Characteristics: Upper Panel - Gender, Education and Job Title for the Intervention Condition (N=54); Lower Panel – Gender, Education and Job Title for the Control condition (N=45)

<table>
<thead>
<tr>
<th>Gender</th>
<th>Percentage of Participants</th>
<th>Education</th>
<th>Number of Participants</th>
<th>Job Title</th>
<th>Number of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>24.1%</td>
<td>Leaving cert.</td>
<td>5.6%</td>
<td>Careworker</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td>75.9%</td>
<td>Certificate</td>
<td>7.4%</td>
<td>Programme facilitator</td>
<td>18.5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diploma</td>
<td>24.1%</td>
<td>Programme supervisor</td>
<td>11.1%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Undergrad.</td>
<td>42.6%</td>
<td>Team leader</td>
<td>7.4%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>degree</td>
<td></td>
<td>Service manager</td>
<td>3.7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Postgraduate</td>
<td>16.7%</td>
<td>Other</td>
<td>9.3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>qualification</td>
<td>3.7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td>Other</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
<th>Percentage of Participants</th>
<th>Education</th>
<th>Number of Participants</th>
<th>Job Title</th>
<th>Number of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>20%</td>
<td>Leaving cert.</td>
<td>2.2%</td>
<td>Careworker</td>
<td>55.6%</td>
</tr>
<tr>
<td></td>
<td>80%</td>
<td>Certificate</td>
<td>11.1%</td>
<td>Programme facilitator</td>
<td>22.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diploma</td>
<td>15.6%</td>
<td>Programme supervisor</td>
<td>4.4%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Undergrad.</td>
<td>42.2%</td>
<td>Team leader</td>
<td>8.9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>degree</td>
<td></td>
<td>Service manager</td>
<td>2.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Postgraduate</td>
<td>20%</td>
<td>Other</td>
<td>6.7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>degree</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Other</td>
<td>4.4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Missing</td>
<td>4.4%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Participant characteristics – Time 2 (Post-intervention testing). Only participants in the intervention condition participated in testing at Time 2. All 54 of these participants completed the reinforcement and systematic prompting knowledge assessments. Fifty-three of the participants completed the functional communication training knowledge assessment and 50 completed the task analysis knowledge assessment, Test of Knowledge questionnaire and Attitudes to Evidence-Based Practice Questionnaire.
One of the 54 participants missed the functional communication training and task analysis modules due to illness and was exempt from completing the associated assessments at Time 2. Similarly, two participants fell ill prior to the final post-intervention testing session and were excused from participation. Finally, one participant was present for the final post-intervention testing session but the relevant documents were not submitted to the research team. All four participants were employed in different service sites.

**Participant characteristics – Time 3.** A total of 92 participants across 12 service sites participated in testing at Time 3. A total of 46 participants were assigned to the intervention condition and 46 participants were assigned to the control condition. Table 5 characterises participants according to gender, education and job title for the intervention and control conditions, respectively.

There was no significant relationship between experimental allocation and gender ($\chi^2 (1, n = 92) = 0.28, p = .599$), level of education (i.e., pre-undergraduate degree & other; undergraduate degree; postgraduate qualification) ($\chi^2 (2, n = 91) = 0.37, p = .832$) or length of service within the host organisation ($\chi^2 (3, n = 92) = 6.11, p = .106$).
Table 53

Summary of Participant Characteristics: Upper Panel - Gender, Education and Job Title for the Intervention Condition (N=46); Lower Panel – Gender, Education and Job Title for the Control condition (N=46)

<table>
<thead>
<tr>
<th>Gender</th>
<th>Percentage of Participants</th>
<th>Education</th>
<th>Number of Participants</th>
<th>Job Title</th>
<th>Number of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>21.7%</td>
<td>Leaving cert. Certificate</td>
<td>6.5%</td>
<td>Careworker</td>
<td>50%</td>
</tr>
<tr>
<td>Female</td>
<td>78.3%</td>
<td>Certificate</td>
<td>6.5%</td>
<td>Programme facilitator</td>
<td>21.7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diploma</td>
<td>23.9%</td>
<td>Programme supervisor</td>
<td>13%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Undergrad. degree</td>
<td>45.7%</td>
<td>Team leader</td>
<td>8.7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Postgraduate qualification</td>
<td>15.2%</td>
<td>Service manager</td>
<td>2.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Other</td>
<td>2.2%</td>
<td>Other</td>
<td>4.3%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
<th>Percentage of Participants</th>
<th>Education</th>
<th>Number of Participants</th>
<th>Job Title</th>
<th>Number of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>17.4%</td>
<td>Leaving cert. Certificate</td>
<td>2.2%</td>
<td>Careworker</td>
<td>54.3%</td>
</tr>
<tr>
<td>Female</td>
<td>82.6%</td>
<td>Certificate</td>
<td>10.9%</td>
<td>Programme facilitator</td>
<td>26.1%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diploma</td>
<td>21.7%</td>
<td>Programme supervisor</td>
<td>4.3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Undergrad. degree</td>
<td>41.3%</td>
<td>Team leader</td>
<td>6.5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Postgraduate degree</td>
<td>19.6%</td>
<td>Service manager</td>
<td>2.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Other</td>
<td>2.2%</td>
<td>Other</td>
<td>6.5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Missing</td>
<td>2.2%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Attrition across the study. Analysis of participation rates show that eight participants (14.8%) from the intervention condition did not complete the study, as a result of absences due to illness and annual leave. The rate of attrition for the control condition was slightly lower at four participants (8.9%); two participants had left on maternity leave and two participants had changed to a different employment. This represents a differential attrition rate of 5.9% between the experimental arms. However, five individuals in the control condition, who had been absent during baseline testing (Time 1), completed the battery of questionnaires at Time 3.
Therefore, despite attrition, participant numbers in the control condition increased over the period of the study. Figure 17 provides a diagram of participant flow during the randomised control trial.

Further analyses showed that participant samples in both conditions remained comparable, despite attrition. For the intervention group there was no significant relationship between time point and gender ($\chi^2 (2, n = 150) = 0.10, p = .953$), time point and level of education (i.e., pre-undergraduate degree & other, undergraduate degree, postgraduate qualification) ($\chi^2 (4, n = 150) = 0.10, p = .999$) or time point and length of service within the host organisation ($\chi^2 (6, n = 150) = 0.24, p = 1.000$). Similarly, for the control group there was no significant relationship between time point and gender ($\chi^2 (1, n = 91) = 0.10, p = .750$), time point and level of education (i.e., pre-undergraduate degree & other, undergraduate degree, postgraduate qualification) ($\chi^2 (2, n = 88) = 0.08, p = .961$) or time point and length of service within the host organisation ($\chi^2 (3, n = 91) = 0.12, p = .990$). The latter result should be interpreted with caution given that the expected count in one cell was slightly less than 5 (i.e., 4.95). However, Fisher’s Exact Probability Test (2-tailed), which is often used in cases like this (Pallant, 2016), supported this non-significant relationship ($p = 1.000$).
Figure 17. Flow diagram showing participant attrition during the randomised control trial.
Missing data. An analysis was undertaken to evaluate the extent of missingness in the current data set and to determine if the incomplete data were missing at random or not (Little, 1988). If the data were missing completely at random, statistical analyses of the data could be confidently carried out, without the risk of incurring bias (Gelman & Hill, 2006; Seltman, 2015).

Knowledge assessments for each evidence-based practice module and the Test of Knowledge. Targeted knowledge acquisition was measured using five instruments; a multiple-choice, content specific assessment designed for each of the four training modules and the Test of Knowledge (Denne et al., 2015), which assesses participants’ knowledge of behavioural principles and applications, more generally.

For the intervention condition, analyses revealed missing data across all knowledge measures (Table 54 – upper panel). Attrition was the primary reason for incompleteness. Little’s Missing Completely at Random (MCAR) test, which was carried out simultaneously across all knowledge assessments, indicated that the data were missing completely at random (i.e., no identifiable pattern exists in the missing data): $\chi^2 (38) = 34.79, p = .619$. In addition, Little’s Missing Completely at Random (MCAR) test, which was obtained separately for each knowledge assessment, also indicated that data were missing completely at random; reinforcement knowledge assessment: $\chi^2 (2) = 0.07, p = .965$, systematic prompting knowledge assessment: $\chi^2 (2) = 1.64, p = .441$, functional communication training knowledge assessment: $\chi^2 (3) = 0.65, p = .885$, task analysis knowledge assessment: $\chi^2 (5) = 0.65, p = .260$, and Test of Knowledge: $\chi^2 (5) = 3.525, p = .620$. 

249
Likewise, data from the control condition showed evidence of missingness across all knowledge measures (Table 54 – lower panel) and attrition was the main reason for incompleteness. Little’s Missing Completely at Random (MCAR) test, which was carried out simultaneously across all knowledge assessments, indicated that the data were missing completely at random; \( \chi^2 (10) = 2.60, p = .989 \). In addition, Little’s Missing Completely at Random (MCAR) test, which was obtained separately for each knowledge assessment, also indicated that data were missing completely at random: reinforcement knowledge assessment: \( \chi^2 (2) = 0.76, p = .884 \), systematic prompting knowledge assessment: \( \chi^2 (2) = 0.85, p = .655 \), functional communication training knowledge assessment \( \chi^2 (2) = 1.29, p = .524 \), task analysis knowledge assessment: \( \chi^2 (2) = 1.88, p = .391 \), and Test of Knowledge: \( \chi^2 (2) = 1.18, p = .555 \).
Table 54

Summary of Missing and Valid Cases across the Measures Assessing Knowledge Acquisition: Upper Panel - Intervention Condition; Lower Panel – Control Condition

<table>
<thead>
<tr>
<th>Variable</th>
<th>Number of Missing Cases</th>
<th>Percentage of Missing Cases</th>
<th>Number of Valid Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reinforcement Assessment (Time 1)</td>
<td>0</td>
<td>0</td>
<td>54</td>
</tr>
<tr>
<td>Reinforcement Assessment (Time 2)</td>
<td>0</td>
<td>0</td>
<td>54</td>
</tr>
<tr>
<td>Reinforcement Assessment (Time 3)</td>
<td>9</td>
<td>16.7%</td>
<td>46</td>
</tr>
<tr>
<td>Prompting Assessment (Time 1)</td>
<td>0</td>
<td>0</td>
<td>54</td>
</tr>
<tr>
<td>Prompting Assessment (Time 2)</td>
<td>0</td>
<td>0</td>
<td>54</td>
</tr>
<tr>
<td>Prompting Assessment (Time 3)</td>
<td>9</td>
<td>16.7%</td>
<td>46</td>
</tr>
<tr>
<td>FCT Assessment (Time 1)</td>
<td>0</td>
<td>0</td>
<td>54</td>
</tr>
<tr>
<td>FCT Assessment (Time 2)</td>
<td>1</td>
<td>1.9%</td>
<td>53</td>
</tr>
<tr>
<td>FCT Assessment (Time 3)</td>
<td>9</td>
<td>16.7%</td>
<td>46</td>
</tr>
<tr>
<td>Task Analysis Assessment (Time 1)</td>
<td>0</td>
<td>0</td>
<td>54</td>
</tr>
<tr>
<td>Task Analysis Assessment (Time 2)</td>
<td>4</td>
<td>7.4%</td>
<td>50</td>
</tr>
<tr>
<td>Task Analysis Assessment (Time 3)</td>
<td>9</td>
<td>16.7%</td>
<td>46</td>
</tr>
<tr>
<td>Test of Knowledge (Time 1)</td>
<td>0</td>
<td>0</td>
<td>54</td>
</tr>
<tr>
<td>Test of Knowledge (Time 2)</td>
<td>4</td>
<td>7.4%</td>
<td>50</td>
</tr>
<tr>
<td>Test of Knowledge (Time 3)</td>
<td>9</td>
<td>16.7%</td>
<td>46</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Number of Missing Cases</th>
<th>Percentage of Missing Cases</th>
<th>Number of Valid Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reinforcement Assessment (Time 1)</td>
<td>5</td>
<td>10%</td>
<td>45</td>
</tr>
<tr>
<td>Reinforcement Assessment (Time 3)</td>
<td>4</td>
<td>8%</td>
<td>46</td>
</tr>
<tr>
<td>Prompting Assessment (Time 1)</td>
<td>5</td>
<td>10%</td>
<td>45</td>
</tr>
<tr>
<td>Prompting Assessment (Time 3)</td>
<td>4</td>
<td>8%</td>
<td>46</td>
</tr>
<tr>
<td>FCT Assessment (Time 1)</td>
<td>5</td>
<td>10%</td>
<td>45</td>
</tr>
<tr>
<td>FCT Assessment (Time 3)</td>
<td>4</td>
<td>8%</td>
<td>46</td>
</tr>
<tr>
<td>Task Analysis Assessment (Time 1)</td>
<td>5</td>
<td>10%</td>
<td>45</td>
</tr>
<tr>
<td>Task Analysis Assessment (Time 3)</td>
<td>4</td>
<td>8%</td>
<td>46</td>
</tr>
<tr>
<td>Test of Knowledge (Time 1)</td>
<td>5</td>
<td>10%</td>
<td>45</td>
</tr>
<tr>
<td>Test of Knowledge (Time 3)</td>
<td>4</td>
<td>8%</td>
<td>46</td>
</tr>
</tbody>
</table>

**Attitudes to Evidence-Based Practice Questionnaire.** Little’s Missing Completely at Random (MCAR) test indicated that baseline (Time 1) data were indeed missing completely at random across all items on the Attitudes toward Evidence-Based Practice Questionnaire for the intervention group; $\chi^2 (73) = 73.52, p$
\( \chi^2 \) \((81) = 67.41, \ p = .860. \) Likewise, Little’s Missing Completely at Random (MCAR) test indicated that at Time 3, the data were again missing completely at random across all items on the Attitudes toward Evidence-Based Practice Questionnaire for the intervention group; \( \chi^2 \) (73) = 80.38, \( p = .259 \) and the control group; \( \chi^2 \) (85) = 96.38, \( p = .187 \).

Finally, Little’s Missing Completely at Random (MCAR) test indicated that immediately post-intervention (Time 2), the data were not missing completely at random across all items on the Attitudes toward Evidence-Based Practice Questionnaire for the intervention group; \( \chi^2 \) (143) = 178.20, \( p = .024 \). This result may be influenced by the fact that four participants did not respond to any items on the questionnaire, due to illness and failure to submit complete documents. Nonetheless, the post-intervention outcomes for this variable must be interpreted with caution.

**Demographic questionnaire.** All participants, except two individuals in the control condition fully completed the demographic questionnaire. These two participants responded to all items in the questionnaire, except the item relating to the highest level of education attained.

**The Shortened Ways of Coping Questionnaire.** Participants in both experimental conditions completed the Shortened Ways of Coping Questionnaire. Table 55 characterises the missing data on this measure for the intervention (upper panel) and control (lower panel) conditions. Analyses showed that attrition contributed to some of the missingness. Little’s Missing Completely at Random (MCAR) test indicated that data were missing completely at random for the
intervention condition; χ² (4) = 3.43, p = .488 and the control condition; χ² (4) = 4.33, p = .364.

Short Version of the Occupational Self-Efficacy Scale. Participants in both experimental groups completed this measure. Table 55 details the pattern of missing data on this measure for the intervention (upper panel) and control (lower panel) conditions. Analyses showed that attrition accounted for some of the missingness for the control group and all of the missingness for the intervention group. Little’s Missing Completely at Random (MCAR) test indicated that the data were missing completely at random for the intervention condition; χ² (1) = 0.03, p = .861 and the control condition; χ² (2) = 1.96, p = .375.

Maslach Burnout Inventory – Human Services Survey. Participants in both experimental groups completed the Maslach Burnout Inventory. Table 55 summarises the levels of missing data on this measure for the intervention (upper panel) and control (lower panel) groups. Analyses showed that attrition accounted for a portion of the missingness in the data for both groups. Little’s Missing Completely at Random (MCAR) test indicated that data were missing completely at random for the intervention group; χ² (12) = 4.21, p = .979 and the control group; χ² (15) = 9.25, p = .864.

Minnesota Satisfaction Questionnaire – Short Form. Participants in both experimental groups completed the Minnesota Satisfaction Questionnaire. Table 55 provides a summary of the missing data on this measure for the intervention (upper panel) and control (lower panel) groups. Attrition was the main contributor to missing data in this instance. Little’s Missing Completely at Random (MCAR) test
indicated that data were indeed missing completely at random for the intervention group; \( \chi^2 (1) = 0.03, p = .861 \) and the control group; \( \chi^2 (2) = 3.09, p = .213 \).

**The Perceptions of Supervisory Support Scale.** Participants in both experimental groups completed the Perceptions of Supervisory Support Scale. Table 55 characterises the missing data on this measure for the intervention (upper panel) and control (lower panel) condition. Twelve participants in each group did not complete any item in this questionnaire.

**Coping style, occupational self-efficacy, burnout, job satisfaction and supervisory support.** Little’s Missing Completely at Random (MCAR) test obtained for data across all the variables listed, indicated that data were indeed missing completely at random for the intervention group; \( \chi^2 (77) = 89.48, p = .157 \) and the control group; \( \chi^2 (107) = 119.57, p = .191 \).
Table 55

Summary of Missing and Valid Cases across the Predictor Variables: Upper Panel - Intervention Condition; Lower Panel - Control Condition

<table>
<thead>
<tr>
<th>Variable</th>
<th>Number of Missing Cases</th>
<th>Percentage of Missing Cases</th>
<th>Number of Valid Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shortened Ways of Coping Questionnaire</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wishful thinking</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>2</td>
<td>3.7%</td>
<td>52</td>
</tr>
<tr>
<td>Time 3</td>
<td>8</td>
<td>14.8%</td>
<td>46</td>
</tr>
<tr>
<td>Practical coping</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>2</td>
<td>3.7%</td>
<td>52</td>
</tr>
<tr>
<td>Time 3</td>
<td>8</td>
<td>14.8%</td>
<td>46</td>
</tr>
<tr>
<td>Occupational Self-Efficacy Scale</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>0</td>
<td>0</td>
<td>54</td>
</tr>
<tr>
<td>Time 3</td>
<td>8</td>
<td>10%</td>
<td>46</td>
</tr>
<tr>
<td>Maslach Burnout Inventory</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exhaustion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>0</td>
<td>0</td>
<td>54</td>
</tr>
<tr>
<td>Time 3</td>
<td>9</td>
<td>16.7%</td>
<td>45</td>
</tr>
<tr>
<td>Depersonalisation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>0</td>
<td>0</td>
<td>54</td>
</tr>
<tr>
<td>Time 3</td>
<td>9</td>
<td>16.7%</td>
<td>45</td>
</tr>
<tr>
<td>Achievement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>1</td>
<td>1.9%</td>
<td>53</td>
</tr>
<tr>
<td>Time 3</td>
<td>12</td>
<td>22.2%</td>
<td>42</td>
</tr>
<tr>
<td>Minnesota Satisfaction Questionnaire</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>0</td>
<td>0</td>
<td>54</td>
</tr>
<tr>
<td>Time 3</td>
<td>8</td>
<td>10%</td>
<td>46</td>
</tr>
<tr>
<td>The Perceptions of Supervisory Support Scale</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>22.2%</td>
<td>42</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Number of Missing Cases</th>
<th>Percentage of Missing Cases</th>
<th>Number of Valid Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shortened Ways of Coping Questionnaire</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wishful thinking</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>7</td>
<td>14%</td>
<td>43</td>
</tr>
<tr>
<td>Time 3</td>
<td>4</td>
<td>8%</td>
<td>46</td>
</tr>
<tr>
<td>Practical coping</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>7</td>
<td>14%</td>
<td>43</td>
</tr>
<tr>
<td>Time 3</td>
<td>4</td>
<td>8%</td>
<td>46</td>
</tr>
<tr>
<td>Occupational Self-Efficacy Scale</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1</td>
<td>7</td>
<td>14%</td>
<td>43</td>
</tr>
<tr>
<td>Time 3</td>
<td>4</td>
<td>8%</td>
<td>46</td>
</tr>
<tr>
<td>Maslach Burnout Inventory</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exhaustion</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Participants in the current study were sampled from 12 separate service sites. Consequently, there was a possibility that responses from individuals within one service site (cluster) would be more similar to one another than to responses from individuals in the second service site (Rutterford et al., 2015). If this were true, response independence would be compromised. To account for this issue, the general mixed-effects linear model was selected as the method of statistical analysis.

In comparison to the general linear model, the mixed model is extremely robust, accounting for correlated data (including that associated with repeated measures), missing data and the clustering of data within contextual variables (e.g., service sites; (Field, 2016; Larsen et al., 2010; Seltman, 2015). However, as the mixed model is an extension of the basic linear model, the assumptions of linearity and residual normality and constant variance continue to apply (Field, 2016). In addition, the random coefficients (i.e., intercepts, slopes) should approximate a normal distribution around the overall model (Field, 2016).

For the current study, preliminary analyses were carried out to evaluate which random effects should be applied. According to Field (2016), it is recommended that

<table>
<thead>
<tr>
<th></th>
<th>Time 1</th>
<th>Time 3</th>
<th>Depersonalisation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>7</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Achievement</td>
<td>8</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Minnesota Satisfaction Questionnaire</td>
<td>8</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>The Perceptions of Supervisory Support Scale</td>
<td>12</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

Clustered data
analyses of mixed effects start with a model, in which all parameters are fixed. The subsequent addition of any random effects (e.g., intercepts, slopes) can then be evaluated by comparing model fit. In order to assess model fit, a chi-square likelihood ratio test should be conducted, using maximum likelihood estimation and the deviance statistic, which is minus twice the log-likelihood (-2LL) (Field, 2016). It should be noted that time was treated as a categorical variable in this study. Therefore, given that there was only one observation per participant per repeat (i.e., at each time point), it was impossible to fit a random slope model for this data set (due to the absence of residuals). As such, only random intercept models were tested; random intercepts across participants, random intercepts across service sites, and random intercepts across participants and across service sites. The fixed parameters tested in all models that compared the intervention and control conditions, were as follows: time of measurement; experimental allocation; interaction between time, and experimental allocation.

Assessment of model fit.

Reinforcement knowledge assessment. The initial model, which included fixed parameters only, produced a -2LL of 1645.73. According to results from the chi-square likelihood ratio test, all subsequent random intercept models significantly improved on this value, with the random intercept across participants model producing the most parsimonious solution, \( \text{Var}(u_{ij}) = 137.89, \chi^2(1) = 15.37, p < .01 \).

Systematic prompting knowledge assessment. The initial model, which included fixed parameters only, produced a -2LL of 1612.64. According to results from the chi-square likelihood ratio test, the random intercepts across participants model and the random intercepts across participants and across service sites model
significantly improved on this value, with the random intercept across participants model producing the most parsimonious solution, \( \text{Var}(u_{0j}) = 136.44, \chi^2 (1) = 19.55, p < .01. \)

*Functional communication training knowledge assessment.* The initial model, which included fixed parameters only, produced a -2LL of 1743.54. According to results from the chi-square likelihood ratio test, the random intercepts across participants model and the random intercepts across participants and across service sites model significantly improved on this value, with the random intercept across participants model producing the most parsimonious solution, \( \text{Var}(u_{0j}) = 216.37, \chi^2 (1) = 11.98, p < .01. \)

*Task analysis knowledge assessment.* The initial model, which included fixed parameters only, produced a -2LL of 1772.96. According to results from the chi-square likelihood ratio test, the random intercepts across participants model was the only model to significantly improve on this value, \( \text{Var}(u_{0j}) = 368.23, \chi^2 (1) = 24.74, p < .01. \)

*Test of Knowledge.* The initial model, which included fixed parameters only, produced a -2LL of 1425.05. According to results from the chi-square likelihood ratio test, all subsequent random intercept models significantly improved on this value, with the random intercept across participants model producing the most parsimonious solution, \( \text{Var}(u_{0j}) = 46.67, \chi^2 (1) = 18.82, p < .01. \)

*Maslach Burnout Inventory – Exhuastion dimension.* The initial model, which included fixed parameters only, produced a -2LL of 1271.54. According to results from the chi-square likelihood ratio test, all subsequent random intercept models significantly improved on this value, with the random intercept across participants model producing the most parsimonious solution, \( \text{Var}(u_{0j}) = 216.37, \chi^2 (1) = 11.98, p < .01. \)
participants model producing the most parsimonious solution, \( \text{Var}(u_{0j}) = 32.13, \chi^2 (1) = 38.46, p < .01. \)

*Maslach Burnout Inventory – Depersonalisation dimension.* The initial model, which included fixed parameters only, produced a -2LL of 1023.22. According to results from the chi-square likelihood ratio test, all subsequent random intercept models significantly improved on this value, with the random intercept across participants model producing the most parsimonious solution, \( \text{Var}(u_{0j}) = 7.48, \chi^2 (1) = 30.77, p < .01. \)

*Maslach Burnout Inventory – Personal achievement dimension.* The initial model, which included fixed parameters only, produced a -2LL of 1224.33. According to results from the chi-square likelihood ratio test, the random intercepts across participants model and the random intercepts across participants and across service sites model significantly improved on this value, with the random intercept across participants model producing the most parsimonious solution, \( \text{Var}(u_{0j}) = 28.91, \chi^2 (1) = 26.66, p < .01. \)

*The Shortened Ways of Coping Questionnaire – Wishful coping dimension.* The initial model, which included fixed parameters only, produced a -2LL of 973.36. According to results from the chi-square likelihood ratio test, all subsequent random intercept models significantly improved on this value, with the random intercept across participants model producing the most parsimonious solution, \( \text{Var}(u_{0j}) = 7.27, \chi^2 (1) = 52.11, p < .01. \)

*The Shortened Ways of Coping Questionnaire – Practical coping dimension.* The initial model, which included fixed parameters only, produced a -2LL of 973.36. According to results from the chi-square likelihood ratio test, the random intercepts
across participants model and the random intercepts across participants and across service sites model significantly improved on this value, with the random intercept across participants model producing the most parsimonious solution, \( \text{Var}(u_{0j}) = 4.76, \chi^2(1) = 24.44, p < .01 \).

**Minnesota Satisfaction Questionnaire – Short Form.** The initial model, which included fixed parameters only, produced a \(-2\text{LL}\) of 1447.92. According to results from the chi-square likelihood ratio test, all subsequent random intercept models significantly improved on this value, with the random intercept across service sites model producing the most parsimonious solution, \( \text{Var}(u_{0j}) = 42.41, \chi^2(1) = 51.96, p < .01 \).

**The Short Version of the Occupational Self-Efficacy Scale.** The initial model, which included fixed parameters only, produced a \(-2\text{LL}\) of 1092.62. According to results from the chi-square likelihood ratio test, the random intercepts across participants model and the random intercepts across participants and across service sites model significantly improved on this value, with the random intercept across participants model producing the most parsimonious solution, \( \text{Var}(u_{0j}) = 8.75, \chi^2(1) = 17.96, p < .01 \).

**The Perceptions of Supervisory Support Scale.**

**Assessment of model fit.** A general linear mixed model analysis was run to examine the relationship between perception of supervisory support and participant scores on the Test of Knowledge and each of the four knowledge assessments, at Time 3. According to results from the chi-square likelihood ratio tests, including random intercepts across participants, service sites or participants and service sites did not significantly improve on the \(-2\text{LL}\) value for any models across either the
intervention condition or the control condition. Therefore, a fixed parameters only model was used to quantify the relationship between perception of supervisory support and the Test of Knowledge and each of the four knowledge assessments.

**Likert data.** Likert scale data are based on ranked categories, the intervals between scale items are not equal. Therefore, it is generally recommended that nonparametric statistical tests should be used to analyse this type of data (Jamieson, 2004). However, research has shown that the $F$-statistic can return accurate $p$-values on Likert items (Glass, Peckham, & Sanders, 1972). General linear mixed-effect models employ the $F$-test, as part of their computation and given their level of precision, it is recommended that they can be used with Likert scale data (Kizach, 2014; Norman, 2010). Therefore, as the general mixed-effects linear model is being applied in the current study, outcome measures that employ Likert items will be analysed using this statistical method.

**Results**

**Participant Training History**

Results showed that prior to the Behaviour Skills Training intervention, the majority of participants in the intervention condition had not received training in any of the four target evidence-based practices (Figure 18). A total of 41 (75.9%), 47 (87%), 46 (85.2%), and 43 (79.6%) participants reported not having received any training in reinforcement, systematic prompting, functional communication training, and task analysis, respectively. Furthermore, of those who did receive training, only 3 (23.1%), 3 (42.9%), 2 (25%), and 1 (1.9%) participant(s) stated that a Behaviour Skills Training approach was used when they received training in reinforcement,
systematic prompting, functional communication training and task analysis respectively.

![Bar chart showing percentage of participants in intervention (N=54) condition who had already received training in each of the four evidence-based practices.

**Figure 18.** The percentage of participants in the intervention (N=54) condition who had already received training in each of the four evidence-based practices.

At baseline, a sizeable portion of the control group also reported that they had not received prior training in reinforcement (n=22; 48.9%), systematic prompting (n=32; 71.1%), functional communication training (n=31; 68.9%) and task analysis (n=30; 66.7%) (Figure 19). In addition, of those who did receive training, only 9 (39.1%), 2 (15.4%), 3 (21.4%) and 1 (6.7%) participant(s) stated that a Behaviour Skills Training approach was used when they received training in reinforcement, systematic prompting, functional communication training and task analysis, respectively.
Figure 19. The percentage of participants (N=45) in the control condition who had already received training in each of the four evidence-based practices.

**Knowledge Acquisition**

**Reinforcement knowledge assessment.** Time of measurement $F(1, 96.08) = 20.93, p < .001$, experimental allocation $F(1, 104.15) = 30.42, p < .001$, and the interaction between time of measurement and experimental allocation $F(1, 96.08) = 66.52, p < .001$ all significantly predicted participant scores on the reinforcement knowledge assessment.

The interaction was initially analysed by applying separate fixed parameters models to the reinforcement knowledge assessment data from Time 1 and Time 3 (the inclusion of random effects was not supported in this instance). Pairwise comparisons, which were based on estimated marginal means and used a Bonferroni adjustment, showed no significant difference between the two experimental conditions’ reinforcement knowledge assessment scores at Time 1; $M = 0.63, SE =$
3.85; $t(97) = 0.16$, $p = .870$. However, there was a significant difference in reinforcement knowledge assessment scores at Time 3; $M = 35.39$, $SE = 4.00$; $t(88) = 8.84$, $p < .001$.

The interaction was further analysed by applying separate mixed effects models to data from the intervention group and the control group. The models used were identical to the main model, with the main effect and interaction term relating to experimental allocation removed. Pairwise comparisons, which were based on estimated marginal means and used a Bonferroni adjustment, showed that participant scores on the reinforcement knowledge assessment significantly increased from Time 1 to Time 3 ($M = 28.01$, $SE = 2.48$) for the intervention group, $t(52.2) = 11.29$, $p < .001$ and significantly decreased from Time 1 to Time 3 ($M = -7.87$, $SE = 3.69$) for the control group, $t(45.35) = -2.13$, $p = .039$.

Figure 20 graphically depicts these trends in the data. Finally, confidence in the results is enhanced by the fact that the assumptions of linearity, residual normality and constant variance were met. In addition, the random intercept coefficients tended to be normally distributed around relevant models.
Figure 20. Average percentage of items answered correctly on the reinforcement knowledge assessment (based on the estimated marginal mean) for participants in the intervention condition at Time 1 (N=54) and Time 3 (N=44) and participants in the control condition at Time 1 (N=45) and Time 3 (N=46).

**Systematic prompting knowledge assessment.** Time of measurement $F(1, 90.80) = 17.42, p < .001$, experimental allocation $F(1, 100.23) = 47.15, p < .001$, and the interaction between time of measurement and experimental allocation $F(1, 90.80) = 41.49, p < .001$ all significantly predicted participant scores on the systematic prompting knowledge assessment.

The interaction was initially analysed by applying separate fixed parameters models to the systematic prompting knowledge assessment data from Time 1 and Time 3 (the inclusion of random effects was not supported in this instance). Pairwise comparisons, which were based on estimated marginal means and used a Bonferroni adjustment, showed that the systematic prompting knowledge assessment scores at Time 1 were significantly higher for the intervention group compared to the control.
group; $M = 7.89, SE = 3.55; t(97) = 2.22, p = .028$. Knowledge assessment scores at Time 3 were also significantly higher for the intervention group; $M = 33.47, SE = 3.64; t(88) = 9.20, p < .001$.

The interaction was further analysed by applying separate mixed effects models to data from the intervention group and the control group. The models used were identical to the main model, with the main effect and interaction term relating to experimental allocation removed. Pairwise comparisons, which were based on estimated marginal means and used a Bonferroni adjustment, showed that participant scores on the systematic prompting knowledge assessment significantly increased from Time 1 to Time 3 ($M = 20.62, SE = 2.93$) for the intervention group, $t(46.21) = 7.04, p < .001$ but did not change significantly from Time 1 to Time 3 ($M = -4.22, SE = 2.47$) for the control group, $t(43.82) = -1.71, p = .095$.

Figure 21 graphically depicts these trends in the data. Finally, confidence in the results is enhanced by the fact that the assumptions of linearity, residual normality and constant variance were met. In addition, the random intercept coefficients tended to be normally distributed around relevant models.
Figure 21. Average percentage of items answered correctly on the systematic prompting knowledge assessment (based on the estimated marginal mean) for participants in the intervention condition at Time 1 (N=54) and Time 3 (N=44) and participants in the control condition at Time 1 (N=45) and Time 3 (N=46).

**Functional communication training knowledge assessment.** Time of measurement $F(1, 90.49) = 19.9, p < .001$, experimental allocation $F(1, 99.54) = 37.438, p < .001$, and the interaction between time of measurement and experimental allocation $F(1, 90.49) = 74.50, p < .001$ all significantly predicted participant scores on the functional communication training knowledge assessment.

The interaction was initially analysed by applying separate fixed parameters models to the functional communication training knowledge assessment data from Time 1 and Time 3 (the inclusion of random effects was not supported in this instance). Pairwise comparisons, which were based on estimated marginal means and used a Bonferroni adjustment, showed no significant difference between the two experimental conditions’ functional communication training knowledge assessment.
scores at Time 1; \( M = 1.37, SE = 5.54; t(97) = 0.25, \ p = .805 \). However, there was a significant difference in functional communication training knowledge assessment scores at Time 3; \( M = 49.23, SE = 4.45; t(88) = 11.07, \ p < .001 \).

The interaction was further analysed by applying separate mixed effects models to data from the intervention group and the control group. The models used were identical to the main model, with the main effect and interaction term relating to experimental allocation removed. Pairwise comparisons, which were based on estimated marginal means and used a Bonferroni adjustment, showed that participant scores on the functional communication training knowledge assessment significantly increased from Time 1 to Time 3 (\( M = 35.93, SE = 3.91 \)) for the intervention group, \( t(47.68) = 9.19, \ p < .001 \) and significantly decreased from Time 1 to Time 3 (\( M = -11.47, SE = 4.39 \)) for the control group, \( t(45.56) = -2.61, \ p = .012 \).

Figure 22 graphically depicts these trends in the data. Finally, confidence in the results is enhanced by the fact that the assumptions of linearity, residual normality and constant variance were met. In addition, the distribution of random intercept coefficients around relevant models did not deviate substantially from a normal distribution.
**Figure 22.** Average percentage of items answered correctly on the functional communication training knowledge assessment (based on the estimated marginal mean) for participants in the intervention condition at Time 1 (N=54) and Time 3 (N=44) and participants in the control condition at Time 1 (N=45) and Time 3 (N=46).

**Task analysis knowledge assessment.** Time of measurement $F(1, 87.12) = 7.86, p = .006$, experimental allocation $F(1, 97.84) = 24.32, p < .001$, and the interaction between time of measurement and experimental allocation $F(1, 87.12) = 33.91, p < .001$ all significantly predicted participant scores on the task analysis knowledge assessment.

The interaction was initially analysed by applying separate fixed parameters models to the task analysis knowledge assessment data from Time 1 and Time 3 (the inclusion of random effects was not supported in this instance). Pairwise comparisons, which were based on estimated marginal means and used a Bonferroni adjustment, showed no significant difference between the two experimental
conditions’ task analysis knowledge assessment scores at Time 1; $M = 6.00, SE = 5.90; t(97) = 1.02, \ p = .312$. However, there was a significant difference in task analysis knowledge assessment scores at Time 3; $M = 40.93, SE = 4.94; t(88) = 8.28, \ p < .001$.

The interaction was further analysed by applying separate mixed effects models to data from the intervention group and the control group. The models used were identical to the main model, with the main effect and interaction term relating to experimental allocation removed. Pairwise comparisons, which were based on estimated marginal means and used a Bonferroni adjustment, showed that participant scores on the task analysis knowledge assessment significantly increased from Time 1 to Time 3 ($M = 24.46, SE = 3.98$) for the intervention group, $t(40.54) = 6.15, \ p < .001$ and significantly decreased from Time 1 to Time 3 ($M = -8.11, SE = 3.97$) for the control group, $t(43.40) = -2.04, \ p = .047$.

Figure 23 graphically depicts these trends in the data. Finally, confidence in the results is enhanced by the fact that the assumptions of linearity, residual normality and constant variance were met. In addition, the distribution of random intercept coefficients around relevant models did not deviate substantially from normality.
Figure 23. Average percentage of items answered correctly on the task analysis knowledge assessment (based on the estimated marginal mean) for participants in the intervention condition Time 1 (N=54) and Time 3 (N=44) and participants in the control condition at Time 1 (N=45) and Time 3 (N=46).

**Test of Knowledge.** Time of measurement $F(1, 96.02) = 22.35, p < .001$, experimental allocation $F(1, 104.79) = 37.18, p < .001$, and the interaction between time of measurement and experimental allocation $F(1, 96.02) = 38.31, p < .001$ all significantly predicted participant scores on the Test of Knowledge.

The interaction was initially analysed by applying separate fixed parameters models to the Test of Knowledge data from Time 1 and Time 3 (the inclusion of random effects was not supported in this instance). Pairwise comparisons, which were based on estimated marginal means and used a Bonferroni adjustment, showed no significant difference between the two experimental conditions’ Test of Knowledge scores at Time 1; $M = 3.33, SE = 1.94; t(97) = 1.72, p = .089$. However,
there was a significant difference in Test of Knowledge scores at Time 3; $M = 18.33$, $SE = 2.45$; $t(88) = 7.49$, $p < .001$.

The interaction was further analysed by applying separate mixed effects models to data from the intervention group and the control group. The models used were identical to the main model, with the main effect and interaction term relating to experimental allocation removed. Pairwise comparisons, which were based on estimated marginal means and used a Bonferroni adjustment, showed that participant scores on the Test of Knowledge significantly increased from Time 1 to Time 3 ($M = 12.99$, $SE = 1.75$) for the intervention group, $t(52.21) = 7.42$, $p < .001$ but did not change significantly from Time 1 to Time 3 ($M = -1.61$, $SE = 1.58$) for the control group, $t(44.50) = -1.02$, $p = .314$.

Figure 24 graphically depicts these trends in the data. Finally, confidence in the results is enhanced by the fact that the assumptions of linearity and residual constant variance were met. There was a slight deviation from normality within the distribution of residuals. However, the distribution still approximated normality and this assumption is generally a concern only with small sample sizes (i.e., < 30). Therefore, a slight deviation from normality should not pose a problem in this instance (Field, 2016; Gravetter & Wallnau, 2004; Seltman, 2015). In addition, the distribution of random intercept coefficients around relevant models did not deviate substantially from a normal distribution.
Figure 24. Average percentage of items answered correctly on the Test of Knowledge (based on the estimated marginal mean) for participants in the intervention condition at Time 1 (N=54) and Time 3 (N=44) and participants in the control condition at Time 1 (N=45) and Time 3 (N=46).

Psychological Well-being

Maslach Burnout Inventory – Human Services Survey.

Exhaustion dimension. Time of measurement $F(1, 88.08) = 0.52, p = .473$, experimental allocation $F(1, 100.24) = 0.11, p = .746$, and the interaction between time of measurement and experimental allocation $F(1, 88.08) = 0.45, p = .505$ did not predict participant scores on this dimension. Confidence in the results is enhanced by the fact that the assumptions of linearity, residual normality and constant variance were met. In addition, the distribution of random intercept coefficients around the overall model did not deviate significantly from normality.
**Depersonalisation dimension.** Time of measurement $F(1, 93.24) = 1.98, p = .163$, experimental allocation $F(1, 103.37) = 0.13, p = .718$, and the interaction between time of measurement and experimental allocation $F(1, 93.24) = 2.78, p = .099$ did not predict participant scores on this dimension. Confidence in the results is enhanced by the fact that the assumptions of linearity, residual normality and constant variance were met. In addition, the distribution of random intercept coefficients around the overall model did not deviate significantly from normality.

**Personal achievement dimension.** Time of measurement $F(1, 84.05) = 0.21, p = .646$, experimental allocation $F(1, 96.96) = 1.03, p = .313$, and the interaction between time of measurement and experimental allocation $F(1, 84.05) = 0.50, p = .482$ did not predict participant scores on this dimension. Confidence in the results is enhanced by the fact that the assumptions of linearity, residual normality and constant variance were met. In addition, the distribution of random intercept coefficients around the overall model did not deviate significantly from normality.

**The Shortened Ways of Coping Questionnaire.**

**Wishful thinking dimension.** Time of measurement $F(1, 87.57) = 0.17, p = .677$, experimental allocation $F(1, 100.74) = 0.07, p = .786$, and the interaction between time of measurement and experimental allocation $F(1, 87.57) = 0.68, p = .413$ did not predict participant scores on this dimension. Confidence in the results is enhanced by the fact that the assumptions of linearity, residual normality and constant variance were met. In addition, the distribution of random intercept coefficients around the overall model did not deviate significantly from normality.

**Practical coping dimension.** Time of measurement $F(1, 88.59) = 1.49, p = .226$, experimental allocation $F(1, 98.58) = 2.65, p = .107$, and the interaction
between time of measurement and experimental allocation $F(1, 88.59) = 0.27, p = .604$ did not predict participant scores on this dimension. Confidence in the results is enhanced by the fact that the assumptions of linearity, residual normality and constant variance were met. In addition, the distribution of random intercept coefficients around the overall model did not deviate from normality.

**Minnesota Satisfaction Questionnaire.** Time of measurement $F(1, 176.16) = 1.61, p = .206$, experimental allocation $F(1, 9.95) = 1.57, p = .239$, and the interaction between time of measurement and experimental allocation $F(1, 176.16) = 0.03, p = .874$ did not predict participant scores on this questionnaire. Confidence in the results is enhanced by the fact that the assumptions of linearity, residual normality and constant variance were met. In addition, the distribution of random intercept coefficients around the overall model did not deviate significantly from normality.

**The Short Version of the Occupational Self-Efficacy Scale.** Time of measurement $F(1, 88.07) = 2.53, p = .115$, experimental allocation $F(1, 97.17) = 0.30, p = .586$, and the interaction between time of measurement and experimental allocation $F(1, 88.07) = 2.74, p = .101$ did not predict participant scores on this measure. Confidence in the results is enhanced by the fact that the assumptions of linearity, residual normality and constant variance were met. In addition, the distribution of random intercept coefficients around the overall model did not deviate from normality.

**The Perceptions of Supervisory Support Scale**

For the intervention group, results from the general mixed-effects linear model analysis revealed no significant relationship between perception of supervisory support and participants’ scores on the reinforcement knowledge
assessments ($F(1, 39) = 0.14, p = .712$), the systematic prompting knowledge assessment ($F(1, 39) = 0.01, p = .905$), the functional communication training knowledge assessment ($F(1, 39) = 3.97, p = .053$), the task analysis knowledge assessment ($F(1, 39) = 0.52, p = .475$), and the Test of Knowledge ($F(1, 39) = 1.27, p = .267$) at Time 3. Confidence in these results is enhanced by the fact that the assumptions of linearity and residual normality and constant variance were met.

For the control group, results from the general mixed-effects linear model analysis also revealed no significant relationship between perception of supervisory support and participants’ scores on the reinforcement knowledge assessment ($F(1, 36) = 0.13, p = .720$), the systematic prompting knowledge assessment ($F(1, 36) = 3.24, p = .080$), the functional communication training knowledge assessment ($F(1, 36) = 0.37, p = .548$), the task analysis knowledge assessment ($F(1, 36) = 0.02, p = .895$), and the Test of Knowledge ($F(1, 36) = 1.02, p = .320$) at Time 3. Confidence in these results is enhanced by the fact that the assumptions of linearity, residual normality and constant variance were met.

**Attitudes to Evidence-Based Practice Questionnaire**

Tables 56 and 57 (Appendix H) summarise the percentage of participants selecting a specific response on each item of the Attitudes to Evidence-Based Practice Questionnaire. Direct comparisons of the percentage data were not made, either across time or across group, because patterns of missing data varied across questionnaire item, group and time point. Instead, a more general comparison of response trends (based on percentage data) was made for each questionnaire item, across time and group. Finally, given that the intervention group’s post-intervention (Time 2) data was not missing completely at random, the post-intervention results for
this group were interpreted with caution and validated by reference to the results from Time 3.

**Intervention condition.** Results show that participants generally displayed positive attitudes towards evidence-based practice throughout the study. The majority of participants “agreed” or “strongly agreed” that implementing evidence-based practice would benefit them professionally and that there were advantages to changing their approach according to recommendations from high quality research. Although, over 40% of the participants were unsure about their ability to evaluate available research at Time 1 (baseline testing), responses at Time 2 (post-intervention testing) and at Time 3 (follow-up testing) showed an increase in confidence across participants. Likewise, despite the majority of participants agreeing that they experienced difficulty when implementing change in the workplace at Time 1, the majority “disagreed” with this statement at Time 2 and Time 3. However, notwithstanding this increase in confidence, the majority of participants continued to report that they would feel more assured if a “professional skilled in research” was available to provide them with relevant information.

**Control condition.** Similar to the intervention group, the majority of the control group also “agreed” or “strongly agreed” that implementing evidence-based practices would benefit them professionally and that there were advantages to changing their approach according to recommendations from high quality research. However, at Time 3, a larger proportion of participants in the control condition “agreed” or “strongly agreed” that research findings do not easily transfer into their own work. Furthermore, the control group consistently reported time and cost restraints as barriers to the implementation of evidence-based practice in their workplace. The intervention group, on the other hand, were not so decisive on these
issues at either Time 1 or Time 3. Finally, despite a sizeable portion of participants in the control condition repeatedly indicating that they trust their own abilities to evaluate research, the majority continued to report that they would feel more confident if a “professional skilled in research” was available to provide them with relevant information.

Discussion

The goal of the current study was to implement a randomised control trial to determine whether Behaviour Skills Training was an effective method for disseminating evidence-based practice to frontline staff working with adults with intellectual disabilities. The first objective was to test whether the Behaviour Skills Training intervention would positively impact participants’ knowledge of the target evidence-based practices. It was hypothesised that when compared to baseline testing (Time 1), participants in the Behaviour Skills Training condition would score significantly higher on knowledge assessments at Time 3. However, this pattern was not predicted for participants in the wait list control condition.

Consistent with predictions, participants exposed to Behaviour Skills Training scored significantly higher on all four knowledge assessments and the Test of Knowledge at the 4-week follow-up. In contrast, participants in the wait list control condition showed either no change or a statistically significant decrease in their knowledge scores over the study period. Furthermore, there was no significant difference between the two experimental groups’ baseline scores on three of the four knowledge assessments or on the Test of Knowledge. As a result, the improvement in knowledge scores among participants in the intervention condition, can be more confidently attributed to the impact of the Behaviour Skills Training package.
However, given that the time constraints governing the current research programme resulted in the control group being designated to receive the Behaviour Skills Training intervention outside the time-frame allocated for the current study, it is impossible to determine whether exposure to the intervention would have resulted in similar knowledge gains among participants in the wait list control condition.

The second objective investigated potential barriers to successful dissemination of the four target evidence-based practices across both experimental groups. Previous research reported that the impact of supervision can outweigh specific training methods (Rapp et al., 2008). However, results from the current study showed no relationship between perceptions of supervisory support and knowledge scores at Time 3, for participants in either condition. An analysis of participant responding to the Attitudes to Evidence-Based Practice Questionnaire also revealed no significant impediment to the dissemination of evidence-based practice among participants in either experimental group. However, a larger portion of the wait list control group reported cost and time restraints as barriers to the implementation of evidence-based practice in their workplace.

The final objective focused on evaluating the impact of the Behaviour Skills Training intervention on participants’ psychological well-being. It was hypothesised that exposure to this training programme would produce an improvement in staff-reported measures of job satisfaction, burnout, occupational self-efficacy, and the use of effective coping strategies. However, it was predicted that a similar change would not be observed among participants in the wait list control condition. Contrary to predictions, the results consistently showed no statistically significant change in measures of psychological well-being over the course of the study, for participants in either of the two experimental conditions.
A key limitation of the current study was the inability to definitively conclude that the Behaviour Skills Training intervention had no impact on participants’ psychological well-being. Firstly, considering the complexity of the psychological constructs being measured (Fukui et al., 2014; Lazarus, 1999; Locke, 1969; Maslach, 1993; Maslach, 1999), the follow-up period of 4 weeks may have been inadequate to facilitate meaningful change. Hätinen, Mäkikangas, Kinnunen, and Pekkonen (2013) conducted a longitudinal study to evaluate the impact of a 1-year intervention, specifically designed to target recovery from burnout and at the 6-month follow-up, burnout recovery was only consistently evident for exhaustion. Such results demonstrate the complexity of psychological variables. Therefore, future studies must facilitate follow-up over a substantially longer period than 4 weeks, in order to adequately evaluate the impact of Behaviour Skills Training on relevant psychological constructs.

Furthermore, the sample size employed in the current study may have been insufficient for detecting changes to participants’ psychological well-being. A priori estimations of sample size were based on the large effect sizes observed during the pilot study reported in Chapter 8 and these effect sizes were exclusively related to participant performance on the Test of Knowledge and the four knowledge assessments. Furthermore, the pilot study had failed to yield adequate data to calculate reliable effect sizes pertaining to the psychological well-being variables. Nevertheless, given that the primary focus of the randomised control trial was the dissemination of evidence-based practice and there is an ethical obligation to avoid resource wastage in applied research (Devane et al., 2004), the approach taken for sample size calculation was considered prudent.
A second limitation to this study was identified with the analysis of baseline knowledge assessment scores. Findings showed that participants in the Behaviour Skills Training condition scored significantly higher than participants in the control condition on the systematic prompting assessment. Therefore, at follow-up testing, improved scores on the systematic prompting assessment for participants in the intervention group cannot be readily accredited to Behaviour Skills Training, given that at baseline, these participants were already outperforming their counterparts in the control condition. In addition, the edumetric reliability of the systematic prompting assessment was compromised. These issues were not a concern for the remaining three knowledge assessments and the Test of Knowledge.

Notwithstanding these limitations, results from the randomised control trial clearly support Behaviour Skills Training as an effective method for disseminating knowledge relating to evidence-based practices, to frontline staff working with adults with intellectual disabilities; scores on target knowledge assessments improved significantly from baseline (Time 1) to follow-up (Time 3). However, it is recognised that knowledge acquisition alone will not lead to successful change in procedural practices in the workplace (Embry & Biglan, 2008; Grey et al., 2007); there must also be an emphasis on skill acquisition (Parsons et al., 2012). As a result, Chapter 10 will provide a detailed analysis of the extent of knowledge maintenance over time, as well as the degree of skill acquisition and generalisation that was demonstrated by participants in the intervention condition. An indepth analysis of these outcomes can be valuable in informing staff training within clinical practice. Given the logistical constraints controlling the current research programme, an evaluation of knowledge maintenance and skill acquisition among participants
assigned to the wait list control condition was not possible as part of the randomised control trial reported in this chapter.
Chapter 10

Dissemination of Evidence-Based Practice to Frontline Staff Working in the Field of Intellectual Disability: An Indepth Analysis of the Behaviour Skills Training Intervention
Abstract

Frontline staff frequently report not receiving training specialised to their work with people with intellectual disabilities (Campbell, 2010). Furthermore, the training literature states that many instructional programmes fail to adequately train staff employed in the human services sector; these programmes fail to bridge the gap between classroom-based learning of knowledge and skills and daily practices of staff in their typical work environments (Casey & McWilliam, 2011; Nitasha, Lisa, & Craig, 2004; Parsons et al., 2012; Sturmey, 1998). As a result, the extent of knowledge maintenance, skill acquisition, and skill generalisation was analysed for the 54 participants assigned to the intervention condition during the randomised control trial, reported in Chapter 9. However, given logistical constraints, it was not possible to assess these outcomes for participants assigned to the wait list control condition. Participant performance across all four knowledge assessments and the Test of Knowledge improved significantly between baseline (Time 1) and post-intervention testing (Time 2) and performance was maintained at the 4-week follow-up (Time 3) across the reinforcement, functional communication training, and task analysis knowledge assessments, as well as the Test of Knowledge. However, participant performance on the systematic prompting knowledge assessment decreased significantly at follow-up testing (Time 3). Although knowledge maintenance was not achieved, data indicate that participant knowledge on this topic did improve as a result of Behaviour Skills Training. Results also provide clear evidence of skill acquisition and subsequent generalisation to the workplace environment. Finally, participants evaluated the behaviour skill training package favourably.
Frontline staff are an extremely valuable resource within the intellectual disability sector; the standard of their work dictates the overall quality of care within an organisation (Campbell, 2010; Devereux, Hastings, & Noone, 2009; Salyers et al., 2015). Research has shown that compromised skill sets among staff can adversely affect staff-service user relations (Finn & Sturmey, 2009), the learning opportunities available (Schepis et al., 2001), and quality of life outcomes for service users (Jahr, 1998). With the successful implementation of interventions largely dependent on the behaviour of frontline staff, their level of knowledge and skills represent an important factor in effective service provision (Lowe et al., 2007).

However, a sizeable portion of the frontline staff, who are supporting people with intellectual disabilities may not be considered adequately qualified (Campbell, 2010; Joint Committee on Social Care Professionals, 2002). Furthermore, even among those staff with relevant professional qualifications, the skills required to effectively support this population are typically underdeveloped or absent at the time of qualification (Campbell, 2010). As a result, high quality, effective staff training has been identified as a central component in superior service provision in the intellectual disability sector (Lowe et al., 2007).

In spite of this, frontline staff frequently report not receiving training specialised to their work with people with intellectual disabilities (Campbell, 2010). This is reflected in results from the training needs analysis described in Chapter 6, which show that the majority of staff reported receiving no preparatory training from the host organisation, in either behaviour management strategies or skills teaching procedures, prior to beginning direct work with service users. In addition, more than half of the participants who had received preparatory training did not believe that it adequately equipped them to support service users. Results from the study also
revealed that less than 40% of staff received ongoing training in how to teach their service users new skills, despite 93.5% of staff being employed with the host organisation for at least a year.

Therefore, although the literature has suggested that the successful implementation of evidence-based practice often fails due to staff resistance (Ager & O'May, 2001; Grey et al., 2007), deficient skill-sets among frontline staff may also contribute to this problem. The evidence base for effective behavioural interventions is well-established (Grey & Hastings, 2005; Kurtz et al., 2011; National Autism Center, 2009; 2015; Sturmey & Didden, 2014; Wong et al., 2014), but in comparison the evidence base for effective staff training is relatively weak (Campbell, 2010).

Results from the systematic literature review of training practices reported in Chapter 5 show that many established evidence-based practices were not addressed. For example, while functional communication training is an empirically supported treatment for individuals with intellectual disabilities (Kurtz et al., 2011), no studies in the review focused on providing training in the implementation of this skill. In addition, only eight of the 107 studies reviewed, contrasted the effects of different training packages on staff and service user outcomes.

Furthermore, the training literature reports that many instructional programmes fail to adequately train staff employed in the human services sector in target skills and procedures (Casey & McWilliam, 2011; Nitasha et al., 2004; Parsons et al., 2012; Sturmey, 1998). According to Grey et al. (2007), focusing explicitly on knowledge acquisition as a training goal, is ineffective in improving procedural integrity. In contrast, training programmes such as Behaviour Skills Training, which emphasise the teaching and measurement of overt staff behaviour, have demonstrated superior outcomes (e.g., Ducharme, Williams, Cummings,
Murray, & Spencer, 2001; Lambert et al., 2013; Lavie & Sturmey, 2002; Nabeyama & Sturmey, 2010; Schepis et al., 2001). In 2014, Fetherston and Sturmey provided a systematic replication of previous studies which had demonstrated the efficacy of Behaviour Skills Training across the implementation of discrete trial teaching (Dib & Sturmey, 2007; Lafasakis & Sturmey, 2007), incidental teaching (Ryan et al., 2008), and teaching activity schedules (Ryan et al., 2008; Sarokoff & Sturmey, 2004).

There was substantial evidence of skill acquisition and positive behaviour change for staff trainees and their clients, across all three experiments. In addition, staff generalised new practices to novel situations and rated the training procedure as highly acceptable (Fetherston & Sturmey, 2014).

In order to demonstrate clinical value and social validity, a training programme must effectively transition between classroom-based learning of knowledge and skills and daily practices of staff in their typical work environment (van Oorsouw et al., 2009). Results from the randomised control trial reported in Chapter 9, clearly support Behaviour Skills Training as an effective method for disseminating knowledge relating to evidence-based practice, to frontline staff working with adults with intellectual disabilities. However, knowledge acquisition alone has been shown to be insufficient in producing change (Grey et al., 2007).

Therefore, the current study sought to provide a closer examination of the impact of the Behaviour Skills Training intervention, tested in Chapter 9. Of particular interest was the extent of knowledge maintenance, as well as the degree of skill acquisition and generalisation that was demonstrated by participants in the Behaviour Skills Training condition.

According to Campbell (2010), relatively few studies, which focus on effective training interventions, incorporate an evaluation of the long-term training
outcomes for frontline staff in the experimental design. Of those studies that do, results frequently show a deterioration of knowledge and skill performance at follow-up assessment (Campbell, 2010). Therefore, the first objective in the current study was to examine the extent of knowledge maintenance at 4 weeks post-intervention (i.e., follow-up testing). Consistent with the Behaviour Skills Training literature, it was hypothesised that participants would maintain the knowledge that they had acquired, during Behaviour Skills Training, at follow-up testing. In addition, the impact of factors, including educational level and length of service, on knowledge acquisition among staff was investigated.

The second objective focused on the performance- and competency-based components of Behaviour Skills Training. Previous research has not examined the efficacy of Behaviour Skills Training in coaching staff in the specific combination of evidence-based practices reported in Chapter 9. Therefore, the current study focused on investigating the speed of target skill mastery, as well as levels of procedural integrity on completion of each training module.

The training literature also reports that successful post-intervention generalisation of trained skills to the workplace is rare; it remains “the exception rather than the rule in reported studies” (Campbell, 2010, p.190). Again, these findings are contrary to those reported in the Behaviour Skills Teaching literature, which show clear evidence of skill generalisation across people and settings (e.g., Fetherston & Sturmey, 2014; Sarokoff & Sturmey, 2008). Therefore, the third objective of the current study was to examine the generalisation of functional communication training skills from the classroom setting, to the typical workplace environment, with service users. Based on results from the pilot study reported in
Chapter 8, it was hypothesised that high levels of procedural generalisation would be observed.

Finally, research has shown that “just because a program is considered effective does not mean that it will be considered appropriate by those closely involved in the implementation of the program” (Carter, 2010, p. 7). Instead, the “user friendliness” of a programme or intervention appears integral to effective implementation (McConnachie & Carr, 1997). Therefore, the fourth and final objective of this chapter was to evaluate participants’ perception of the Behaviour Skills Training package, in terms of its relevancy, usefulness, organisation, and potential for interaction and engagement.

**Phase I – Behaviour Skills Training**

**Method**

**Service sites.** Six of the 12 service sites, recruited for the randomised control trial reported in Chapter 9, were assigned to the “Behaviour Skills Training” intervention condition. Four of these service sites provided residential support to service users, while two offered support through day-based services.

**Participants.** In total, 54 participants were assigned to the intervention condition. Tables 50 and 51, which are reported in Chapter 9, characterise participants according to service site, length of service with the host organisation, gender, education, and job title. A total of 41 participants were female and 37 participants listed their job title as careworker (n=27) or programme facilitator (n=10). Twenty-three participants had achieved an undergraduate degree and nine had earned a postgraduate qualification. Overall, 30 participants had been employed with the host organisation for 5 years or less, 14 participants reported an employment
period of 5-10 years, and 10 participants reported an employment period of 10+ years.

**Trainers.** Trainer characteristics are presented in Chapter 9.

**Experimenters.** Experimenter characteristics are presented in Chapter 9.

**Setting.**

*Trainer instruction.* Trainer instruction was identical to that described in Chapter 9.

*Administration of measures.* The administration of measures was identical to that described in Chapter 9.

*Behaviour Skills Training.* Behaviour Skills Training was identical to that described in Chapter 9.

**Materials.**

*Trainer instruction.*

*Trainer pack.* The trainer pack was identical to that described in Chapter 9.

*Behaviour Skills Training.*

*Evidence-based practice training packs.* The evidence-based practice training packs were identical to those described in Chapter 9.

*PowerPoint presentations.* The PowerPoint presentations were identical to that described in Chapter 9.

*Rehearsal props.* The rehearsal props were identical to that described in Chapter 9.
Chapter 10

Measures. The battery of measures was identical to that used during the randomised control trial (Chapter 9) and outlined in detail in Chapter 8: (1) demographic questionnaire; (2) knowledge assessments for each evidence-based practice module; (3) mastery assessments for each evidence-based practice module; (4) Training Acceptability Survey for each evidence-based practice module; (5) Attitudes to Evidence-Based Practice Questionnaire (McKenna et al., 2004); (6) Maslach Burnout Inventory – Human Services Survey (MBI-HSS; Maslach et al., 1996); (7) Minnesota Satisfaction Questionnaire – Short Form (MSQ-SF; Weiss et al., 1967); (8) Short Version of the Occupational Self-Efficacy Scale (Rigotti et al., 2008); (9) Test of Knowledge (Denne et al., 2015); (10) The Perceptions of Supervisory Support Scale (PSS; Fukui et al., 2014), and (11) The Shortened Ways of Coping Questionnaire (SWC-R; Hatton & Emerson, 1995).

Knowledge assessments for each evidence-based practice module (available from the author on request). The analysis of edumetric validity and edumetric reliability was reported in Chapter 9.

Training Acceptability Survey for each evidence-based practice module (Underwood et al., 2002). The training acceptability survey is a six-item respondent-based measure that examines training, in relation to its relevance, planning, opportunities for participation, resource use, areas of difficulty, providing motivation to learn more, and overall usefulness. Responses are measured on a four-point Likert Scale, with “not true” scoring a 1 and “completely true” scoring a 4. In the current study the Cronbach alpha coefficient was 0.88 for the reinforcement module, 0.87 for the systematic prompting module, 0.91 for the functional communication training module, and 0.90 for the task analysis module.
**Design.** A within-subjects, pre-post-test design was used to evaluate the impact of an evidence-based Behaviour Skills Training programme on participant knowledge of each target evidence-based practice. Observations to criterion, a type of event recording, was also employed during training to measure the number of practice opportunities required by participants to reach a pre-determined level of proficiency/mastery in the implementation of skills related to each target evidence-based practice (Cooper et al., 2014).

**Procedure.**

_Randomisation of service sites._ The randomisation process was identical to that outlined in Chapter 9.

_Trainer instruction._ The instruction programme provided was identical to that used in the randomised control trial (Chapter 9) and outlined in the pilot study reported in Chapter 8.

_Administration of measures – Time 1 (Baseline testing)._ The administration of measures was identical to the procedures used in the randomised control trial (Chapter 9) and outlined in the pilot study reported in Chapter 8. Figure 25 provides a timeline of study implementation for the intervention group.

_Intervention (Behaviour Skills Training)._ The Behaviour Skills Training intervention was identical to that used in Chapter 9.

_Administration of measures - Time 2 (Post-intervention testing)._ The administration of measures was identical to the procedures used in the randomised control trial (Chapter 9) and outlined in the pilot study reported in Chapter 8.
Chapter 10

Administration of measures – Time 3 (Follow-up testing). The administration of measures was identical to the procedures used in the randomised control trial (Chapter 9) and outlined in the pilot study reported in Chapter 8.

![Flow diagram showing the timeline of study implementation.](image)

Procedural fidelity. The procedural fidelity data are presented in Chapter 9.

Interobserver agreement. Interobserver agreement (IOA) data were collected for 79.2% of reinforcement, 75.5% of systematic prompting, 79.9% of functional communication training to access items, 83% of functional communication training to access a break, and 90.2% of task analysis rehearsal.
sessions. The trainer and the lead experimenter independently recorded the accuracy of participants’ delivery of each of the target practices during role-play scenarios. A point-by-point method was used to calculate IOA. The number of agreements was divided by the number of agreements and disagreements. This figure was multiplied by 100 to establish the percentage IOA. The mean percentage of agreement ranged from 98.7% to 99.3%; reinforcement module (98.7%), systematic prompting module (98.9%), functional communication training module (access to items – 98.7%; access to breaks – 99.3%), and task analysis module (98.7%).

A second interrater agreement analysis was conducted using the Kappa statistic (Cohen, 1960). According to Cohen (1960), the statistic can be interpreted as follows: < 0 indicates “no agreement”, 0.01-0.20 indicates “none to slight” agreement, 0.21-0.40 indicates “fair” agreement, 0.41-0.60 indicates “moderate” agreement, 0.61-0.80 indicates “substantial” agreement, and 0.81-1.00 indicates “almost perfect” agreement. In the current study, overall agreement among raters was characterised as “almost perfect”; reinforcement module (Kappa = 0.93, p<.001); systematic prompting module (Kappa = 0.85, p<.001); functional communication training module (access to items: Kappa = 0.89, p<.001; access to break: Kappa = 0.88, p<.001), and the task analysis module (Kappa = 0.80, p<.001).

Table 58 provides the interobserver agreement and interrater reliability (Kappa) data across service sites for all rehearsal sessions.
Table 58

Interobserver agreement (IOA) and interrater reliability (Kappa) across Service Sites for all Rehearsal Sessions

<table>
<thead>
<tr>
<th>Service Site</th>
<th>Trainer</th>
<th>Rehearsal Session</th>
<th>Observations with</th>
<th>IOA Score</th>
<th>Kappa Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>Positive reinforcement</td>
<td>42.9%</td>
<td>100%</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Least prompting</td>
<td>42.9%</td>
<td>96.4%</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FCT (access items)</td>
<td>61.9%</td>
<td>97.5%</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FCT (access breaks)</td>
<td>57.1%</td>
<td>99.7%</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Backward chaining</td>
<td>100%</td>
<td>98.7%</td>
<td>0.54</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>Positive reinforcement</td>
<td>100%</td>
<td>99.3%</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Least prompting</td>
<td>100%</td>
<td>99.5%</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FCT (access items)</td>
<td>100%</td>
<td>99.9%</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FCT (access breaks)</td>
<td>100%</td>
<td>99.6%</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Backward chaining</td>
<td>100%</td>
<td>94.6%</td>
<td>0.38</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>Positive reinforcement</td>
<td>100%</td>
<td>99.3%</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Least prompting</td>
<td>100%</td>
<td>99.2%</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FCT (access items)</td>
<td>100%</td>
<td>99.3%</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FCT (access breaks)</td>
<td>100%</td>
<td>100%</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Backward chaining</td>
<td>100%</td>
<td>99.6%</td>
<td>0.95</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>Positive reinforcement</td>
<td>100%</td>
<td>97.3%</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Least prompting</td>
<td>100%</td>
<td>98.8%</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FCT (access items)</td>
<td>100%</td>
<td>98.3%</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FCT (access breaks)</td>
<td>100%</td>
<td>98.6%</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Backward chaining</td>
<td>100%</td>
<td>99.2%</td>
<td>0.86</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>Positive reinforcement</td>
<td>100%</td>
<td>98.8%</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Least prompting</td>
<td>100%</td>
<td>98.6%</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FCT (access items)</td>
<td>100%</td>
<td>98.8%</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FCT (access breaks)</td>
<td>100%</td>
<td>99.1%</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Backward chaining</td>
<td>100%</td>
<td>98.9%</td>
<td>0.92</td>
</tr>
<tr>
<td>11</td>
<td>4</td>
<td>Positive reinforcement</td>
<td>46.2%</td>
<td>98.6%</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Least prompting</td>
<td>30.8%</td>
<td>100%</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FCT (access items)</td>
<td>38.5%</td>
<td>99.2%</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FCT (access breaks)</td>
<td>53.8%</td>
<td>99.1%</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Backward chaining</td>
<td>61.5%</td>
<td>100%</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Data analysis. As described in Chapter 9, a number of assumptions must be met, if a consumer is to be confident in the validity of the statistical outcomes from a research study. Therefore, in order to test the relevant assumptions, analyses were conducted across the following areas: (1) Sample size calculation; (2) Participation across study; (3) Attrition across study; (4) Missing data, and (5) Clustering of data.

Sample size calculation. The calculation of sample size is reported in Chapter 9.

Patterns of participation and attrition. The relevant findings are reported in Chapter 9.

Missing data. An analysis was undertaken to evaluate the extent of missingness in the current data set and to determine if the incomplete data were missing at random or not (Little, 1988). If the data were missing completely at random, statistical analyses of the data could be confidently carried out, without the risk of incurring bias (Gelman & Hill, 2006; Seltman, 2015).

Knowledge assessments for each evidence-based practice module and the Test of Knowledge. The relevant findings are reported in Chapter 9.

Skill assessment. Procedural integrity was assessed across five target skills. Table 59 shows the extent of missing data during rehearsal and feedback sessions for each evidence-based practice module. Analyses indicated low levels of missing data and the reason for missingness was participant absence due to illness. The affected participants were employed by different service sites and attended different training sessions.
Table 59

Summary of Missing and Valid Cases across the Rehearsal and Feedback Sessions for the Intervention Group

<table>
<thead>
<tr>
<th>Module and Target Skill</th>
<th>Number of Missing Cases</th>
<th>Percentage of Missing Cases</th>
<th>Number of Valid Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reinforcement:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive reinforcement</td>
<td>0</td>
<td>0</td>
<td>54</td>
</tr>
<tr>
<td>Prompting:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Least-most prompting</td>
<td>0</td>
<td>0</td>
<td>54</td>
</tr>
<tr>
<td>Functional Communication Training:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FCT to access items/activities</td>
<td>1</td>
<td>1.9%</td>
<td>53</td>
</tr>
<tr>
<td>FCT to access a break</td>
<td>1</td>
<td>1.9%</td>
<td>53</td>
</tr>
<tr>
<td>Task Analysis:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task analysis with backward chaining</td>
<td>3</td>
<td>5.6%</td>
<td>51</td>
</tr>
</tbody>
</table>

Demographic questionnaire. The findings are reported in Chapter 9.

Training Acceptability Survey. Table 60 provides a summary of the data missing across the training acceptability survey for each training module. There were no responses missing for any item on surveys relating to the reinforcement and systematic prompting modules. Three participants did not respond to any item on the training acceptability survey for the functional communication training module.

Similarly, four participants did not provide any responses on the training acceptability survey for the task analysis module. The primary reason for failure to complete the surveys was absence due to illness. Little’s Missing Completely at Random (MCAR) test indicated that the data was indeed missing completely at random across the four training acceptability survey (i.e., for each of the four evidence-based practice training modules); \( \chi^2 (8) = 14.31, p = .074 \).
Table 60

*Summary of Missing and Valid Cases on the Training Acceptability Survey for each Evidence-Based Practice Module*

<table>
<thead>
<tr>
<th>Module</th>
<th>Number of Missing Cases</th>
<th>Percentage of Missing Cases</th>
<th>Number of Valid Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reinforcement</td>
<td>0</td>
<td>0</td>
<td>54</td>
</tr>
<tr>
<td>Prompting</td>
<td>0</td>
<td>0</td>
<td>54</td>
</tr>
<tr>
<td>Functional Communication Training</td>
<td>3</td>
<td>5.6%</td>
<td>51</td>
</tr>
<tr>
<td>Task Analysis</td>
<td>4</td>
<td>7.4%</td>
<td>50</td>
</tr>
</tbody>
</table>

*Clustering and assessment of model fit.* In order to account for correlated data (including that associated with repeated measures), missing data and the clustering of data within contextual variables (e.g., service sites; (Field, 2016; Larsen et al., 2010; Seltman, 2015), the general mixed-effects linear model was selected as the method of statistical analysis in this instance. Preliminary analyses were carried out to evaluate which random effects should be applied. It should be noted that time was treated as a categorical variable in this study, making it impossible to fit a random slope model for this data set (due to the absence of residuals). As such, only random intercept models were tested; random intercepts across participants, random intercepts across service sites and random intercepts across participants and across service sites. The fixed parameter tested in all models was time of measurement.

*Reinforcement knowledge assessment.* The initial model, which included fixed parameters only, produced a -2LL of 1250.23. According to results from the chi-square likelihood ratio test, the random intercept across participants model was the only model to significantly improve on this value, $\text{Var}(u_{0j}) = 56.16, \chi^2 (1) = 9.02, p < .01$.

*Systematic prompting knowledge assessment.* The initial model, which included fixed parameters only, produced a -2LL of 1249.03. According to results
from the chi-square likelihood ratio test, the random intercept across participants model was the only model to significantly improve on this value, $\text{Var}(u_{0j}) = 47.44, \chi^2(1) = 6.31, p < .05$.

*Functional communication training knowledge assessment.* The initial model, which included fixed parameters only, produced a -2LL of 1314.44. According to results from the chi-square likelihood ratio test, the random intercepts across participants model and the random intercepts across participants and across service sites model significantly improved on this value, with the random intercept across participants model producing the most parsimonious solution, $\text{Var}(u_{0j}) = 90.29, \chi^2(1) = 7.81, p < .01$.

*Task analysis knowledge assessment.* The initial model, which included fixed parameters only, produced a -2LL of 1303.51. According to results from the chi-square likelihood ratio test, the random intercepts across participants model was the only model that significantly improved on this value, $\text{Var}(u_{0j}) = 84.07, \chi^2(1) = 4.81, p < .05$.

*Test of Knowledge.* The initial model, which included fixed parameters only, produced a -2LL of 1061.92. According to results from the chi-square likelihood ratio test, the random intercepts across participants model was the only model that significantly improved on this value, $\text{Var}(u_{0j}) = 19.4, \chi^2(1) = 8.35, p < .01$.

**Results**

**Knowledge acquisition.**

*Reinforcement knowledge assessment.* Time of measurement $F(2, 100.60) = 94.27, p < .001$ significantly predicted participant scores on the reinforcement knowledge assessment. Pairwise comparisons, which were based on estimated
marginal means and used a Bonferroni adjustment, showed that participant scores on
the reinforcement knowledge assessment significantly increased from Time 1 to
Time 2, $M = 31.30, SE = 2.47, p < .001$ and from Time 1 to Time 3, $M = 28.07, SE =
2.64, p < .001$. However, there was no significant change in participants’ scores
from Time 2 to Time 3, $M = -3.23, SE = 2.64, p = .672$.

Figure 26 graphically depicts these trends in the data. Confidence in the
results is enhanced by the fact that the assumptions of linearity and residual
normality and constant variance were met. In addition, the random intercept
coefficients tended to be normally distributed around the overall model.

Figure 26. Average percentage of items answered correctly on the reinforcement
knowledge assessment (based on the estimated marginal mean) at Time 1 (N=54),
Time 2 (N=54), and Time 3 (N=44).

Systematic prompting knowledge assessment. Time of measurement $F(2,
99.91) = 67.35, p < .001$ significantly predicted participant scores on the systematic
prompting knowledge assessment. Pairwise comparisons, which were based on
estimated marginal means and used a Bonferroni adjustment, showed that participant scores on the systematic prompting knowledge assessment significantly increased from Time 1 to Time 2, $M = 28.52$, $SE = 2.53$, $p < .001$, and from Time 1 to Time 3, $M = 20.83$, $SE = 2.69$, $p < .001$. There was also a significant decrease in participants’ scores from Time 2 to Time 3, $M = -7.69$, $SE = 2.69$, $p = .016$.

Figure 27 graphically depicts these trends in the data. Confidence in the results is enhanced by the fact that the assumptions of linearity and residual normality and constant variance were met. In addition, the random intercept coefficients tended to be normally distributed around the overall model.

![Figure 27](image)

Figure 27. Average percentage of items answered correctly on the systematic prompting knowledge assessment (based on the estimated marginal mean) at Time 1 (N=54), Time 2 (N=54), and Time 3 (N=44).
Functional communication training knowledge assessment. Time of measurement \( F(2, 97.12) = 97.61, p < .001 \) significantly predicted participant scores on the functional communication training knowledge assessment. Pairwise comparisons, which were based on estimated marginal means and used a Bonferroni adjustment, showed that participant scores on the functional communication training knowledge assessment significantly increased from Time 1 to Time 2, \( M = 41.22, SE = 3.18, p < .001 \), and from Time 1 to Time 3, \( M = 35.95, SE = 3.37, p < .001 \). There was no significant change in participants’ scores from Time 2 to Time 3, \( M = -5.27, SE = 3.38, p = .366 \).

Figure 28 graphically depicts these trends in the data. Confidence in the results is enhanced by the fact that the assumptions of linearity and residual normality and constant variance were met. In addition, the random intercept coefficients tended to be normally distributed around the overall model.
Figure 28. Average percentage of items answered correctly on the functional communication training knowledge assessment (based on the estimated marginal mean) at Time 1 (N=54), Time 2 (N=53), and Time 3 (N=44).

Task analysis knowledge assessment. Time of measurement $F(2, 91.91) = 40.97, p < .001$ significantly predicted participant scores on the task analysis knowledge assessment. Pairwise comparisons, which were based on estimated marginal means and used a Bonferroni adjustment, showed that participant scores on the task analysis knowledge assessment significantly increased from Time 1 to Time 2, $M = 29.54, SE = 3.50, p < .001$, and from Time 1 to Time 3, $M = 28.84, SE = 3.65, p < .001$. There was no significant change in participants’ scores from Time 2 to Time 3, $M = -4.70, SE = 3.71, p = .626$.

Figure 29 graphically depicts these trends in the data. Confidence in the results is enhanced by the fact that the assumptions of linearity and residual normality and constant variance were met. In addition, the distribution of random
intercept coefficients around the overall model did not deviate grievously from a normal distribution.

Figure 29. Average percentage of items answered correctly on the task analysis knowledge assessment (based on the estimated marginal mean) at Time 1 (N=54), Time 2 (N=50), and Time 3 (N=44).

**Test of Knowledge.** Time of measurement $F(2, 97.94) = 51.72, p < .001$ significantly predicted participant scores on the Test of Knowledge. Pairwise comparisons, which were based on estimated marginal means and used a Bonferroni adjustment, showed that participant scores on the Test of Knowledge significantly increased from Time 1 to Time 2, $M = 13.63, SE = 1.50, p < .001$, and from Time 1 to Time 3, $M = 12.95, SE = 1.57, p < .001$. There was no significant change in participants’ scores from Time 2 to Time 3, $M = -0.68, SE = 1.59, p = 1.00$.

Figure 30 graphically depicts these trends in the data. Confidence in the results is enhanced by the fact that the assumptions of linearity and residual
normality and constant variance were met. In addition, the random intercept coefficients tended to be normally distributed around the overall model.

**Figure 30.** Average percentage of items answered correctly on the Test of Knowledge (based on the estimated marginal mean) at Time 1 (N=54), Time 2 (N=50), and Time 3 (N=44).

**Relationship between knowledge acquisition and education level.** A general linear mixed model analysis was run to examine the impact of education level on knowledge acquisition. A random intercepts across participants model was used, with time of measurement, education level and the interaction between these two factors as the fixed parameters.

Analyses showed that time of measurement significantly predicted participant scores on the reinforcement ($F(2, 96.80) = 82.63, p < .001$), systematic prompting ($F(2, 96.62) = 52.54, p < .001$), functional communication training ($F(2, 94.26) = 75.72, p < .001$) and task analysis knowledge assessments ($F(2, 92.41) = 32.30, p < .001$), as well as the Test of Knowledge ($F(2, 94.40) = 46.28, p < .001$).
However, level of education did not significantly predict scores on the Test of Knowledge or three of the four knowledge assessments: Test of Knowledge \(F(2, 51.41) = 2.45, p = .096\); reinforcement knowledge assessment \(F(2, 52.49) = 1.04, p = .361\); systematic prompting knowledge assessment \(F(2, 51.76) = 2.21, p = .120\), and the functional communication training knowledge assessment \(F(2, 50.10) = 2.18, p = .124\). In addition, there was no evidence of an interaction between time of measurement and level of education across the Test of Knowledge or any of the four knowledge assessments: Test of Knowledge \(F(4, 94.07) = 0.89, p = .476\); reinforcement knowledge assessment \(F(4, 96.70) = 0.54, p = .705\); systematic prompting knowledge assessment \(F(4, 96.51) = 1.02, p = .399\); functional communication training knowledge assessment \(F(4, 94.19) = 1.08, p = .373\), and the task analysis knowledge assessment \(F(4, 92.00) = 1.12, p = .350\).

Level of education did significantly predict scores on the task analysis knowledge assessment \(F(2, 48.20) = 5.79, p = .006\). According to pairwise comparisons, with a Bonferroni adjustment, participants who had not yet achieved an undergraduate degree consistently scored lower than those with an undergraduate degree \((M = -12.09, p = .010)\) and those with a postgraduate degree \((M = -13.00, p = .049)\). However, as stated above, there was no interaction between time of measurement and education level.

Confidence in all results obtained was enhanced by the fact that the assumptions of linearity and residual normality and constant variance were met. In addition, the random intercept coefficients tended to be normally distributed around their respective models.
Relationship between knowledge acquisition and length of service. A general linear mixed model analysis was run to examine the impact of length of service on knowledge acquisition. A random intercepts across participants model was used, with time of measurement, length of service and the interaction between these two factors as the fixed parameters.

Analyses showed that time of measurement significantly predicted participant scores on the reinforcement ($F(2, 95.01) = 84.25, p < .001$), systematic prompting ($F(2, 93.90) = 62.93, p < .001$), functional communication training ($F(2, 91.17) = 89.09, p < .001$) and task analysis knowledge assessments ($F(2, 86.93) = 37.45, p < .001$), as well as the Test of Knowledge ($F(2, 92.77) = 45.49, p < .001$).

However, length of service did not significantly predict scores on the Test of Knowledge or any of the four knowledge assessments: Test of Knowledge ($F(3, 51.13) = 0.45, p = .716$); reinforcement knowledge assessment ($F(3, 53.51) = 2.07, p = .116$); systematic prompting knowledge assessment ($F(3, 50.37) = 0.10, p = .960$); functional communication training knowledge assessment ($F(3, 48.10) = 0.66, p = .583$), and the task analysis knowledge assessment ($F(3, 44.82) = 0.25, p = .859$). In addition, there was no evidence of an interaction between time of measurement and length of service across the four knowledge assessments or the Test of Knowledge: reinforcement knowledge assessment ($F(6, 94.90) = 0.66, p = .680$); systematic prompting knowledge assessment ($F(6, 93.78) = 0.54, p = .773$); functional communication training knowledge assessment ($F(6, 91.00) = 0.79, p = .580$); task analysis knowledge assessment ($F(6, 86.67) = 0.78, p = .590$), and the Test of Knowledge ($F(6, 92.53) = 0.46, p = .837$).
Confidence in all results obtained was enhanced by the fact that the assumptions of linearity and residual normality and constant variance were met. In addition, the distribution of random intercept coefficients around their respective overall model did not deviate significantly from a normal distribution.

**Relationship between knowledge acquisition and prior training in the target evidence-based practices.** A general linear mixed model analysis was run to examine the impact of prior training in each of the evidence-based practices on knowledge acquisition. A random intercepts across participants model was used, with time of measurement, prior training and the interaction between these two factors as the fixed parameters.

Analyses showed that time of measurement significantly predicted participant scores on the reinforcement ($F(2, 99.16) = 64.88, p < .001$), systematic prompting ($F(2, 99.37) = 28.54, p < .001$), functional communication training ($F(2, 98.88) = 36.84, p < .001$) and task analysis knowledge assessments ($F(2, 91.07) = 29.47, p < .001$)

However, prior training in the target evidence-based practices did not predict scores on the relevant knowledge assessments: reinforcement knowledge assessment ($F(1, 54.25) = 0.04, p = .846$); systematic prompting knowledge assessment ($F(1, 54.65) = 0.49, p = .489$); functional communication training knowledge assessment ($F(1, 55.36) = 0.26, p = .610$), and the task analysis knowledge assessment ($F(1, 47.54) = 0.06, p = .809$). In addition, there was no evidence of an interaction between time of measurement and prior training across any of the knowledge assessments: reinforcement knowledge assessment ($F(2, 99.16) = 1.22, p = .301$); systematic prompting knowledge assessment ($F(2, 99.37) = 0.02, p = .978$);
functional communication training knowledge assessment \( F(2, 98.88) = 1.36, p = .262 \), and task analysis knowledge assessment \( F(2, 91.07) = 0.35, p = .705 \).

Confidence in all results obtained was enhanced by the fact that the assumptions of linearity and residual normality were met. In addition, the distribution of random intercept coefficients around their respective overall model did not deviate substantially from a normal distribution.

**Target skill acquisition.** Figure 31 represents the average number of observations taken by participants in the intervention group to reach the mastery criterion during the rehearsal and feedback sessions for each of the evidence-based practice modules. The average number of observations required for participants was 3.2 (range = 3-6 observations), 3.4 (range = 3-9 observations), 3.9 (range = 3-6 observations), 3.3 (range = 3-8 observations), and 3.6 (range = 3-7 observations) for the reinforcement, systematic prompting, functional communication training (access to items/activities & access to breaks), and task analysis modules, respectively.
Figure 31. Average number of observations taken by participants in the intervention group to reach mastery during the rehearsal and feedback sessions for each evidence-based practice module.

Table 61 displays the average percentage of steps completed correctly on the first observation and the final observation for each module. The data show that following the rehearsal and feedback sessions, participants were achieving average procedural integrity levels exceeding 97% across all target skills.
Table 61

**Mean Percentage of Steps Correctly Implemented on First and Final Observations Across Each Evidence-Based Practice Module**

<table>
<thead>
<tr>
<th>Module</th>
<th>Mean Percentage of Steps Correctly Implemented on First Observation</th>
<th>Mean Percentage of Steps Correctly Implemented on Final Observation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reinforcement</td>
<td>93.9% (80-100%)</td>
<td>97.8% (90-100%)</td>
</tr>
<tr>
<td>Prompting</td>
<td>93.9% (71.5-100%)</td>
<td>98.7% (90-100%)</td>
</tr>
<tr>
<td>Functional Communication Training Items/Activities</td>
<td>87.5% (63.5-100%)</td>
<td>99.5% (90-100%)</td>
</tr>
<tr>
<td>Break</td>
<td>94.3% (71.5-100%)</td>
<td>99.7% (91.5-100%)</td>
</tr>
<tr>
<td>Task Analysis</td>
<td>92.9% (71-100%)</td>
<td>99.7% (93-100%)</td>
</tr>
</tbody>
</table>

**Social validity of Behaviour Skills Training package.** Table 62 (Appendix I) provides a summary of participant responding to each item on the training acceptability survey across each evidence-based practice module. The data show that participants evaluated the behaviour skill training package positively, in terms of its relevancy, organisation and potential for interaction and engagement. Table 63 summarises the potential score range, actual score range and the mean and median scores for participants. The potential score range on the training acceptability survey was 8-32 and the median rating score was at least 29 across all modules. The functional communication training and task analysis modules recorded the highest mean and median scores.
Table 63

Potential Score Range, Actual Score Range and the Mean and Median Scores for Participants’ Evaluations of the Four Evidence-Based Practice Modules

<table>
<thead>
<tr>
<th>Method</th>
<th>Reinforcement (N=54)</th>
<th>Prompting (N=54)</th>
<th>Functional Communication Training (N=51)</th>
<th>Task Analysis (N=50)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potential score range</td>
<td>8-32</td>
<td>8-32</td>
<td>8-32</td>
<td>8-32</td>
</tr>
<tr>
<td>Actual score range</td>
<td>18-32</td>
<td>21-32</td>
<td>19-32</td>
<td>22-32</td>
</tr>
<tr>
<td>Mean score</td>
<td>27.6</td>
<td>28.1</td>
<td>28.7</td>
<td>28.8</td>
</tr>
<tr>
<td>Median score</td>
<td>28</td>
<td>28.5</td>
<td>30</td>
<td>30</td>
</tr>
</tbody>
</table>

Phase II – Skill Generalisation

Method

Participants. Participants were the same 54 frontline staff that participated in Phase I.

Trainers. Trainers were the same four trainers that delivered training during Phase I.

Experimenters. The lead and second experimenters from Phase I also acted as experimenters in Phase II.

Service users. Nineteen service users whose direct support staff had participated in the training sessions during Phase I participated in Phase II. To be considered a direct support staff (or “keyworker”) for a service user, staff must have known the service user and had the opportunity to observe them in one or more environments for 3 months or more. Specialised consent forms were designed to facilitate service user understanding that their keyworker(s) would be observed by a trainer and experimenter while they worked with them (available from the author on
request). Guardian consent was obtained for those service users unable to provide assent (consent forms available from the author on request). All service users were at least 18 years of age and had an intellectual disability.

**Setting.** Sessions were conducted in settings, where the participants typically supported their service users. These settings included residential care and day support services.

**Measures.**

*Mastery assessment for functional communication training in the natural environment.* The mastery assessment data sheets used during the functional communication training rehearsal session were adapted to record staff participant performance when they were supporting key service users in the natural work environment (available from the author on request).

**Design.** Observations to criterion was employed during the practical observations to measure the number of opportunities required by participants to reach a pre-determined level of proficiency/mastery in the implementation of the target functional communication skill with a service user in the typical work environment (Cooper et al., 2014).

**Procedure.** The procedure was identical to that used during Phase II of the pilot study reported in Chapter 8.

**Interobserver agreement.** During Phase II, interobserver agreement (IOA) was documented for 78.9% of sessions. The trainer and one of the experimenters independently recorded the accuracy of participants’ delivery of functional communication training to service users. A point-by-point method was used and the mean percentage of agreement was 99.8%. The Kappa statistic (Cohen, 1960) was
also calculated, in order to evaluate interrater reliability. Results indicated “almost
perfect” agreement among raters (Kappa = 0.99, p < .001) (Cohen, 1960). Table 64
provides the interobserver agreement and interrater reliability (Kappa) across service
sites.

Table 64

Interobserver Agreement (IOA) and interrater reliability (Kappa) across Service
Sites for all Observations

<table>
<thead>
<tr>
<th>Service Site</th>
<th>Trainer</th>
<th>FCT Skill</th>
<th>Observations with IOA</th>
<th>IOA Score</th>
<th>Kappa Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>Access to items</td>
<td>83.3%</td>
<td>98.7%</td>
<td>0.93</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>Access to items</td>
<td>33.3%</td>
<td>100%</td>
<td>1.00</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>Access to items</td>
<td>87.5%</td>
<td>100%</td>
<td>1.00</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>Access to items</td>
<td>75%</td>
<td>100%</td>
<td>1.00</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>Access to breaks</td>
<td>42.8%</td>
<td>100%</td>
<td>1.00</td>
</tr>
<tr>
<td>11</td>
<td>4</td>
<td>Access to breaks, Access to items</td>
<td>50%</td>
<td>100%</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Data Analysis.

**Missing data.** Fifteen participants (27.8%) had no opportunity to display the
target skill due to an absence of motivation on the part of their service site to request.
Nine participants (16.7%) were unavailable for observation due to illness, annual
leave or service site closure.

Results

Data show that 30 (55.6%) participants had the opportunity to implement
functional communication training with a service user in their typical work setting.
Participants who did not have the opportunity to engage in the in-situ observations (as a result of either limited service user motivation or their own unavailability) were scheduled for additional training after the current research programme was complete.

Twenty-nine of the 30 participants (96.7%) who had the opportunity to implement functional communication training achieved mastery within three observations. Of these 30 participants, 25 (83.3%) displayed mastery on the first observation. Three participants required two observations to achieve mastery and only one participant required the maximum of three observations to reach mastery. One participant did not achieve mastery within three observations. This participant had displayed 83.3% procedural integrity on the target skill for the first observation. However, the participant had no additional opportunities to practice the skill due to an absence of motivation on the part of the service user to request. Figure 32 shows the percentage of staff participants that reached mastery after the first, second and third observation.
Figure 32. Percentage of participants reaching mastery after the first, second and third observation during the implementation of functional communication training with a service user in the typical work environment.

Discussion

Results from the randomised control trial reported in Chapter 9, confirm that Behaviour Skills Training was an effective method for disseminating knowledge relating to evidence-based practice to frontline staff. However, successful change to procedural practices in the workplace is typically unsuccessful with knowledge acquisition alone. Training programmes that prioritise skill development, consistently demonstrate superior outcomes (Grey et al., 2007). It is important to note, however, that relatively few studies evaluate the maintenance of newly acquired knowledge or skills post-training (Campbell, 2010). Therefore, the goal of the current study was to more thoroughly evaluate the specific impact of the Behaviour Skills Training package, which was tested in Chapter 9. However, given
logistical constraints, it was not possible to conduct an equivalent analysis for participants assigned to the wait list control condition.

The first objective was to examine the extent of knowledge maintenance, which was displayed by the intervention group, over the course of the study. Participant performance across all four knowledge assessments and the Test of Knowledge improved significantly between baseline (Time 1) and post-intervention testing (Time 2) and performance on the Test of Knowledge and the reinforcement, functional communication training, and task analysis knowledge assessments was maintained at follow-up testing (Time 3). However, participant performance on the systematic prompting knowledge assessment decreased significantly between post-intervention and follow-up testing; participants scored, on average, 8% lower at follow-up. It must be noted that despite the decline, participants’ knowledge scores remained significantly higher than those recorded at baseline. Therefore, although knowledge maintenance was not achieved for the systematic prompting module, participant knowledge on this topic, nevertheless, improved as a result of Behaviour Skills Training.

As part of the first objective, the relationship between knowledge scores and the following participant characteristics was evaluated: (1) educational level; (2) length of service, and (3) prior training in the target evidence-based practices. There was no relationship between scores on the Test of Knowledge or the knowledge assessments and either length of service with the host organisation or prior training in the target evidence-based practices. As such, participants, regardless of length of service or prior training, benefited equally from the Behaviour Skills Training package. An analysis of the relationship between educational level and knowledge assessment scores yielded similar results. The only exception related to the task
analysis knowledge assessment; participants who had not yet achieved an undergraduate degree consistently scored lower than those with an undergraduate or postgraduate degree. However, there was no interaction between time of measurement and education level, indicating that all educational groups were improving and maintaining knowledge from the task analysis module at a comparable rate.

The second objective focused on investigating the speed of target skill mastery, as well as levels of procedural integrity on completion of each training module. In line with results from the pilot study reported in Chapter 8, participants mastered the implementation of target skills relatively quickly. The average number of observations required to achieve mastery was 3.2 (range = 3-6 observations), 3.4 (range = 3-9 observations), 3.9 (range = 3-6 observations), 3.3 (range = 3-8 observations), and 3.6 (range = 3-7 observations) for the reinforcement, systematic prompting, functional communication training (access to items/activities & access to breaks), and task analysis modules, respectively. These results indicate rapid levels of skill acquisition, considering that the mastery criterion was 90% correct implementation over three consecutive observations.

An analysis of the average percentage of steps completed correctly on the first observation and the final observation for each module was also conducted. Data show that prior to the first feedback session, participants were achieving average procedural integrity levels of at least 87% across all target skills. This figure exceeded 97%, following the rehearsal and feedback sessions. According to results from the randomised control trial reported in Chapter 9, a total of 75.9%, 87%, 85.2%, and 79.6% of participants in the intervention condition stated that they had not received previous training in reinforcement, systematic prompting, functional
communication training and task analysis, respectively. Considering that the majority of participants had no prior training in the target skills, the instruction and modelling components of the current Behaviour Skills Training intervention, appeared to play an important role in the high levels of procedural integrity observed during participants’ first rehearsal of the respective target skills. Previously, it has been the rehearsal and feedback components that have been shown to be critically important in the Behaviour Skills Training approach (van Oorsouw et al., 2009; Ward-Horner & Sturmey, 2012). However, findings from the current study support the position that training outcomes are optimised when all four components of Behaviour Skills Training are used in combination (Barnes et al., 2011; Homlitas et al., 2014).

As discussed earlier in this chapter, in order to demonstrate clinical value and social validity, a training programme must effectively transition between the classroom-based learning of knowledge and skills and the daily practice of staff in their typical work environment (van Oorsouw et al., 2009). Therefore, the third objective of the current study was to examine the generalisation of functional communication training skills from the classroom setting, to the typical workplace environment, with service users. Data show that 29 of the 30 participants who had the opportunity to implement functional communication training with a service user, achieved mastery within three observations and 25 of these participants displayed mastery on the first observation. Although the general training literature reports that successful post-intervention generalisation of trained skills to the workplace is difficult (Ager & O'May, 2001; Campbell, 2010), findings from the present study did not support this position. Rather, the current results are consistent with previous Behaviour Skills Training research, which demonstrate clear evidence of skill
generalisation across people and settings (e.g., Fetherston & Sturmey, 2014; Moore et al., 2002; Sarokoff & Sturmey, 2008).

The final objective of this study was to evaluate participants’ perception of the Behaviour Skills Training package, in terms of its relevancy, usefulness, organisation, and potential for interaction and engagement. Research has shown that “just because a program is considered effective does not mean that it will be considered appropriate by those closely involved in the implementation of the program” (Carter, 2010, p. 7). Furthermore, the training literature has documented “a staff culture frequently hostile to the structure and language of behavioural interventions…the most soundly evidence-based form of intervention” (Ager & O'May, 2001, p. 252). This “culture” was not reflected in feedback from the current study. Instead, the data show that participants evaluated the Behaviour Skill Training package positively, in terms of its relevancy and usefulness in the workplace, with the functional communication training and task analysis modules rated most favourably.

The current study had a number of limitations. A key limitation relates to the evaluation of target skill acquisition over the course of the study. Observations to criterion, a type of event recording, was employed during training to measure the number of practice opportunities required by participants to reach a pre-determined level of proficiency/mastery in the implementation of skills related to each target evidence-based practice (Cooper et al., 2014). However, no measure of procedural integrity across the target skills was recorded prior to initiating Behaviour Skills Training. Therefore, although it was hypothesised that the instruction and modelling components played an important role in the high levels of procedural integrity.
observed during participants’ first rehearsal of the target skills, the impact of a pre-existing skill-set cannot be ignored.

The duration of the follow-up period is considered a second limitation in the current study. Campbell (2010) reported that relatively few studies that examine the effectiveness of training interventions, include an evaluation of the long-term outcomes for frontline staff. As a result, in the current study knowledge maintenance was assessed 4 weeks following the completion of the Behaviour Skills Training programme. In addition, the generalisation of functional communication training skills from the classroom setting, to the typical workplace environment was appraised during this follow-up period. However, the long-term outcomes of the training intervention may be better assessed by including two additional follow-up assessments; one at 6 months and one at 12 months. Longitudinal data of this nature would increase confidence in the clinical value of the training package, both in terms of knowledge maintenance and accurate skill implementation in the workplace (e.g., McClean et al., 2005). An extended follow-up period may also provide an opportunity to assess the generalisation of all target skills, not just functional communication training, to the workplace environment. Therefore, future studies may wish to address this limitation by adopting a longitudinal approach to the research design.

Another limitation of the current study relates to the loss of participants during Phase II, which examined the generalisation of functional communication training skills to the workplace environment. Although positive results were recorded during this phase, a sizeable portion of participants from Phase I did not participate in Phase II. Fifteen participants (27.8%) had no opportunity to display the target skill due to an absence of motivation on the part of their service user to
request. Nine participants (16.7%) were unavailable for observation due to illness, annual leave or service site closure. Rearrangement of the workplace sessions was difficult, due to the limited follow-up period, coupled with the trainers’ demanding work schedules. Therefore, an extended follow-up period may have facilitated the observation of additional participants and should be factored into future research.

Notwithstanding these limitations, results clearly support Behaviour Skills Training as an effective method for disseminating knowledge and skills relating to evidence-based practice, to frontline staff working with adults with intellectual disabilities. However, the extent to which staff training can affect outcomes for people with intellectual disabilities is, as yet, unclear (van Oorsouw et al., 2009). Positive results have been reported in the Behaviour Skills Training literature. For example, Gianoumis et al. (2012) used Behaviour Skills Training to successfully train staff in the appropriate application of the Natural Language Paradigm. Alongside improved teacher performance, a greater frequency of vocalisations and reduced challenging behaviour for the majority of children was observed. Nevertheless, results from the systematic literature review of training practices reported in Chapter 5 indicate that the number of Behaviour Skills Training research studies, which examine the impact of staff training on service user outcomes is relatively sparse. Therefore, the goal of the subsequent study was to investigate the effects of the Behaviour Skills Training intervention reported in this chapter on the quality of support received by service users, as well their functional communication repertoires.
Chapter 11

Dissemination of Evidence-Based Practice to Frontline Staff Working in the Field of Intellectual Disability: An Examination of Service User Impact
Abstract

Although the effectiveness of any staff training intervention is inextricably linked to service user outcomes (Jahr, 1998), a relatively small number of studies have examined the impact of Behaviour Skills Training packages, which are delivered to frontline staff, on such outcomes. As a result, the goal of the current study was to evaluate the consequences for service users, whose direct-support staff had participated in the Behaviour Skills Training intervention, reported in Chapters 9 and 10. Four case studies were conducted. Each case study focused on one service user and involved two direct observations; one occurring the day before the service user’s direct support staff received Behaviour Skills Training and one occurring during the 4-week follow-up period. The first objective was to identify the impact of the staff training intervention on the quality of support received by service users. Consistent with predictions, Active Support Measure ratings were reliably higher for all participants during post-intervention observations. The following areas showed substantial improvement across all or most of the participants: (1) “Sufficient contact by staff”; (2) “Graded assistance”, and (3) “Differential reinforcement of other behaviour”. The second objective of the current study sought to quantify the impact of the Behaviour Skills Training intervention, on the functional communication repertoires of service users. Post-intervention observations showed an improvement in the appropriate communication of service users. However, given the nature of case study design, definite conclusions cannot be drawn from these results. Therefore, future research must emphasise the collateral measurement of service user outcomes, while employing substantial levels of scientific rigor and experimental control.
It has been reported that personal ability is the most significant predictor of quality of life, with individuals who display lower levels of adaptive skills achieving less personal outcomes than their peers with higher functioning levels (Maes et al., 2007; Schalock et al., 1994). Therefore, considering that an intellectual disability is defined as a developmental “disability characterised by significant limitations both in intellectual functioning (reasoning, learning, problem solving) and in adaptive behavior” (American Association on Intellectual and Developmental Disabilities, 2017), people with intellectual disabilities are particularly vulnerable to poorer quality of life outcomes. According to research, this group experiences less choice and variation in their daily activities, have less opportunities for social participation and meaningful engagement, are less likely to live in a typical community setting, and overall lead a less active lifestyle (Beadle-Brown et al., 2015; Maes et al., 2007).

The movement away from larger congregated, institutional settings towards smaller community-based services, has typically resulted in more positive outcomes for people with intellectual disabilities (Emerson & Hatton, 1996; Health Service Executive, 2011). In a review, which focused on deinstitutionalisation across the UK and Ireland, Emerson and Hatton (1996) reported that these community-based services were associated with: (1) increased engagement with community facilities and on-going activities; (2) more contact with staff; (3) improvements in adaptive behaviour, as well as challenging behaviour; (4) increased choice, and (5) a better standard of material living. However, according to Mansell and Beadle-Brown (2012), individuals with less severe disabilities are more likely to experience these positive outcomes, while those with more complex intellectual and developmental needs typically benefit less. Beadle-Brown et al. (2015) propose that individuals
with more severe intellectual disabilities are unable to capitalise on the opportunities provided by this type of service model without the assistance of skilled support staff.

Frontline staff are an extremely valuable asset within the intellectual disability sector; they make it more or less likely that an individual will contact and benefit from functional learning opportunities. Frontline staff ultimately shape the behaviour of the service users they support, through the assistance, reinforcement, and feedback that they provide (Beadle-Brown et al., 2015). Research has shown that compromised skill sets among staff can adversely affect staff-service user relations (Finn & Sturmey, 2009), the learning opportunities available (Schepis et al., 2001), and quality of life outcomes for service users (Jahr, 1998). However, Luckasson et al. (2003) report that when personalised supports, which foster growth, education, interests, and well-being, are in place for an extended period, the functioning of a person with an intellectual disability will subsequently improve.

The standard of service provision within the intellectual disability sector, is dependent on the behaviour of staff who are actually delivering the supports, which in turn is impacted by the quality and relevancy of the training provided to these staff (Reed & Henley, 2015). Although the existing evidence-base for effective training interventions is considered weak (Campbell, 2010), research investigating the effects of Behaviour Skills Training has yielded particularly positive results (e.g., Dib & Sturmey, 2007; Lavie & Sturmey, 2002; Nigro-Bruzzi & Sturmey, 2010; Sarokoff & Sturmey, 2004; 2008). While a limited number of these studies have examined the impact of Behaviour Skills Training on service user outcomes and reported positive results (e.g., Bingham et al., 2007; Dib & Sturmey, 2007; Maggin et al., 2012; McBride & Schwartz, 2003; Nigro-Bruzzi & Sturmey, 2010), findings from the systematic literature review of training practices reported in Chapter 5, indicate that
the number of Behaviour Skills Training research studies, which examine the impact of staff training on service user outcomes is relatively sparse.

As a result, with the effectiveness of the current intervention, including its socially validity, inextricably linked to service user outcomes (Jahr, 1998), a series of case studies were conducted to investigate the impact of the Behaviour Skills Training intervention, reported in Chapters 9 and 10, on service users in the host organisation. Each case study focused on one service user and involved two direct observations; one occurring the day before the service user’s direct support staff received Behaviour Skills Training and one occurring during the 4-week follow-up period. A case study design, rather than a single-subject experimental design, was employed in this instance given the logistical and time constraints controlling the current research programme.

The first objective was to evaluate the quality of support received by service users and it was hypothesised that the standard of service provision would improve following staff engagement in the Behaviour Skills Training intervention. Similarly, the second objective sought to examine service users’ functional communication repertoires. Again it was hypothesised that the use of appropriate communication among service users would increase following staff participation in the Behaviour Skills Training intervention.

**Method**

**Participants**

Participants were four service users whose direct support staff had participated in the Behaviour Skills Training intervention reported in Chapters 9 and 10. Three of the four service users were male and all were aged between 18 and 39
years (M=26.25 years). Two of the services users were diagnosed with moderate intellectual disabilities, one service user had been diagnosed with autism and one service user had a diagnosis of autism and an intellectual disability. Specialised consent forms were designed to facilitate service user understanding that their keyworker(s) would be observed by a trainer and experimenter while they worked with them (available from the author on request). Guardian consent was obtained for those service users unable to provide assent (consent form available from the author on request).

Observer

The lead experimenter from the pilot study and the randomised control trial, which were reported in Chapters 8 and 9, respectively, conducted all observations during this study.

Setting

Observations took place in the settings where participants availed of services delivered by the host organisation. For participants 1 and 4, the settings were community-based day support services and for participants 2 and 3, the settings were residential in nature.

Measures

**Active Support Measure (ASM) (Mansell & Elliott, 1996).** The Active Support Measure is a tool used to evaluate the quality of support that a particular service user is receiving from their frontline staff. The Active Support Measure, which has demonstrated good internal consistency (e.g., Beadle-Brown, Hutchinson, & Whelton, 2012) contains 15 items, which examine the opportunities for service user engagement, as well as the skills used by staff to facilitate and foster this
involvement. Each service user is observed for 2 hours and the Active Support Measure is completed at the end of the 2-hour observation period. Each of the following sections received a rating between 0 (poor, inconsistent support/performance) and 3 (good, consistent support/performance): (1) Age appropriateness of activities and materials; (2) ‘Real’ rather than pretend or very simple activities; (3) Choice of activities; (4) Demands presented carefully; (5) Tasks appropriately analysed to facilitate service user involvement; (6) Sufficient staff contact for service user; (7) Graded assistance to ensure service user success; (8) Speech matches developmental level of service user; (9) Interpersonal warmth; (10) Differential reinforcement of adaptive behaviour; (11) Staff notice and respond to service user communication; (12) Staff manage serious challenging behaviour well; (13) Staff work as a coordinated team to support service users; (14) Teaching is embedded in everyday activities, and (15) Written plans in routine use.

**Functional Communication Assessment.** A data sheet was designed (available from the author on request) for the purposes of this study to record instances of appropriate and inappropriate requests displayed by the participants. A requesting opportunity was defined as a participant demonstrating overt signs of motivation to either access items/activities/attention or escape from a particular item/activity/interaction. These overt signs of motivation could include gestures, reaching actions, particular body postures (e.g. orienting towards a preferred item/activity or moving body away from a non-preferred item/activity/person) and vocalisations. An instance of an appropriate request was defined as the emission of a relevant word, phrase, gesture, sign or picture exchange that would be intelligible to an unfamiliar communication partner, given a requesting opportunity. An instance of an inappropriate request was defined as the absence of the emission of a relevant
word, phrase, gesture, sign or picture exchange that would be intelligible to an unfamiliar communication partner, given a requesting opportunity. For every requesting opportunity, the observer recorded if the behaviour displayed met the criteria for an appropriate or an inappropriate request.

**Design**

A case study research design was used to evaluate the impact of the Behaviour Skills Training intervention, which is reported in Chapters 9 and 10, on service users whose direct support staff had received this training package.

**Procedure**

Each case study focused on one participant and involved two direct observations; one occurring the day before a participant’s direct support staff received Behaviour Skills Training and one occurring during the 4-week period, post-training. Observations were scheduled for a 2-hour period and the observer shadowed the participant, at a distance, throughout this time. Instances of appropriate and inappropriate functional communication were recorded by the observer in real-time. The Active Support Measure was completed by the observer at the end of each 2-hour observation period.

**Results**

**Case Study 1**

**Background information.** Cillian was a 26 year old male, diagnosed with moderate intellectual disabilities. He lived at home with his parents and attended a ‘hub service’, which was linked to a community-based day service site. In the ‘hub service’, support was provided to Cillian to help foster his independence and
integration into community activities. According to results on the Vineland Adaptive Behaviour Scale (VABS), Cillian displayed adequate expressive and receptive communication skills.

**Area(s) of concern.** According to staff reports and direct observations, Cillian displayed inconsistencies in requesting permission from others to engage in/escape from activities that impacted the overall peer group (e.g., eloping to the shop while his peer group were waiting for him, in order to depart on the bus to attend a scheduled activity).

**Intervention.** Staff supporting Cillian received a 3-day training programme in evidence-based practices (i.e., reinforcement, systematic prompting, functional communication training, and task analysis). Training was delivered using the Behaviour Skills Training method.

**Direct observations.** Cillian was observed on two separate occasions, for a 2-hour period, while preparing for a recreational activity (swimming), taking the bus to a local leisure centre, using the facilities in the leisure centre and returning to the ‘hub’ setting.

**Active Support Measure.** Figures 33 and 34 show Cillian’s scores on each item of the Active Support Measure pre- and post-intervention. Data revealed that the quality of support that Cillian received improved post-intervention, with 12 out of 15 items scoring a ‘3’ (maximum score) post-intervention. This is compared to 5 items scoring a ‘3’ pre-intervention. The areas demonstrating the most improvement were: (1) “The appropriate analysis of tasks to facilitate service user engagement”; (2) “Sufficient staff contact for service user”; (3) “Graded assistance to ensure
service user success”, and (4) Embedding teaching opportunities in everyday activities

**Note.** Cillian did not engage in any instances of serious challenging behaviour. Therefore, no scores were recorded for the “Management of Challenging Behaviour” item pre- and post-intervention.

**Figure 33.** Pre- and post-intervention scores on the first eight items of the Active Support Measure.
Figure 34. Pre- and post-intervention scores on the final seven items of the Active Support Measure.

Functional Communication Assessment. Data from the communication observations revealed that Cillian displayed 49 appropriate requests out of a total of 74 requesting opportunities pre-intervention and 51 appropriate requests out of a total of 52 requesting opportunities post-intervention. This represented an percentage increase of 31.9% in appropriate requests (66.2% pre-intervention to 98.1% post-intervention). Similarly, inappropriate requests reduced from 25 out of 74 requesting opportunities (33.8%) pre-intervention to 1 out of 52 requesting opportunities (1.9%) post-intervention.
**Conclusion.** Data from pre- and post-intervention observations tentatively show an improvement in the quality of support received by Cillian, in addition to an improvement in his requesting skills.

**Case Study 2**

**Background information.** Michael was an 18 year old male, diagnosed with autism spectrum disorder, hydrocephalus and dystonia. He lived alone with the constant support of at least two direct support staff. Michael did not attend an educational setting or day support service site and according to the Assessment of Functional Life Skills (AFLS) he displayed limited expressive communication skills but comparatively good receptive skills. A review of files revealed that Michael inconsistently communicated using one-word vocal requests and had a history of using the Picture Exchange Communication System (PECS).

**Area(s) of concern.** According to staff reports and direct observations (i.e., narrative Antecedent-Behaviour-Consequence (ABC) records), Michael engaged in serious challenging behaviour, including aggression toward others and self-injurious behaviour. These behaviours were hypothesised to serve a communicative function to gain access to desired items/activities or to escape or avoid unwanted situations/interactions. His existing behaviour support plan directed staff to facilitate learning opportunities for the following skills: 1) asking for help; 2) indicating too much/too little interpersonal interaction, and 3) making choices.

**Intervention.** Staff supporting Michael received a 3-day training programme in evidence-based practices (i.e., reinforcement, systematic prompting, functional communication training, and task analysis). Training was delivered using the Behaviour Skills Training method.
**Direct observations.** Michael was observed in his own home on two separate occasions.

**Active Support Measure.** Figures 35 and 36 show Michael’s scores on each item of the Active Support Measure pre- and post-intervention. During the pre-intervention observation, Michael was observed to be either sleeping or lying awake in bed for the 2-hour period. Therefore, a score of ‘0’ was assigned to each item of the Active Support Measure. During the post-intervention observation, Michael was again observed to be sleeping. However, on this occasion, staff immediately attempted to engage Michael in a number of activities (e.g., mealtime, watching a preferred DVD) having observed overt signs of waking. As such, the quality of support provided to Michael was substantially higher across all items of the Active Support Measure during the post-intervention observation. Data revealed that 10 out of the 15 Active Support Measure items were scored as a ‘3’ (maximum score) post-intervention. These 10 items were as follows: (1) “Choice of activities”; (2) “Demands presented carefully”; (3) “Sufficient staff contact”; (4) “Graded assistance frequently given”; (5) “Speech matches developmental level of service user”; (6) “Interpersonal warmth”; (7) “Differential reinforcement of other behaviour”; (8) “Staff notice and respond to service user behaviour”; (9) “Staff work as a team”, and (10) “Teaching embedded in everyday activities”.

**Note.** Michael did not engage in any instances of serious challenging behaviour. Therefore, no scores were recorded for the “Management of Challenging Behaviour” item pre- and post-intervention.
Figure 35. Pre- and post-intervention scores on the first eight items of the Active Support Measure.
Figure 36. Pre- and post-intervention scores on the final seven items of the Active Support Measure.

Functional Communication Assessment. Post-intervention, Michael displayed six appropriate vocal requests for various items and actions and did not resort to inappropriate requests or challenging behaviour as a means of communication. Three of these requests consisted of two or more words. Considering that Michael’s file indicated that he was inconsistently using one-word requests pre-intervention, this is a noticeable improvement. However, given that there was no data pre-intervention, no direct comparison could be made.

Conclusion. Observations provide tentative support for an improvement in Michael’s requesting skills and the quality of support he received post-intervention. However, no definitive conclusions can be drawn given the nature of the data, particularly the limitations to the pre-intervention data.
Case Study 3

**Background information.** Orla was a 22 year old female, diagnosed with autism spectrum disorder and a severe intellectual disability. She lived with three adult males and her twin sister and was supported by one staff member at all times. Orla attended a day community-based day support service site for 3 hours per day, Monday to Friday. According to keyworker reports and direct observations she displayed limited expressive communication skills but comparatively good receptive skills.

**Area(s) of concern.** Staff-reports and direct observations indicated that Orla frequently engaged in relatively low-level-intensity challenging behaviour (i.e., grabbing items and people, spitting, kicking, poking and high-pitched vocalisations). These behaviours were hypothesised to serve a communicative function to gain access to desired items/activities/interactions.

**Intervention.** Staff supporting Orla received a 3-day training programme in evidence-based practices (i.e., reinforcement, systematic prompting, functional communication training, and task analysis). Training was delivered using the Behaviour Skills Training method.

**Direct observations.** Orla was observed in her own home on two separate occasions.

**Active Support Measure.** Figures 37 and 38 display Orla’s scores on each item of the Active Support Measure pre- and post-intervention. Data revealed that the quality of support that Orla received improved post-intervention, with 9 out of 15 items scoring a ‘3’ (maximum score) post-intervention. This is compared to no items receiving a score of ‘3’ pre-intervention. Records showed an obvious transition
away from Orla’s passive participation in everyday activities (e.g., having dinner prepared for her and given to her at the table, having tea prepared for her) to active engagement in these activities, with systematic assistance from direct support staff. The items demonstrating the most improvement were: (1) “Demands presented carefully”; (2) “The appropriate analysis of tasks to facilitate service user engagement”; (3) “Sufficient staff contact”; (4) “Graded assistance frequently given”; (5) “Differential reinforcement of other behaviour”; (6) “Staff notice and respond to service user behaviour”; (7) “Good management of challenging behaviour”; (8) “Staff work as a team”, and (9) “Teaching embedded in everyday activities”.

![Figure 37](image)

**Figure 37.** Pre- and post-intervention scores on the first eight items of the Active Support Measure.
Figure 38. Pre- and post-intervention scores on the final seven items of the Active Support Measure.

Functional Communication Assessment. Data from the communication observations showed that Orla displayed 8 appropriate requests out of a total of 57 requesting opportunities pre-intervention (14%) and 49 inappropriate requests out of a total of 57 requesting opportunities pre-intervention (86%). However, based on the recommendations of the supervising behaviour therapist, the introduction of a functional communication plan for Orla was delayed until staff had successfully implemented a similar plan with her twin sister. The rationale behind this decision will not be addressed in detail here but it is suffice to say that the behavioural issues in Orla’s case were more complex and it was thought that the plan would be more
successful in the long-term if staff had the opportunity to carry out a similar plan in a
less complex case initially. As a result, there is no post-intervention data available
for Orla’s communication skills.

**Conclusion.** Data from pre- and post-intervention observations tentatively
show an improvement in the quality of support received by Orla. For reasons beyond
the control of the research programme, post-intervention data on Orla’s
communication skills was not recorded. This prevented pre-post-intervention
comparisons from being carried out.

**Case Study 4**

**Background information.** Peter was a 39 year old male, diagnosed with a
moderate intellectual disability and co-morbid mental health issues (not yet
specified). He lived in supported living accommodation with three adult males and
availed of the supports offered in the community-based day support service site.
According to reports from the manager of the day support service site, Peter
displayed adequate expressive and receptive communication skills.

**Area(s) of concern.** Reports from keyworkers and direct observations
revealed that Peter frequently ruminated on negative topics of conversation (e.g.,
health risks of smoking, his inability to quit smoking, family conflict, lack of
stimulation), displayed a reluctance to decide how to spend his leisure time (e.g.,
choosing which coffee shop he would like to go to), and experienced difficulty
expressing his desire not to engage in a particular activity (e.g., his voice became
strained and his body rigid and he stated that he “can’t do X because he isn’t able”
rather than calmly saying “I don’t want to do X”).
**Intervention.** Staff supporting Peter received a 3-day training programme in evidence-based practices (i.e., reinforcement, systematic prompting, functional communication training, and task analysis). Training was delivered using the Behaviour Skills Training method.

**Direct observations.** Peter was observed, on two separate occasions, engaging in a community outing into the city centre. This was facilitated by direct support staff in the day support service site. The trip itself involved taking a private bus or walking to the city centre, buying cigarettes, going for coffee, and carrying out errands for the manager in the day service site.

**Active Support Measure.** Figures 39 and 40 display Peter’s scores on each item of the Active Support Measure pre- and post-intervention. According to the Active Support Measure ratings pre-intervention, the quality of support that Peter was receiving was already quite high, with 10 of the 15 items scoring a ‘3’ (maximum score). Peter had not engaged in challenging behaviour pre-intervention. Post-intervention, the areas in which improvement was demonstrated were: (1) “The differential reinforcement of other behaviour”, and (2) “Graded assistance to ensure service user success”.

*Note:* Peter did not engage in any instances of serious challenging behaviour during pre-intervention observations. Therefore, no scores were recorded for the “Management of Challenging Behaviour” item pre-intervention. However, he did elope from the service site during the post-intervention observation and this was deemed a serious challenging behaviour by the staff team. Therefore, based on staff performance, a score of ‘3’ was given for the “Management of Challenging Behaviour” item post-intervention.
Figure 39. Pre- and post-intervention scores on the first eight items of the Active Support Measure.
Conclusion. Although the quality of service provision was already quite high, data from pre- and post-intervention observations tentatively show progress in
the quality of support received by Peter, in addition to an improvement in his requesting skills.

**Discussion**

The goal of the current study was to examine the impact on service users, whose direct-support staff had participated in the Behaviour Skills Training intervention, reported in Chapters 9 and 10. The first objective was to evaluate the effects of the staff training intervention on the quality of support received by service users. Consistent with predictions, Active Support Measure ratings were reliably higher for all participants during post-intervention observations. The following areas showed substantial improvement across all or most of the participants: (1) “Sufficient contact by staff”; (2) “Graded assistance”, and (3) “Differential reinforcement of other behaviour”. These areas are directly related to content covered during the evidence-based practice modules, with the reinforcement and systematic prompting modules having particular relevance. Considering that research shows that frontline staff actively shape the behaviour of the service users they support, through the assistance, reinforcement, and feedback that they provide (Beadle-Brown et al., 2015), the improvements observed are significant for the service users affected.

The second objective of the current study sought to quantify the impact of the Behaviour Skills Training intervention, on the functional communication repertoires of service users. Observations, supported the hypothesis that the use of appropriate communication would increase among service users, following the Behaviour Skills Training intervention. Compared to pre-intervention, Cillian and Peter both reduced their display of inappropriate communication, while simultaneously expanding on their use of appropriate communication, during the post-intervention observation. Similarly, Michael displayed six appropriate vocal requests for various items and
actions and did not resort to inappropriate requests or challenging behaviour as a means of communication, post-intervention. In addition, three of the six requests consisted of two or more words. This was a significant improvement for Michael, given reports that he was inconsistently relying on one-word requests to communicate, pre-intervention.

However, the current study is limited, in terms of its research design. Firstly, given the nature of support provided by frontline staff during the pre-intervention observation, no opportunities to communicate arose for Michael. As a result, comparisons of Michael’s pre- and post-intervention functional communication repertoires cannot be made. Furthermore, no functional communication assessment data was available for Orla, either pre- or post-intervention. A substantial criticism of case study design is the limited external validity of findings (Zainal, 2007). Therefore, incomplete data for two of the four participants, further compromises the already restricted generalisability of this study.

Another critique of case study design is the insufficient scientific rigor and experimental control that is demonstrated (Zainal, 2007). Inadequate methodological rigor precludes researchers from drawing definite conclusions from research findings (National Autism Center, 2009; 2015). Therefore, while results from the current study appear to be supportive of the Behaviour Skills Training intervention, findings may be compromised by unidentified and uncontrolled confounding variables. As a result, it is imperative that future Behaviour Skills Training research, not only emphasises the collateral measurement of service user outcomes, but also employs substantial levels of scientific rigor and experimental control during this process.
Chapter 12

General Discussion
According to the most recent annual report, there were 28,275 people registered with the Irish National Intellectual Disability Database (NIDD) at the end of December 2016 (Doyle et al., 2017). Records indicate that 98.5% of persons registered with the NIDD are currently in receipt of some type of support, and of those availing of support, 27,813 (99.8%) have participated in at least one “day programme” in 2016. This represents the highest percentage since the register was first developed in 1996. In addition, 7,582 of these individuals are also receiving full-time residential care (Doyle et al., 2017). Therefore, with a growing population of adults with moderate to severe intellectual disabilities (Doyle et al., 2017) and an increasing demand for day services and residential care, there is an ever-growing need for high quality, evidence-based supports for people with an intellectual disability in the Republic of Ireland.

However, the quality of service provision in the field of intellectual disabilities can vary substantially across providers in the Republic of Ireland (National Economic and Social Council, 2012). According to the National Economic and Social Council (2012), “services have been free to develop as independent republics, providing care as they see fit, with little protection for vulnerable groups” (p. xiii). In 2002, the Joint Committee on Social Care Professionals in the Republic of Ireland published a report outlining the qualifications of staff employed in residential and community-based childcare settings, as well as residential settings for people with intellectual disabilities (Joint Committee on Social Care Professionals, 2002). Over half of the 2,791 participants sampled were employed by intellectual disability providers and 36% of these individuals did not possess a qualification in social care.
Furthermore, it is estimated that 71% (19,975) of persons registered with the National Intellectual Disability Database (NIDD) will require new or improved specialised supports during the 2017-2021 period (Doyle et al., 2017). However, restricted access to skilled professionals, including speech and language therapists, psychologists and psychiatrists (Fallon et al., 2017; McClean et al., 2005; Yacoub, 2015), may result in insufficient support and training for frontline staff and a continued reliance on the “caring” approach rather than active support of service users (Linehan et al., 2014; Tatlow-Golden et al., 2014). Considering that compromised skill sets among intellectual disability staff has been associated with significant adverse repercussions for service users (e.g., Finn & Sturmey, 2009; Jahr, 1998; Schepis et al., 2001), the lack of professional qualifications in this sector in the Republic of Ireland is concerning.

However, research has shown that staff with varying backgrounds and educational qualifications can be effectively trained to carry out procedures in line with evidence-based practice (Allen & Tynan, 2000; Dench, 2005; Grey & Hastings, 2005). In turn, training in evidence-based practice can enhance service user outcomes by improving the interactions between staff and service users (Emerson et al., 2000) and increasing the learning opportunities available (Rose et al., 1998). Given that the intellectual disability sector in the Republic of Ireland is currently in transition (National Economic and Social Council, 2012), there is an opportunity for innovation and positive change.

As a result, the current research programme aimed to examine the training needs of frontline staff employed by one of the largest intellectual disability service providers in the Republic of Ireland and subsequently develop an effective and socially valid method for disseminating a package of evidence-based practices, in
order to prepare these frontline staff to better support their service users with intellectual disabilities. The objectives were as follows: (1) to conduct a systematic literature review of the training provided to staff working in the intellectual disability sector; (2) to establish a training needs analysis among a representative sample of frontline staff and their service users within the host organisation. This analysis sought to identify areas of need for service users, by acquiring information on current levels of adaptive behaviour skills, quality of life, and management of challenging behaviour through staff reports of training and work-related experiences; (3) to design and employ a Behaviour Skills Training intervention, matched to the training needs of the host organisation and founded in evidence-based practice, and (4) evaluate the impact of this intervention on a range of staff- and service user-based outcomes.

Contributions of Current Research Programme and Overview of Findings

Overview of Study 1 (Chapter 5)

The first objective of this programme of research was to conduct a systematic literature review of the training provided to staff working in the intellectual disability sector (Chapter 5). The review comprised 107 studies, published between the years 2000 and 2015. Results show that while the range of practices targeted for training were generally broad, training in the application of many established evidence-based practices was not addressed. For example, functional communication training is an evidence-based treatment for individuals with intellectual disabilities (Kurtz et al., 2011). However, the systematic literature review did not source any studies, which provided training in the implementation of this practice. The following evidence-based behavioural interventions, were also notably absent from the research studies
included in the review: a) shaping; b) discrimination training; c) reinforcement schedules; d) differential reinforcement techniques, and e) non-contingent reinforcement. Therefore, despite rigorous and consistent scientific support for these strategies, the literature did not concurrently evaluate training protocols to successfully disseminate these practices to frontline staff. Such findings provide a potential explanation for the current disconnect between theory and effective practice in the field of intellectual disability (Campbell, 2010; Hile & Walbran, 1991; Rapp et al., 2010; Swain et al., 2010).

**Overview of Study 2 (Chapter 6)**

The second objective of the current research programme was to establish a training needs analysis among a representative sample of frontline staff and their service users within RehabCare. RehabCare is one of the largest providers of intellectual disability services in the Republic of Ireland. This organisation provides a range of health and social care services to more than 3,000 people with an intellectual disability across 88 geographical locations throughout the country. Over 1,000 staff are employed to deliver these services, across a variety of community-based applied settings, including resource centres and residential settings. The goal of the training needs analysis was to obtain a broad, comprehensive analysis of the diagnostic, adaptive functioning, and behavioural profiles of a representative sample of service users, as well as the training needs of staff providing direct care to these individuals. The information yielded from this analysis would be utilised in designing a subsequent staff training programme to address deficient skills in selecting and applying evidence-based practices for persons with intellectual disability.
Results from the training needs analysis demonstrated a disconnect between the training received by staff and the adaptive and behavioural needs of the service users. Communication, social skills, and daily living skills were identified as priority areas for service users within the host organisation. However, the majority of staff reported receiving no preparatory training from RehabCare, in either behaviour management strategies or skills teaching procedures, prior to starting work with service users and among those who did receive preparatory training, over 50% did not believe that it adequately equipped them to support service users. Furthermore, less than 40% of staff received ongoing training in how to teach their service users new skills, despite 93.5% of staff surveyed being employed with RehabCare for at least a year. Given that applied behaviour analytic research has developed a range of empirically tested interventions that target deficits in communication, social skills, and daily living skills (Carr et al., 2009; Grey & Hastings, 2005; Kurtz et al., 2011; National Autism Center, 2009; 2015; Wong et al., 2014), staff training in these evidence-based practices would assist frontline staff in supporting service users to achieve a quality of life commensurate with their peers.

Overview of Study 3 (Chapter 7)

The next objective focused on the design and implementation of a Behaviour Skills Training intervention, matched to the training needs of the host organisation and founded in evidence-based practice. As discussed in Chapter 3, the National Autism Centre and the National Professional Development Centre independently conducted two large-scale evaluations of interventions for children, adolescents, and adults with autism (National Autism Center, 2009; 2015; Wong et al., 2014). According to the results reported by these projects, the following behavioural interventions are considered established evidence-based practices for addressing
deficits in communication, social, and daily living skills among people with autism: (1) reinforcement; (2) systematic prompting; (3) functional communication training, and (4) task analysis. However, a similar large-scale analysis of interventions for children and adults with intellectual disabilities was not evident in the extant literature. Therefore, results from the National Standards Project (National Autism Center, 2009; 2015) and the National Development Centre report (Wong et al., 2014) were used to direct a series of systematic literature reviews (Chapter 7), which evaluated the evidence-base supporting the use of reinforcement, systematic prompting, functional communication training, and task analysis with people with intellectual and developmental disabilities. Results showed a substantial evidence-base supporting the use of these behavioural interventions with this population.

**Overview of Studies 4 and 5**

According to results from the systematic literature review reported in Chapter 5, there is limited research examining the effective dissemination of evidence-based practices to frontline staff in the intellectual disability sector. Therefore, an effective method of dissemination must be identified to bridge the theory-practice gap (Campbell, 2010; Jahr, 1998). In 2009, van Oorsouw and colleagues conducted a meta-analysis to investigate the structural components necessary for effective staff training (van Oorsouw et al., 2009). The authors concluded that optimal results are achieved through the combined use of in-service training and on-the-job coaching, with target information being delivered to staff in a number of different ways. Furthermore, Reid et al. (2003) suggest that staff training should be performance- and competency-based as knowledge acquisition alone will not produce improved staff-service user interactions (Embry & Biglan, 2008; Grey et al., 2007); there must also be an emphasis on skill acquisition (Parsons et al., 2012).
Behaviour Skills Training, an intervention that utilises a combination of instruction, modelling, rehearsal, and feedback, meets these criteria (Reid et al., 2003; van Oorsouw et al., 2009), and has been successfully applied to train a wide range of groups, including frontline staff, a variety of new work-related skills (e.g., Homlitas et al., 2014; Maggin et al., 2012; Parsons et al., 2012; Sarokoff & Sturmey, 2008). However, to the author’s knowledge the efficacy of Behaviour Skills Training had never previously been tested in a large group-based experimental design nor had it been evaluated in the context of delivering the following combination of evidence-based practices to frontline staff, working in the intellectual disability sector: reinforcement, systematic prompting, functional communication training, and task analysis.

**Study 4 (Chapter 8).** Therefore, the principal aim of the current research programme was to conduct a clustered randomised control trial to comprehensively evaluate the effectiveness of Behaviour Skills Training as a method for disseminating this package of evidence-based practices to frontline staff working with adults with intellectual disabilities. A pilot study was conducted, prior to the randomised control trial, to assess the feasibility of implementing an evidence-based Behaviour Skills Training package within the host organisation. Results showed that the experimental design and associated methodology were robust in measuring outcomes related to staff participants, and the intervention, itself, could be successfully delivered within the allocated time-period.

**Study 5 (Chapters 9, 10, and 11).** The fourth and final objective of the present programme of research was to evaluate the impact of the Behaviour Skills Training intervention on a range of staff- and service user-based outcomes. In the interest of clarity, Chapter 9 compared the Behaviour Skills Training intervention to
a wait list control condition to evaluate the differential impact on target knowledge acquisition and psychological well-being among frontline staff, while Chapter 10 introduced an analysis of outcomes for the intervention group only. One hundred and four frontline staff were recruited from twelve service sites within the RehabCare organisation; 54 participants were assigned to the Behaviour Skills Training condition and 50 participants were assigned to the wait list control condition. Results showed that participants who received Behaviour Skills Training demonstrated statistically significant improvements on four knowledge assessments and the Test of Knowledge at follow-up testing. In contrast, participants in the wait list control condition showed either no change or a statistically significant decrease in their knowledge scores over the study period. No statistically significant changes to psychological well-being were observed for either experimental group and no salient barriers to dissemination were reported by participants.

The goal of Chapter 10 was to provide a more rigorous analysis of the impact of the Behaviour Skills Training on the intervention group within the randomised control trial. Results show that participant performance across all four knowledge assessments and the Test of Knowledge improved significantly between baseline and post-intervention testing and performance was maintained at the 4-week follow-up across the reinforcement, functional communication training, and task analysis assessments, as well as the Test of Knowledge. Further analyses found that regardless of length of service, educational attainment or prior training, participants benefited equally from the Behaviour Skills Training package. Results also provided clear evidence of skill acquisition and generalisation of functional communication training skills to the workplace environment.
Finally, the goal of Chapter 11 was to evaluate outcomes for service users, whose direct-support staff had participated in the Behaviour Skills Training intervention, reported in Chapters 9 and 10. Four case studies were conducted; each focused on one service user. The aim of the case studies was to examine service users’ functional communication repertoires and assess the quality of support they received from frontline staff, pre-and post-intervention. There was evidence of improvement across both outcomes. The frequency of appropriate communication increased post-intervention and the following areas of service provision showed substantial improvement: (1) “Sufficient contact by staff”; (2) “Graded assistance”, and (3) “Differential reinforcement of other behaviour” (Mansell & Elliott, 1996).

Limitations, Challenges, and Future Directions

Notwithstanding the results already outlined, which clearly support Behaviour Skills Training as an effective method for disseminating knowledge and skills relating to evidence-based practice, to frontline staff working with adults with intellectual disabilities, the current research programme has a number of key limitations. The following section elaborates on these limitations, while also suggesting directions for future research.

Duration of Follow-up Period

Relatively few studies that examine the effectiveness of training interventions, include an evaluation of the long-term outcomes for frontline staff (Campbell, 2010). In an attempt to address this shortcoming, participants’ psychological well-being and knowledge maintenance was assessed 4 weeks following the completion of the Behaviour Skills Training programme. In addition, the generalisation of functional communication training skills from the classroom
setting, to the typical workplace environment was appraised during this follow-up period. However, the long-term outcomes of the training intervention may be better assessed by including two additional follow-up assessments; one at 6 months and one at 12 months. Firstly, longitudinal data of this nature would increase confidence in the clinical value of the training package, both in terms of knowledge maintenance and accurate skill implementation in the workplace (e.g., McClean et al., 2005). In addition, an extended follow-up period may also provide an opportunity to assess the generalisation of multiple target skills to the workplace environment. Secondly, considering the complexity of the psychological constructs being assessed during the randomised control trial reported in Chapter 9 (Fukui et al., 2014; Lazarus, 1999; Locke, 1969; Maslach, 1993; Maslach, 1999), a follow-up period of at least 6 months may be necessary to observe measurable changes in these variables (e.g., Hätinen et al., 2013). Therefore, future research must adopt a longitudinal approach, in order to adequately assess the long-term impact of Behaviour Skills Training on relevant staff- and service user-based outcomes.

**Data Analysis Methodologies**

**Evaluation of psychological variables.** Throughout the current programme of research, all psychological variables of interest were assessed through participant self-report. However, it is widely recognised that respondents are often reluctant or unable to accurately report on sensitive topics due to “egodefensive or impression management reasons” (Fisher, 1993, p. 303). This phenomenon is known as the social desirability effect and can result in a pattern of responding that is systematically biased toward what the respondent perceives a socially acceptable (Fisher, 1993). Considering the nature of the psychological variables assessed in the current research programme, including job satisfaction, perceived supervisory
support, levels of burnout, and service user quality of life, there is a significant risk that social desirability may have impacted results. Attempts were made to minimise this possibility by prioritising participant anonymity and assuring participants that the raw data collected would only be reviewed by the first and second experimenter. Nonetheless, given that the social desirability bias is a significant issue across the social science literature (Fisher, 1993), it is likely that this source of bias impacted upon the results of the training needs analysis, pilot study, and randomised control trial studies. This issue of bias may be addressed by assessing the psychological variables of interest through a variety of methods, including but not restricted to participant self-report. This approach was not taken in the current programme of research, thereby introducing a second source of bias; single source bias (Campbell & Fiske, 1959). Single source bias relates to a reliance on data collected from a single source (Campbell & Fiske, 1959) and future research must account for this type of bias, in order to address the potential impact of subjective responding and increase confidence in the results obtained.

**Evaluation of participant skill acquisition.** A further limitation may include failing to record participants’ ability to implement skills related to each target evidence-based practice prior to engaging in the Behaviour Skills Training intervention. Although observations to criterion, a type of event recording, facilitated the assessment of the number of practice opportunities required by participants to reach a pre-determined level of proficiency in the implementation of the target skills (Cooper et al., 2014), without baseline data, the full extent of post-intervention performance improvement cannot be evaluated. Furthermore, data from the rehearsal and feedback sessions show procedural integrity levels of at least 87% for each of the target evidence-based practices prior to the first feedback session.
However, 75.9%, 87%, 85.2%, and 79.6% of staff who participated in the Behaviour Skills Training intervention reported that they had not received previous training in reinforcement, systematic prompting, functional communication training, and task analysis procedures, respectively. Therefore, although it would appear that the instruction and modelling components of the intervention have an important role in facilitating skill development, without baseline data the impact of potentially pre-existing skills cannot be disregarded. The measurement of baseline skill performance within a large-scale group design may be logistically and financially challenging (Proctor et al., 2011). However, staff training should be performance- and competency-based (Reid et al., 2003). Knowledge acquisition alone will not produce improved staff-service user interactions (Embry & Biglan, 2008; Grey et al., 2007); training programmes must focus on the development and measurement of new target skills (Parsons et al., 2012). Therefore, future group-based research should endeavour to address this within the experimental design, to allow for a more comprehensive assessment of Behaviour Skills Training.

**Evaluation of service user outcomes.** The reliance on case studies to evaluate the impact of the Behaviour Skills Training intervention on service user outcomes may also be considered a significant limitation of the current research programme. Case study design is criticised for demonstrating poor experimental control and limited external validity (Zainal, 2007), which prevents consumers from drawing definite conclusions from research findings (National Autism Center, 2009; 2015). Therefore, while results from the case studies reported in Chapter 11 appear to be supportive of the Behaviour Skills Training intervention, findings are inconclusive given the nature of the research design. As discussed in Chapter 4, Callahan et al. (2016) have proposed a number of indicators to optimise the social
validity of experimental evaluations, including: (1) were socially important
dependent variables selected?, and (2) was the impact of the training package
assessed in the natural context? Given that the purpose of training is not only to
upskill frontline staff in effective practice (Allen & Tynan, 2000; Dench, 2005; Grey
& Hastings, 2005), but also to enhance service user outcomes (e.g., Emerson et al.,
2000; Jahr, 1998; Rose et al., 1998), the impact of a training intervention on service
users should be considered a socially important dependent variable and prioritised for
assessment in the natural context. Therefore, it is imperative that future Behaviour
Skills Training research, not only emphasises the measurement of service user
outcomes, but applies the highest level of scientific rigor during the process.

Sampling

The reliance on convenience sampling of service sites during the pilot study
and randomised control trial, reported in Chapters 8 and 9, respectively, represents
another key limitation. Participating service sites were not randomly selected from
all available service sites within the RehabCare organisation. For both the pilot
study and the randomised control trial inclusion criteria were minimally restrictive
and selection was determined by logistical constraints, such as a service site’s ability
to facilitate a service closure for 3 consecutive days. The difficulties experienced,
with regards to service site recruitment, precluded the use of more stringent inclusion
criteria as was orginally intended. As a result, this reliance upon convenience
sampling may have inadvertantly introduced selection bias into the current research
programme, which may limit the external validity of the results. In addition, the
generalisabilty of findings may be further compromised given that all participating
service sites were sampled from a single intellectual disability service provider
organisation. Therefore, future research must address these limitations in order to definitively demonstrate the robustness of the Behaviour Skills Training intervention.

**Role of Supervisory Support in Successful Training Outcomes**

Finally, successful and enduring change within an organisation cannot be achieved by targeting frontline staff alone (Cullen, 2000; Rapp et al., 2010; Swain et al., 2010). Ongoing training is required for professionals at all levels, particularly for those in managerial and supervisory positions (DiGennaro Reed & Henley, 2015). Swain et al. (2010) reported that across 49 mental health organisation in America, the largest differences between “sustainers” (continued use of evidence-based practices over a 2-year period) and “non-sustainers” (failure to implement evidence-based practices 2 years after initial implementation) were evident on staff training measures, followed closely by scores on staff supervision measures. Likewise, Rapp et al. (2008) found that high quality supervision from programme leaders plays a critical and unique role in the context of evidence-based service provision. Programme leaders have a responsibility to guide, develop, supervise and enforce evidence-based practice on a daily basis (Rapp et al., 2008).

Therefore, in addition to effective training, the provision of high standard, empirically validated supervision for frontline staff is critical, not only to ensure the fidelity and consistency of evidence-based practice implementation over time, but also to ensure the best possible outcomes for service users. Although results from the randomised control trial reported in Chapter 9 show no relationship between perceptions of supervisory support and knowledge of target evidence-based practices, future research must investigate the factors related to effective supervision, considering its role in sustaining the accurate implementation of evidence-based practice among frontline staff (Rapp et al., 2008; Swain et al., 2010).
Chapter 12

**Final Conclusions**

Notwithstanding these limitations, results from the current research programme clearly demonstrate the effectiveness of Behaviour Skills Training in disseminating knowledge and skills relating to evidence-based practice to frontline staff working with adults with intellectual disabilities. From the outset, the following social validity indicators were used to guide experimental evaluations of the training intervention: (1) were socially important dependent variables selected?; (2) were consumers satisfied with the results?; (3) was it a time and cost effective training package?; (4) was the impact of the training package assessed in the natural context?; (5) did the training package impact the behaviour of frontline staff, and (6) were clinically significant levels of behaviour change achieved (Callahan et al., 2016)?

Therefore, all efforts were made to select socially significant dependent variables by focusing on skill development, as well as knowledge acquisition among frontline staff. In addition, an evaluation of service user outcomes was conducted, even though the research design may not have been optimal. For participants who engaged in the Behaviour Skills Training intervention, clinically significant levels of target behaviour change were observed both in the classroom and in the natural context with service users. Finally, staff who received the training intervention consistently rated the evidence-based practice modules positively.
Dissemination of evidence-based practice to frontline staff working in the field of intellectual disability: An evaluation of training protocols, intervention compliance, and generalisation.

Volume II

Laura Gormley

A thesis submitted to Trinity College Dublin, the University of Dublin, in partial fulfillment of the requirements for the Degree of Doctor of Philosophy (PhD) in Psychology

2018

Supervisor: Dr. Olive Healy (Trinity College Dublin)
# Table of Contents

## Volume II

<table>
<thead>
<tr>
<th>References</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>References</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Appendices</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendix A</td>
<td>101</td>
</tr>
<tr>
<td>Appendix B</td>
<td>109</td>
</tr>
<tr>
<td>Appendix C</td>
<td>120</td>
</tr>
<tr>
<td>Appendix D</td>
<td>132</td>
</tr>
<tr>
<td>Appendix E</td>
<td>140</td>
</tr>
<tr>
<td>Appendix F</td>
<td>145</td>
</tr>
<tr>
<td>Appendix G</td>
<td>149</td>
</tr>
<tr>
<td>Appendix H</td>
<td>153</td>
</tr>
<tr>
<td>Appendix I</td>
<td>159</td>
</tr>
</tbody>
</table>
References


References


References


References


References

disorder: A training and consultation example. *Administration and Policy in Mental Health and Mental Health Services Research, 41*(6), 808-821.


References


References


References


References


27


References


References


listening comprehension for participants with moderate intellectual disability.  


References


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


operate an iPod Touch(R) to listen to music. *Research In Developmental Disabilities*, 32(6), 2987-2992.


References


References

School Psychology Quarterly, 17(4), 341-389. doi: 10.1521/scpq.17.4.341.20872


References


References


discriminated functional communication. *Journal Of Applied Behavior
Analysis, 43*(3), 525-530.

Lerman, D. C., Kelley, M. E., Vorndran, C. M., Kuhn, S. A. C., & LaRue, R. H., Jr.
(2002). Reinforcement magnitude and responding during treatment with

evaluation of a brief, intensive teacher-training model. *Journal Of Applied
Behavior Analysis, 41*(2), 243-248.

teachers in evidence-based practices for young children with autism. *School

Liakopoulou, M., Panaretaki, I., Papadakis, V., Katsika, A., Sarafidou, J., Laskari,
Oncology. *Supportive Care in Cancer, 16*(2), 143-150.

to persons with mental retardation. *Education and Training in Mental
Retardation and Developmental Disabilities, 34*(1), 90-98.

Light, J. C., Binger, C., Ramsay, K. N., & Agate, T. L. (1999). Teaching Partner-
Focused Questions to Individuals Who Use Augmentative and Alternative
Communication To Enhance Their Communicative Competence. *Journal of
Speech, Language, and Hearing Research, 42*(1), 241-255.
References


MacArthur, J., Ballard, K. D., & Artinian, M. (1986). Teaching independent eating to a developmentally handicapped child showing chronic food refusal and
References


References


References


References


References


References


References


References


Robinson, S. E. (2011). Teaching paraprofessionals of students with autism to implement pivotal response treatment in inclusive school settings using a
brief video feedback training package. *Focus on Autism and Other Developmental Disabilities, 26*(2), 105-118.


References


References


References


Smith, B. J., & Fox, L. (2003). Systems of service delivery: A synthesis of evidence relevant to young children at risk of or who have challenging behavior [Executive summary].


constant time delay. *Journal of Developmental and Physical Disabilities, 11*(2), 139-158.


References

University Medical Sciences, 9(2), 110-114. doi:


References


References


References

