Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognize the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library Thesis Consent Form and Deposit Licence.
Improving the effectiveness of evidence-based practices in vocational rehabilitation for people with severe mental illnesses

Helen Amy Boxall née Lockett

Abstract

Background: Advances in the science and practice of vocational rehabilitation have defined and validated evidence-based practices for people with severe mental illnesses, collectively known as Individual Placement and Support (IPS). However, there is wide variation in IPS program performance both within, and across international contexts. Research to improve performance has focused on increasing adherence (fidelity) to known evidence-based practices. An alternative strategy could involve increasing the quality of implementation more broadly. This alternative strategy is reflected in the four aims of this investigation.

1. To examine the association between program fidelity and program performance.
2. To identify and prioritise other factors that could further explain performance.
3. To investigate how a country’s socio-economic context influences program implementation.
4. To conceptualise, construct, and operationalise a measure of general program implementation quality that could, if validated, further explain and predict performance.

Methods: 1. A systematic review and meta-analysis of the empirical evidence on vocational rehabilitation program implementation internationally. 2. A review of the conceptual and empirical literature related to the science of psychosocial program implementation and performance. 3. An analysis of the socio-economic policy context of New Zealand in relation to implementing IPS practices. 4. A conceptual analysis of eight new dimensions of program implementation quality.

Results: 1. Program fidelity is an important predictor of performance. Yet, in the example program studied, fidelity explains only 16.8% of the total variance in program performance. 2. A new Implementation Framework for Vocational Rehabilitation (IFVR) identifies and prioritises candidates for improving the performance of vocational rehabilitation programs. 3. A country’s socio-economic policy context influences the availability and implementation of evidence-based practices. 4. The partial operationalisation of eight dimensions of implementation quality supports further development and testing.
**Conclusions:** Fidelity to evidence-based practices is important for program performance but this strategy alone is not sufficient. Broadening the conceptual framework for implementing vocational rehabilitation, beyond a focus on program fidelity, is an alternative strategy that has identified other promising ways to improve program performance. If measures based on the IFVR are found to be sufficiently valid and reliable, the IFVR promises to assist program leaders, administrators, and policymakers to better describe, define, predict, and improve the performance of vocational rehabilitation programs.
Use of Language

I embarked on this research to generate knowledge which could address the labour force disadvantage of people diagnosed with severe mental illnesses. The term ‘mental illnesses’ is used to refer to clinically diagnosed mental disorders such as schizophrenia, schizoaffective disorders, anxiety disorders, and bi-polar affective disorder, as defined by the DSM-5 or ICD-10 classification systems (American Psychiatric Association, 2013; World Health Organisation [WHO], 1992). The terms ‘severe’ or ‘severe and persistent’ mental illnesses describe symptoms which meet the criteria in the classification systems for a mental disorder and there is a significant loss of functioning, or a lasting impact on a person’s ability to perform activities in daily living. This is approximately 3% of the population, of whom less than 20% are in competitive employment (Morgan et al., 2011).

I use the terms ‘severe mental illnesses’ or ‘psychiatric disability’ in this thesis to distinguish people whose mental health is having a significant and lasting impact on their daily living and vocational participation, from those whose mental health is having less of an impact, or the nature of the issues are transitory. This distinction is important to ensure priority access to evidence-based vocational rehabilitation. Without such support, many people do not fulfil their vocational aspirations. However I also recognise that some people feel uncomfortable with the use of the term ‘severe mental illnesses’, considering the terminology has previously contributed to stigma and discrimination, and takes a negative and deficits framing of their experiences. The recovery movement has encouraged alternative descriptions, such as ‘psychiatric survivor’, ‘person diagnosed with a mental health condition’, or a ‘person with lived experience’ (Cunningham, Peterson, & Collings, 2017, p. 284).

In using these diagnostic terms in this research, I want to ensure that a common language is used so that the nature of the underlying health conditions can be clearly understood and classified with respect to widely used international classification systems of mental illness and psychiatric disability (e.g. DSM-5 and ICD-10). The use of these diagnostic terms or descriptions are not intended to limit a person’s identity. Most psychiatric diagnostic labels represent syndromes (a checklist of signs and symptoms of which only a subset are necessary to meet diagnostic criteria). This means that there can
be wide variation in individual diagnostic profiles even when those individuals are classified by exactly
the same diagnostic category, subtype, and severity of a particular condition. The language used in this
thesis follows the guidelines given in psychiatric treatment and rehabilitation (American Psychological
Association, 2010). However, I want to emphasise from the outset that these terms do not define
peoples’ identities. People diagnosed with severe mental illnesses, or who are referred to as having
psychiatric disabilities, are always people first.
Acknowledgements

I would like to acknowledge the many people who have made the development and submission of this body of research possible. Firstly, I want to recognise and acknowledge the wisdom, kindness, research expertise, and unrelenting responsiveness of my two supervisors, Geoff Waghorn and Rob Kydd. Geoff, I am indebted to the incredible depth and breadth of your knowledge, your unwavering attention to scientific rigour, and for sharing the pleasure you also experienced as the discoveries in this thesis unfolded. Through these five years you have been alongside me, encouraging me to stretch my thinking, to take each idea and explore it even further. I come out of this with an attention to detail and understanding of scientific methods that will stay with me as I continue to progress my academic career. Rob, I am also indebted to you. We had only met a couple of times when I called you out of the blue to ask you to be my supervisor, and you did not hesitate to agree. I am so pleased I asked you. I have always felt in such safe hands, with your experience and success with other students, and the way you complemented Geoff’s perspectives, providing always the bigger picture, challenging any assumptions that I was making, and providing ongoing reassurance and encouragement. One of my fears in starting a thesis was how to deal with situations when my supervisors disagreed, and thankfully I never had to!

I want to acknowledge the financial support provided by my employer, the Wise Group, to undertake this thesis. Thank you particularly to Julie Nelson and Jacqui Graham, who encouraged me to embark on this research journey and provided ongoing support throughout. Thanks also go to Warren, Becki, Caro, Andrea, Claire, Ruth, and other work colleagues and friends who provided a sounding board for my ideas, helping to keep a practical, useful, and applicable focus to the research. I would like to thank David Chant who provided the statistical support for the meta-analysis in Chapter 3 and his wise insights as I interpreted the findings. I also acknowledge Lynne, Jan, Adam and others from the UK IPS Centres of Excellence program who organised for me to give a lecture on my research, while I was in the UK in September 2018. This was a privilege, as you gave me the opportunity to share what I have learned with leading practitioners in the field, to see their enthusiasm
for what had been discovered, and to hear how they would use the findings to improve the implementation of IPS in their regions of the country.

Writing a thesis would not have been possible without the love, support and kindness of my family and friends. Firstly, I want to acknowledge my husband, Gordon, for always believing in me even when I did not believe in myself, for your ongoing patience, and for your genuine interest and excitement in my discoveries and learnings. I am so lucky to have been able to share all these moments with you. Thanks to Kate who came over to New Zealand to have a holiday with me, and spent many evenings of our trip discussing the operationalisation of concepts! Thanks also to my mum, dad, and brother, my stepchildren Katie, Andy, and Mark, to Suzy, Jilly, Deebs, Marion, Viv, Phillipa, Nic, Joanna, Andrew, Gail, and all the many family, friends, and neighbours who have taken an interest in my research and supported my wellbeing along the way. Particular thanks to my fitness buddies, Emma, Mafi, and Kaz. I knew from the start that completing a thesis meant I needed endurance as well as the mindset to persevere when the going got tough, and the three of you made sure we all kept fit and healthy along the journey too. Thanks particularly for listening to the ups and downs of this journey as we ran, walked, or cycled together.

Lastly, I want to give a special acknowledgement to my two mokopuna (grandchildren), Jago and Ruan. Jago was only a year old when I started and Ruan was born during the process. You have brought me a joy I did not know was possible, and a zest for fun and laughter than can get forgotten in the depths of undertaking research. You also contribute to my energy and commitment to make our society a more just and fair one. It is not acceptable that we have evidence-based practices that could support people who experience mental illnesses to secure their vocational aspirations, but we fail to make these practices available for everyone who needs them.
Publishers’ approvals and third-party copyright agreements


   Published by Emerald Publishing Limited, Howard House, Wagon Lane, Bingley BD16 1WA, United Kingdom. Permission to reuse obtained August, 24, 2017. Appears on pages 27 to 54.


   Published by IOS Press, Nieuwe Hemweg 6B, 1013 BG Amsterdam, The Netherlands. Permission to reuse obtained June, 8, 2018. Appears on pages 65 to 85.


   Published by IOS Press, Nieuwe Hemweg 6B, 1013 BG Amsterdam, The Netherlands. Permission to reuse obtained August, 21, 2018. Appears on pages 96 to 118.
# Table of Contents

Abstract .................................................................................................................................................. iii

Use of Language ........................................................................................................................................ v

Acknowledgements ...................................................................................................................................... vii

Co-Authorship Forms ............................................................................................................................... ix

Publishers’ approvals and third-party copyright agreements .................................................................. xii

Table of Contents ..................................................................................................................................... xiii

List of Tables ........................................................................................................................................... xxi

List of Figures .......................................................................................................................................... xxiii

List of Abbreviations ............................................................................................................................... xxiv

Glossary .................................................................................................................................................... xxv

Chapter 1: Overview ................................................................................................................................. 1

  Introduction ........................................................................................................................................... 1

  Aims ....................................................................................................................................................... 4

  Thesis Outline ...................................................................................................................................... 7

Chapter 2: Vocational Rehabilitation for People with Severe Mental Illnesses ............................. 10

  Introduction .......................................................................................................................................... 10

  The evolution of Vocational Rehabilitation in Psychiatry ................................................................. 11

  The Identification of Evidence-Based Practices .................................................................................. 12

  IPS Efficacy and Effectiveness ............................................................................................................. 16

  Enhancements to IPS programs ........................................................................................................... 21

  Addressing Variation in Performance ................................................................................................. 22

  Chapter Conclusions ........................................................................................................................... 24

Chapter 3: Program Fidelity and Program Performance ................................................................. 25

  Prelude ................................................................................................................................................ 25

  xiii
Results ........................................................................................................................................70

A generic conceptual implementation framework to enhance program performance ..........70
A conceptual framework to enhance performance of vocational rehabilitation programs ..70

Discussion ..................................................................................................................................80

New dimensions of implementation quality in vocational rehabilitation ......................80
Employment specialist expertise .................................................................................................81
Program intensity ..........................................................................................................................82
The quality of program delivery .................................................................................................82
The quality of mental health treatment .......................................................................................83
Other dimensions of quality in the new framework .................................................................83
Quality assurance and control .....................................................................................................84
Limitations ....................................................................................................................................85

Conclusions ................................................................................................................................85

Supplementary Analysis .............................................................................................................86

The papers included in phase one of the literature review .......................................................86
Expanding one of the dimensions of implementation quality ...............................................86
Naming the new conceptual implementation framework .........................................................87

Chapter Conclusions ................................................................................................................87

Chapter 5: Applying the Implementation Framework for Vocational Rehabilitation .... 94

Prelude .........................................................................................................................................94

Introduction ..................................................................................................................................96

Methods .......................................................................................................................................99

Data search ..................................................................................................................................99

Data analysis .................................................................................................................................100

Data synthesis ..............................................................................................................................101
Chapter 6: Operationalising the Implementation Framework for Vocational Rehabilitation.

Introduction...............................................................................................................................................120

Methods ................................................................................................................................................122

Step 1: Developing the construct specification ....................................................................................123
Step 2: Identifying the measurable attributes of each construct .........................................................124
Step 3: Identifying existing relevant measures ....................................................................................125
Step 4: Refining the construct specifications and measurable attributes ............................................125
Step 5: Generating targeted statements for each attribute .................................................................127
Step 6: Scale rationalisation................................................................................................................127
Step 7: The development of a measurement item pool .......................................................................127

Results..................................................................................................................................................128
Chapter 7: Operationalising Employment Specialist Expertise

Introduction
Methods
Identifying relevant literature
The Employment Specialist’s Role in Program Performance
Developing the Construct Specification
The attributes of employment specialist expertise
The Practices of Higher Performing Employment Specialists
Establishes a Working Alliance with Caseload Participants
Secures Jobs Aligned with the Desires and Preferences of Participants
Develops Participants’ Job Search Skills
Provides Employment Assistance in the Workplace
Develops a Network of Employers
Builds Effective Working Relationships with a Range of Stakeholders
Puts in Extra Effort to Address Participants’ Assistance Needs and Barriers
Effectively Manages a Caseload of Participants
Other Candidate Attributes for Targeted Measures
Applies relevant evidence-based practices
Engages in self-reflection and self-evaluation of role performance
Discussion
Measurement
Chapter 8: Operationalising Participant Responsiveness ........................................... 242

Introduction ............................................................................................................. 242

Chapter overview ............................................................................................... 243

Methods ................................................................................................................ 245

Developing the Construct Specification ............................................................... 246

Participant responsiveness and program performance ...................................... 246

Program attrition and program performance .................................................... 247

Entry criteria applied by IPS trials ................................................................. 249

The Need for Conceptual Clarity ...................................................................... 250

Participant responsiveness and participant motivation .................................. 250

Re-conceptualising participant responsiveness in vocational rehabilitation programs .... 252

Participants’ Employment Motivation ............................................................... 254

Motivation as a psychological state ................................................................. 256

Motivation as a fixed individual trait ............................................................... 257

Measuring employment motivation ................................................................ 258

The Influences on Participant Responsiveness ............................................... 265

Participants’ general interest in employment ................................................... 265

Participants’ perceptions ................................................................................. 267

Other participant characteristics .................................................................... 271

Participants’ circumstances ............................................................................ 275

Relevant program characteristics that influence participant engagement .... 277

Measuring Participant Responsiveness ............................................................ 280
Chapter 9: Discussion

Introduction

Key Findings

The association between program performance and program fidelity

A framework to search for unexplained variance in program performance

The wider socio-economic context and program implementation

Operationalising the new implementation framework

Implications and Applications

The utility of the new assessment tool

Generalisability

Limitations and Areas for Future Research

Limitations

Areas for further research

Conclusions

References
List of Tables

Table 1. Overview of key publications in the evolution of IPS efficacy and effectiveness research .....19
Table 2. Study inclusion criteria and rationale for each ..........................................................35
Table 3. Studies included in the systematic review and meta-analysis ................................37
Table 4. Precision analysis of the Individual Placement and Support 15-item fidelity scale ........47
Table 5. Precision analysis of the Individual Placement and Support 25-item fidelity scale ........47
Table 6. Multivariate analysis of correlates of commencing competitive employment ...............48
Table 7. A comparison of the hi-hi and the hi-low cohort ....................................................56
Table 8. Multivariate analysis of the correlates of commencing employment within the high-fidelity cohorts ........................................................................................................57
Table 9. Qualitative analysis of the hi-low cohorts ..................................................................60
Table 10. The common dimensions in a generic implementation framework mapped to vocational rehabilitation ........................................................................................................72
Table 11. The empirical studies used to develop the conceptual framework for evidence-based practices in vocational rehabilitation ...............................................................................73
Table 12. The included studies in the first phase of the literature review to develop the generic implementation framework ........................................................................................................88
Table 13. Five distinguishing features of Individual Placement and Support .........................97
Table 14. The labour force activity of working-age population groups in New Zealand compared to international benchmarks ........................................................................................................104
Table 15. Relevant features of the New Zealand social and economic context, 2015-2016 ........107
Table 16. Reviewing contracts offered by two government purchasers of vocational rehabilitation for people with psychiatric disabilities ........................................................................113
Table 17. The challenges, implications, and possible solutions for expanding the implementation of evidence-based vocational rehabilitation .................................................................114
Table 18. The terms used in the partial operationalisation of the IFVR .....................................123
Table 19. Operationalising the construct of technical assistance ......................................................... 139
Table 20. Operationalising the construct of program evaluation .......................................................... 149
Table 21. Operationalising the construct of removal of non-evidence-based practices ....................... 159
Table 22. Operationalising the construct of complementary programs................................................ 171
Table 23. Operationalising the construct of program intensity .............................................................. 181
Table 24. Operationalising the construct of program delivery ............................................................. 193
Table 25. Empirical studies examining the attributes of higher performing employment specialists... 209
Table 26. Operationalising the construct of employment specialist expertise ...................................... 237
Table 27. The different approaches used in vocational rehabilitation to the conceptualisation and measurement of employment motivation ................................................................. 260
Table 28. Operationalising participant responsiveness ........................................................................ 282
List of Figures

Figure 1. Broadening the conceptualisation of program implementation quality ........................................244

Figure 2. Flowchart of article screening and assessment against the inclusion criteria. ............................36

Figure 3. The proportion of employment commencements by fidelity score for cohorts measuring fidelity with the IPS-15 scale ..................................................................................................................44

Figure 4. The relative odds of participants commencing employment in each cohort ordered by fidelity measured with the IPS-15 scale ..................................................................................................................46

Figure 5. The performance of the high performing cohorts compared with the low performing cohorts of the same fidelity score .............................................................................................................58

Figure 6. A conceptual framework for improving the effectiveness of evidence-based practices in vocational rehabilitation ........................................................................................................................................79

Figure 7. The Implementation Framework for Vocational Rehabilitation (IFVR) ..................................93

Figure 8. The steps taken in data search, data analysis, and the synthesis of findings .........................102

Figure 9. The nine dimensions of implementation quality in the IFVR .................................................122

Figure 10. The steps employed in the conceptual analysis for each dimension of implementation quality in the IFVR ........................................................................................................................................126

Figure 11. The observability of the candidate attributes of employment specialist expertise ..............207

Figure 12. Candidate knowledge, skills, and attitudes underlying the behavioural indicators of employment specialist expertise ....................................................................................................................................236

Figure 13. Re-conceptualising the construct participant responsiveness .................................................253

Figure 14. The influences on participant responsiveness in vocational rehabilitation programs ........288
# List of Abbreviations

An explanation of the abbreviations in this investigation.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>EBP</td>
<td>Evidence-Based Practices</td>
</tr>
<tr>
<td>ES</td>
<td>Employment Specialist</td>
</tr>
<tr>
<td>IFVR</td>
<td>Implementation Framework for Vocational Rehabilitation</td>
</tr>
<tr>
<td>IPS</td>
<td>Individual Placement and Support</td>
</tr>
<tr>
<td>IPS-15</td>
<td>The 15-item Individual Placement and Support fidelity scale</td>
</tr>
<tr>
<td>IPS-25</td>
<td>The 25-item Individual Placement and Support fidelity scale</td>
</tr>
<tr>
<td>MBIE</td>
<td>Ministry of Business, Innovation and Employment, New Zealand</td>
</tr>
<tr>
<td>MoH</td>
<td>Ministry of Health, New Zealand</td>
</tr>
<tr>
<td>MSD</td>
<td>Ministry of Social Development, New Zealand</td>
</tr>
<tr>
<td>NICE</td>
<td>National Institute for Health and Care Excellence</td>
</tr>
<tr>
<td>NZ</td>
<td>New Zealand</td>
</tr>
<tr>
<td>NZDSN</td>
<td>New Zealand Disability Support Network</td>
</tr>
<tr>
<td>OECD</td>
<td>Organisation for Economic Co-operation and Development</td>
</tr>
<tr>
<td>RANZCP</td>
<td>Royal Australian and New Zealand College of Psychiatrists</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomised Controlled Trial</td>
</tr>
<tr>
<td>RQ</td>
<td>Research Question</td>
</tr>
<tr>
<td>SE</td>
<td>Supported Employment</td>
</tr>
<tr>
<td>Statistics NZ</td>
<td>Statistics New Zealand</td>
</tr>
<tr>
<td>TVR</td>
<td>Traditional Vocational Rehabilitation</td>
</tr>
<tr>
<td>UK</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>USA</td>
<td>United States of America</td>
</tr>
<tr>
<td>VR</td>
<td>Vocational Rehabilitation</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
</tbody>
</table>
Glossary

An explanation of technical terms and phrases and how they are applied in this investigation. Citations relevant to these definitions can be found in the main text.

**Competitive employment** refers to part-time or full-time employment positions in the open labour market. The two main criteria are that competitive employment positions pay minimum wage or above, and the job is not reserved for people with disabilities or severe mental illnesses. Self-employment can be included if earnings are equivalent to minimum wages. Employment in a family business or farm would normally be excluded as these positions are usually not open to competition. Jobs in Social Firms and Social Enterprises would normally be excluded unless it can be shown that the particular jobs involved are open to competition in the wider labour market and pay at least minimum wages.

**Concepts** are representational ideas concerning entities, classes of entities, and their characteristics. The main concept under investigation in this thesis is the concept of implementation quality. The concept of implementation quality is assumed to be multi-dimensional made up of subordinate concepts. At the next subordinate level these are described as ‘**dimensions of implementation quality**’.

**Constructs** are more specific and represent measurable stand-alone elements of a concept or of a dimension. A **construct specification** is used in this thesis to establish the working definition and description of each construct under investigation to enable precise measurement of the element, entity or object. A construct specification describes elements considered to make up the construct as well as those that do not.

The term **dimension** is used in this thesis to distinguish between the candidate constructs hypothesised to directly influence vocational rehabilitation program performance (e.g. dimensions of implementation quality) from those candidate constructs likely to have an indirect influence (e.g. contextual factors).

The focus of this investigation is on improving the **effectiveness** of psychosocial programs. In this context effectiveness is concerned with the outcomes which can be achieved when a program is delivered in a real-world context under the normal conditions encountered in routine service delivery. This contrasts with the conditions in experimental studies where many elements of the environment may be controlled or restricted to maximise the capacity to detect real differences between two interventions, or between an intervention and a control.

**Efficacy** represents the maximum effect size found under controlled conditions. Efficacy is usually measured through comparison to an alternative intervention, or to a control such as services as usual, prior to the introduction of the new intervention. In vocational rehabilitation, randomised controlled trials conducted across different countries have established the efficacy of the Individual Placement and Support (IPS) in each country. In these trials alternative approaches, or no specific approach to vocational rehabilitation provided the comparison effect.

**Efficiency** in this thesis is considered the ratio of desired outputs to inputs. Because almost all the inputs can be assigned a monetary value, this form of efficiency can be readily measured via methods such as cost-benefit analysis or return on investment. A more efficient program would be the one that demonstrates the lowest cost per unit of quality employment outcome. Efficiency can also be measured as a time sensitive variable. For instance a program could be considered more efficient if the target units of employment outcome were achieved at the same cost, but in a shorter timeframe.
In this case the reason for the greater efficiency could be improved timeliness. Alternatively timeliness could be used to assess efficiency independently of costs.

**Enhancements** are new practices not conceptualised in fidelity scales that are added to a program in order to improve program performance. Enhancements are practices that are not specified in the original model, as represented by the fidelity scale or as described in the practice manual. For example Individual Placement and Support enhancements have included cognitive remediation training, social skills training, and motivational interviewing.

**Evidence-based practices** are a set of well-defined principles and practices which have been validated by research, internationally and shown to be effective, safe, and cost-effective. The current evidence-based practices in vocational rehabilitation are known as Individual Placement and Support. There are standards for what constitutes evidence. For the purpose of this investigation this standard is considered the highest reasonable standard of randomised controlled trial (or equivalent experimental design) possible in the field of interest.

**Fidelity** is the term used to describe the adherence of program delivery to the original principles and practices specified for a particular approach. Fidelity manuals describe how specific practices can be identified and measured. Practice manuals provide broader information about how and why particular practices are recommended. In this thesis the terms adherence, fidelity, and implementation fidelity relate to the measurement of fidelity.

A **fidelity scale** measures the quality of program delivery in relation to how well it represents (aligns with) and reproduces the original model principles and practices. Fidelity scales need to have good psychometric properties, particularly reliability, construct validity, and predictive validity. There are two scales to measure fidelity to evidence-based practices in vocational rehabilitation. The 15-item IPS scale and the 25-item IPS scale.

**Individual Placement and Support** is the term assigned to the current set of defined evidence-based practices in vocational rehabilitation for people with severe mental illnesses.

**Operationalisation** is the process of developing a working definition for a concept to create a defined construct which can be measured. Conceptual analysis is used in this thesis to develop operational definitions so that the underlying concepts and constructs can be described, defined, identified, and measured.

**Supported employment** is an approach to vocational rehabilitation which adopts practices which assist participants to commence competitive employment in the open labour market relatively quickly. Supported employment then provides tailored support once the participant commences employment. The concept of supported employment emerged from research in the 1980s responding to people with intellectual disabilities who wanted competitive employment outside the sheltered employment and other institutions in which they were primarily being assisted.

**Traditional vocational rehabilitation** was the dominant approach to vocational rehabilitation prior to the emergence of supported employment programs. Traditional vocational rehabilitation programs typically conduct extensive assessment followed by pre-employment training in non-competitive employment or protected environments, prior to offering support to commence competitive employment.
Chapter 1: Overview

Introduction

People with mental health disorders have lower rates of labour force participation than people without mental health disorders (Jonsdottir & Waghorn, 2015). For people with a diagnosis of depression and anxiety disorders labour force activity is estimated to be 45%. For people diagnosed with schizophrenia spectrum disorders, labour force activity is estimated to be lower, at 20% (Drake, Bond, Goldman, Hogan, & Karakus, 2016; Jonsdottir & Waghorn, 2015; Morgan et al., 2011, 2016).

These lower levels of labour force activity are not a reflection of low employment aspirations. Research shows that most people with mental health disorders, including people with diagnoses within the schizophrenia spectrum, would like to be employed, and in fact see getting a first job, or returning to employment, as an integral part of recovery (Drake et al., 2016; Morgan et al., 2016; Peterson, Gordon, & Neale, 2017). Westcott, Waghorn, McLean, Statham, and Mowry (2015) found 80-90% of people with a diagnosis of schizophrenia and schizoaffective disorder had an interest in employment. Similarly, Seebohm and Secker (2005) recruited people with severe mental illnesses as interviewers to conduct a survey with 149 unemployed adults currently receiving treatment from mental health services in England. Ninety percent of survey participants were interested in obtaining employment (p. 12).

Therefore there is a large gap between employment aspirations and actual employment status. Applied research can help optimise employment outcomes for people with severe mental illnesses (Hogan, Drake, & Goldman, 2014). Employment has the potential to improve lives, promote social inclusion and financial independence, ease the impact on families, and reduce the costs to government and to the wider community (Arends, Baer, Miranda, Prinz, & Singh, 2014; Organisation for Economic Co-operation and Development [OECD], 2015a; Peterson et al., 2017).

Vocational rehabilitation is the science and practice of doing whatever is needed to help someone with a health condition or injury, to return to, or stay in employment, “It is an idea and an approach as much as an intervention or a service” (Waddell, Burton, & Kendall, 2008, p. 5). In psychiatry, the science and practice of vocational rehabilitation has advanced considerably in the past.
30 years (Drake & Bond, 2017). There is now a defined set of evidence-based practices shown to be more effective than other approaches to vocational rehabilitation at supporting people diagnosed with severe mental illnesses to gain and remain in competitive employment (Drake & Bond, 2017; Kinoshita et al., 2013; Marshall et al., 2014; Modini et al., 2016). Severe mental illnesses are defined as having a diagnosis in one of the following categories: schizophrenia spectrum disorders, bipolar affective disorders, major depressive disorder, or other psychotic disorders, and the disorder is having a moderate or severe impact on the person’s functioning and their participation in daily living, i.e. on mood and feelings, thinking, interpersonal interactions, role performance, or self-care (American Psychiatric Association, 2013; Morgan et al., 2011). Competitive employment is defined as a job that pays at least the minimum wage, the employment setting has disabled and non-disabled workers, the job is not reserved specifically for a person with a disability, or mental illness, and the person is employed by the business and supervised by an employee in that business (Becker & Drake, 2003, p. 77).

The current set of evidence-based practices in vocational rehabilitation for people with severe mental illnesses is also known as the Individual Placement and Support (IPS) approach to vocational rehabilitation. The IPS approach is intended to complement evidence-based psychiatric treatment and care in the community, not in long-stay hospitals or in other institutions. Twenty-five randomised controlled trials (RCTs), across 4 continents and 15 countries, and a growing number of systematic reviews and meta-analyses, including three Cochrane reviews, have consistently endorsed the effectiveness of IPS over other forms of vocational rehabilitation (Bond, Drake, & Becker, 2008a, 2012b; Crowther, Marshall, Bond, & Huxley, 2001; Drake et al., 2016; Drake & Bond, 2017; Kinoshita et al., 2013; Metcalfe, Drake, & Bond, 2017a; Reme et al., 2018; Suijkerbuijk et al., 2017). A recent systematic review and meta-analysis of 17 of the RCTs, across 10 countries, compared IPS to alternative types of vocational rehabilitation. The review found IPS increased the odds of participants commencing employment by more than two times (pooled risk ratio 2.40, 95% CI 1.99-2.90) (Modini et al., 2016).
Despite on average performing higher, there is considerable variation in the proportion of participants commencing competitive employment across IPS interventions. For example, in a systematic review of 15 RCTs, the study with the highest outcomes supported 78.1% of program participants into competitive employment, and the study with the lowest outcomes supported 22.1% of program participants (Bond et al., 2012b). The Bond et al. (2012b) systematic review also found a significant difference among the performances of the nine USA IPS interventions compared to the six non-USA IPS interventions. The proportion employed in the active intervention (IPS) reached 62.1% in USA studies compared to 47.3% in non-USA studies (Bond et al., 2012b). Despite the proven effectiveness of IPS in the USA, performance at some USA locations can be suboptimal (below the USA mean), and there remains considerable variability in performance across all reported IPS interventions.

One way to improve the average performance of IPS programs would be to enhance the performance of all the low performing programs. This would reduce the variability in outcomes attained across programs. Program performance can be improved through working on:

(a) implementation fidelity, i.e. adherence to the original IPS program principles and practices;
(b) program quality and program delivery more generally (Durlak & DuPre, 2008); or
(c) additions (enhancements) to the program (Loveland, Driscoll, & Boyle, 2007).

Measuring current performance of programs in relation to IPS implementation fidelity is an important starting point to identify the extent of variation in performance explained by program fidelity. The next step is to understand the extent of variation in program performance and to determine the extent that local fidelity scores predict local program performance (Mowbray, Holter, Gregory, Teague, & Bybee, 2003; O’Donnell, 2008). A subsequent step is to look beyond program fidelity to identify whether there are any other factors that could distinguish higher performing from lower performing programs. This step is important because we need to effectively apply the practices we already know are effective first, before we start adding new components onto evidence-based practices (Sederer, 2009). An effective program which is poorly implemented or delivered in suboptimal
conditions, will not produce the anticipated outcomes (Durlak, 2016; Fixsen, Blasé, Metz, & Van Dyke, 2013; Latimer, 2010). Once factors for optimal program delivery, other than fidelity, are identified, described and implemented, consistent program delivery and more reliable outcomes can be expected (Fixsen et al., 2013).

The science and implementation of IPS has advanced enough to have developed and validated a measure of implementation fidelity. Being able to measure adherence to the original program design enables fidelity to be taken into account when measuring program implementation, program performance, and program effectiveness (Mowbray et al., 2003). A measure of program implementation fidelity also means that differences in program performance can be compared and contrasted in relation to fidelity, as well as by identifying ways in which programs can be developed further. Thus, examining the relationship between fidelity to IPS principles and practices and program performance is a logical starting place for this investigation.

Aims

This thesis aims to explore how to improve the effectiveness of evidence-based practices in vocational rehabilitation for people with severe mental illnesses. To achieve this a set of four research questions (RQ) were generated, the findings from the previous RQ informing the construction of the next.

RQ1. What is the nature of the relationship between implementation fidelity of IPS vocational rehabilitation programs and effectiveness in an international context?

RQ2. Should a wider theoretical implementation framework be applied to improve vocational rehabilitation for people with severe mental illnesses?

RQ3. How does the socio-economic policy context influence the implementation of evidence-based practices in vocational rehabilitation?
RQ4. Can the new implementation framework for vocational rehabilitation, the IFVR, be operationalised to construct a general tool for measuring the quality of program implementation and program delivery, to complement existing measures of program fidelity?

**RQ1. What is the nature of the relationship between implementation fidelity of vocational rehabilitation programs and effectiveness in an international context?**

Two hypotheses were constructed to investigate the first research question. 1. Vocational rehabilitation programs which adhere more closely to the principles and practices of the IPS approach to vocational rehabilitation will have better outcomes. 2. The strength of the relationship between fidelity of implementation and employment outcomes will be affected by other influences. The research question is investigated by conducting a systematic review and meta-analysis of relevant empirical evidence to examine the association between implementation fidelity and program outcomes in vocational rehabilitation programs internationally.

**RQ2. Should a wider theoretical implementation framework be applied to improve vocational rehabilitation for people with severe mental illnesses?**

This second research question arose from the findings from RQ1. Whilst program fidelity is important, it does not explain all the variance in program performance. A conceptual implementation framework could assist with the prioritisation and organisation of candidate constructs influencing program performance. To investigate this second RQ, a review is undertaken of the broader conceptual and empirical literature related to the science of psychosocial program implementation. This review aims to identify the most relevant dimensions of a general concept of program implementation quality applicable to vocational rehabilitation.

The review also aims to identify any contextual factors that could further explain variation in program performance. These dimensions of implementation quality and contextual factors are selected in terms of their potential applicability to vocational rehabilitation and to their strength of supporting evidence. Eight dimensions of implementation quality in addition to implementation fidelity and 23
contextual factors are organised into a new Implementation Framework for Vocational Rehabilitation (IFVR). The IFVR is intended for improving the performance of vocational rehabilitation programs.

**RQ3. How does the socio-economic policy context influence the implementation of evidence-based practices in vocational rehabilitation?**

The application of the IFVR needs to be country specific because each country represents a unique socio-economic and political context. To answer the third research question, a systematic examination of the economic and social policy context in New Zealand is conducted to identify whether, and how, the wider policy environment is affecting the implementation of evidence-based practices in vocational rehabilitation for people with severe mental illnesses. This investigation also proposes a method for conducting a policy analysis in relation to the implementation of vocational rehabilitation practices and principles. Other countries seeking to improve the implementation of IPS principles and practices could conduct a similar analysis prior to applying the IFVR.

**RQ4. Can the new Implementation Framework for Vocational Rehabilitation (IFVR) be operationalised to construct a general tool for measuring the quality of program implementation and program delivery, to complement existing measures of program fidelity?**

To use the IFVR developed from RQ2 to improve program performance, the key concepts need operationalising. To begin this process, the final part of this thesis undertakes a detailed conceptual analysis for each of the eight proposed dimensions of implementation quality in the IFVR. A clear and well-defined construct specification is developed which includes the measurable attributes thought to best represent the conceptual space for each construct. Targeted statements enable the development of an initial item pool for each measurable attribute, and ways to measure each item are suggested. Once empirically tested for utility, reliability, and validity, it is expected that the resulting measures will help further explain and predict the performance of vocational rehabilitation programs for people with severe mental illnesses. The new measures are expected to be most applicable to programs that have achieved the threshold of good fidelity to evidence-based practices, yet their program outcomes remain
suboptimal. However, the measures could also help improve performance for programs still working on improving implementation fidelity.

**Thesis Outline**

This thesis begins in Chapter 2, with an overview of the evolution of vocational rehabilitation for people with severe mental illnesses. This includes the development of different approaches to vocational rehabilitation and how they have contributed to the emergence of a set of well-defined, evidence-based practices. The Individual Placement and Support (IPS) approach is the current set of evidence-based practices in vocational rehabilitation for people with severe mental illnesses. Chapter 2 therefore discusses the development of the IPS approach and outlines what is currently known about IPS efficacy and effectiveness. The literature on IPS program enhancements is also reviewed. Enhancements have been the primary mechanism utilised to improve the effectiveness of evidence-based practices in vocational rehabilitation, along with improving fidelity to evidence-based practices.

Chapter 3 explores the association between program fidelity and program performance internationally. This association is examined through a systematic review and meta-analysis of relevant empirical evidence utilising two measures of IPS implementation fidelity and a measure of program performance. Thirty studies of IPS implementation fidelity and outcomes were included across 12 countries. The chapter discusses the contribution of program fidelity to program outcomes, then explores the need for the identification of additional influences which could explain more of the variance in program performance.

Chapter 4 draws from the wider science of psychosocial program implementation to help with the search for explanations of unexplained variance in the performance of vocational rehabilitation programs. A review of the broader conceptual and empirical literature related to the science of psychosocial program implementation is undertaken. The findings from the review identify candidates for new dimensions of program implementation quality and program delivery, as well as candidates for contributing contextual factors. These dimensions of implementation quality and contextual factors
were then organised by strength of evidence, and potential to be operationalised, to build a new Implementation Framework for Vocational Rehabilitation, referred to as the IFVR.

The operationalisation of the IFVR starts in Chapter 5 by examining the economic and social context of the country where it is to be applied. The economic and policy context of a country is an important potential influence on the implementation of vocational rehabilitation programs. The chapter gives the rationale for selecting New Zealand to apply the IFVR. The chapter provides the methods and results from a systematic examination of the economic and social policy context in New Zealand. This analysis helps to understand how the socio-economic policy context affects both the availability and the implementation of evidence-based practices in vocational rehabilitation for people with severe mental illnesses. This examination used the operationalisation of IPS implementation fidelity in New Zealand to identify conflicts with existing policy and to determine any policy adjustments that could further improve program implementation quality. The other eight dimensions of implementation quality would also benefit from similar operationalising within a specific social and economic policy context.

The following three chapters, 6, 7, and 8 report on the conceptual analysis to partially operationalise the eight new candidate dimensions of implementation quality which are hypothesised to complement implementation fidelity. Chapter 6 outlines the methods in the conceptual analysis of the eight dimensions of implementation quality in the IFVR. This chapter then reports the results from the conceptual analysis for six of the eight dimensions of implementation quality. The conceptual analysis provides the initial scoping of the conceptual space for each construct, and develop an operational construct specification. The analysis seeks to identify and describe discrete dimensions of implementation quality and program delivery, to enable the development of valid and reliable measures of these six constructs. Chapters 7 and 8 outline and discuss the results of the conceptual analysis of the two remaining dimensions of implementation quality, employment specialist expertise and participant responsiveness, respectively. Employment specialist expertise and participant responsiveness are both complex multi-component dimensions of the IFVR. Each of the two chapters makes the case for clear and explicit conceptualisation of these dimensions, while presenting evidence that previous research
has often failed to establish clear conceptual distinctions. The subsequent lack of conceptual clarity appears to have contaminated previous attempts to measure these constructs.

The conceptual analysis in Chapters 6, 7, and 8 partially operationalise the eight dimensions of implementation in the IFVR ready for further development and empirical testing. The measures are intended to complement, not replace, existing measures of program fidelity, so that the combination can reveal more opportunities to improve program performance.

Chapter 9 summarises the main findings from this body of research and compares and contrasts what has been learned with previous research. The applications and implications of what has been discovered are then discussed. Adherence (fidelity) to known evidence-based practices is an important first step, yet this chapter explores how further improvements in program performance could be attained through broadening the conceptualisation of implementation quality beyond implementation fidelity. Next the chapter examines how the new measures of implementation quality in the IFVR complement the existing fidelity scales, and the extent the findings of this thesis can be generalised. The limitations of the research are acknowledged and discussed. Future research priorities and research directions arising from this investigation that are most likely to benefit the science and practice of vocational rehabilitation are presented. The chapter concludes by highlighting the value of taking a new, and broader approach to conceptualising vocational rehabilitation implementation beyond simply attaining good implementation fidelity.
Chapter 2: Vocational Rehabilitation for People with Severe Mental Illnesses

Introduction

This chapter provides an overview of the development of vocational rehabilitation for people with severe mental illnesses. It examines different approaches to vocational rehabilitation that have been implemented and studied, and considers how these approaches have contributed to the identification of evidence-based practices in vocational rehabilitation. Advancements made in the science and practice of vocational rehabilitation for people with severe mental illnesses have resulted in a well-defined set of evidence-based practices that have been successfully described, measured, and replicated internationally. These practices are collectively referred to as the Individual Placement and Support (IPS) approach to vocational rehabilitation (Becker, Swanson, Reese, Bond, & McLeman, 2015, p. v). The efficacy and effectiveness research for the IPS approach to vocational rehabilitation is examined next.

The chapter then considers how variation in IPS program performance within and across different countries has led researchers and practitioners to investigate factors that could explain this difference in outcomes. Other attempts to improve performance are examined next, these have sought to add-on interventions, referred to as IPS enhancements (Boycott, Schneider, & McMurray, 2012; Suijkerbuijk et al., 2017). An examination of the literature on enhancements to IPS practices finds that to date no new evidence-based practices have emerged that are making a significant difference to the performance of IPS programs. The chapter concludes by raising the question: Is the current implementation framework, which defines implementation quality through program fidelity, sufficient to further improve the effectiveness of evidence-based practices in vocational rehabilitation?
The evolution of Vocational Rehabilitation in Psychiatry

The importance of work for people diagnosed with severe mental illnesses has been recognised for many years (Bond, 2004; Danley, Rogers, MacDonald-Wilson, & Anthony, 1994). However, it is only in the past 40 years that a focus on integrated employment in community settings began (Trochim, Cook, & Setze, 1994). Prior to this, work opportunities were limited to those based in institutional settings, such as sheltered employment, training programs, work enclaves, work crews, and in transitional employment schemes such as in the Clubhouse model in the United States of America (USA) (Beard, Propst, & Malamud, 1982; Wehman, Revell, & Brooke, 2003). These step-wise approaches, referred to as traditional vocational rehabilitation, were underpinned by the philosophy that people with severe mental illnesses needed time to prepare and develop work skills, and improve clinical functioning, prior to entering competitive employment settings (Drake, Bond, & Becker, 2012, p. 24; Waghorn, Lloyd, & Clune, 2009c).

In the 1980s, methods of vocational rehabilitation for people with intellectual disabilities were becoming established. These vocational rehabilitation approaches included the following practice principles: (a) competitive jobs in mainstream workplaces; (b) working alongside people without disabilities; (c) paid at the minimum wage or above; and (d) the provision of on-going support to the employee and employer (Bond, Drake, Mueser, and Becker, 1997; Wehman, 1986; Wehman, Kregel, & West, 1997, p. 11).

In contrast to the practices of traditional vocational rehabilitation or pre-vocational training which adopt a “train then place” approach, Wehman (1981) advocated for supported employment practices using a “place then train” approach. Supported employment practices focus on helping individuals to attain their own personal, competitive employment goals (Gervey, Parrish, & Bond, 1995; Wehman, West, & Kregel 1997, p. 11). Training is provided on the job rather than prior to commencing employment (Wehman, 1986). This type of individualised vocational assistance to secure and sustain competitive jobs, is what people with severe mental illnesses also preferred but services were slow to respond (Seebohm & Secker, 2005, p. 16).
Different approaches to vocational rehabilitation for people with severe mental illnesses came predominantly from the USA (Drebing et al., 2012; Schneider, 2005, p. 37). These included: Choose-Get-Keep from Boston University (Danley & Anthony, 1987); the Clubhouse programs based on Fountain House in New York (Beard et al., 1982; Macias, DeCarlo, Wang, Frey, & Barreira, 2001; Trochim et al., 1994); Indiana supported employment (Vogler, 1998); and Individual Placement and Support (Becker & Drake, 1994; Bond, Becker, Drake, & Vogler, 1997a).

Affirmative businesses or social firms also emerged as workplace environments offering a vocational path for people with severe mental illnesses. These enterprises have an entrepreneurial focus but their primary objectives are social rather than primarily profit driven (Buhariwala, Wilton, & Evans, 2015; Warner & Mandiberg, 2006). These alternative approaches are not in scope for this research because the primary focus is improving the effectiveness of evidence-based practices with a goal of obtaining individual competitive employment.

**The Identification of Evidence-Based Practices**

In the mid-1990s researchers in the USA identified and standardised the practices constituting effective vocational rehabilitation for people with severe mental illnesses (Becker & Drake, 1994; Bond, 2004; Evans & Bond, 2008; Ridgway & Rapp, 1998; Waghorn & Lloyd, 2005).

Practices from several different approaches to vocational rehabilitation contributed to the development of a common set of evidence-based practices in vocational rehabilitation for people with severe mental illnesses (Becker & Drake, 1994). The Choose-Get-Keep model contributed the importance of respecting individual preferences, job-matching, greater attention to utilising existing skills rather than skills acquisition, and the need for service coordination (Rogers, Anthony, & Farkas, 2006). Other practices were not adopted. For instance, in the Choose-Get-Keep approach a time-limit is not placed on the ‘Choose’ phase. This phase typically occurs in pre-employment counselling sessions. The Choose-Get-Keep approach also did not specify the key stakeholders to be involved in the ‘Keep’ phase (Rogers et al., 2006).
The Clubhouse psychosocial rehabilitation program developed transitional employment schemes, offering protected and assisted employment opportunities in mainstream workplaces (Beard et al., 1982; Becker & Drake, 2003, p. 11). Transitional employment provided on-going, on-site support to workers and employers (Macias et al., 2001). The focus is on real jobs for real pay (Schneider, 2005, p. 39). However, in contrast to supported employment practices transitional employment jobs belong to the Clubhouse or agency managing the Clubhouse, and individuals are rotated through them on a time-limited basis usually about six months (Beard et al., 1982).

In the USA state of New Hampshire in the early 1990s, an approach to vocational rehabilitation referred to as Individual Placement and Support (IPS) emerged (Becker & Drake, 1994). Individual Placement and Support evolved out of a combination of practices from the program of assertive community treatment (PACT) (Russert & Frey, 1991) and supported employment (Wehman, 1986), and utilised a team-based approach between mental health treatment services and vocational assistance (Drake, Becker, Clark, & Mueser, 1999a; Drake & Bond, 2017). Evaluation of the conversion of day treatment programs to supported employment programs provided early empirical evidence of effectiveness (Drake et al., 1994, 1999a). At this time, six principles were isolated to operationalise the Individual Placement and Support approach to vocational rehabilitation.

1. A primary goal of competitive employment.
2. Rapid search for jobs (within 4 weeks of commencing onto the program).
3. Attention to individual vocational preferences and interests.
4. A focus on vocational assessment while trying out competitive jobs.
5. Integration of vocational rehabilitation services with mental health treatment.
6. The provision of time-unlimited support once a job is secured (Drake et al., 1999a).

In parallel to these developments on the East coast of the USA, in the state of Indiana Bond, Dietzen, McGrew, and Miller (1995) established supported employment programs within community mental health centres. Similar to IPS, this approach to vocational rehabilitation recognised the importance of an integrated response to mental health and vocational needs, rather than brokering out
services across agencies (Vogler, 1998). This Indiana Supported Employment model emphasised: (a) open access to anyone with severe mental illness interested in employment regardless of symptoms, current functioning, or work history; (b) the coordination of vocational support with other services, in particular state vocational rehabilitation and mental health services; (c) rapid job search soon after referral to the program; (d) intensive individualised job coaching; and (e) on-going support indefinitely once a job is secured (Bond et al., 1995; Vogler, 1998). A difference from other methods of supported employment was that the Indiana approach utilised two vocational specialist roles. One dedicated staff member provided only workplace support, while other vocational generalist staff support the participant through all other stages of the vocational rehabilitation process (Vogler, 1998). The developers of the Indiana approach devised a fidelity scale, the Quality of Supported Employment Implementation Scale (QSEIS), to assist with the operationalisation of the principles, and to measure adherence to the original model (Bond, Picone, Mauer, Fishbein, & Stout, 2000; Vogler, 1998).

A 1997 review of the empirical evidence on methods of vocational rehabilitation for people with severe mental illnesses, found that supported employment programs were assisting on average, 58.0% of participants to commence employment. This compared with 21.0% of participants in traditional vocational rehabilitation programs (Bond et al., 1997b). The review identified two critical features of the effective supported employment programs. 1. Integration between vocational services and mental health treatment services. 2. Avoidance of pre-employment training. The integration of vocational services with treatment services marked an important move away from clinician-directed services. With segregated services, clinicians usually decide when to make an onward referral to vocational rehabilitation. Whereas with integrated services, the emphasis shifts to person-directed services, where vocational rehabilitation is offered to everyone within the mental health treatment service (Bond et al., 1997b). This 1997 review was a collaboration between researchers working in Indiana and those in New Hampshire, namely Bond, Drake, Mueser, and Becker. It appears to mark the time when the available empirical evidence was being brought together to articulate the features of
evidence-based practices, and the terms Individual Placement and Support were being used more widely in vocational rehabilitation research and practice (Bond et al., 1997b).

The Individual Placement and Support (IPS) approach to vocational rehabilitation for people with severe mental illnesses articulated the principles and practices that lead to better employment outcomes (Becker & Drake, 2003 p. 25; Bond, 2004; Evans & Bond, 2008; Suijkerbuijk et al., 2017). This articulation has enabled the IPS approach to be studied and implemented in new communities (Schneider, 2005, p. 43). To illustrate the growth in IPS research, in 2009 a quarter of all published empirical trials of vocational rehabilitation services were IPS interventions (Drebing et al., 2012). The IPS approach currently represents the leading evidence-based approach to vocational rehabilitation for people with severe mental illnesses (Drake & Bond, 2017; Suijkerbuijk et al., 2017). For these reasons, this thesis utilises the IPS approach as the foundation program of focal interest.

The six original underlying principles (Drake et al., 1999a) have been extended to eight core principles. Financial planning and job development were added when empirical evidence identified these as important practices (Evans & Bond, 2008). The eight principles are described as follows.

1. Program eligibility is based on individual choice.
2. The focus of the employment assistance is on commencing and maintaining competitive employment.
3. The search for jobs starts quickly, on average within four weeks of commencing on the program.
4. Mental health treatment and employment assistance services are integrated.
5. Attention is given to individual preferences, interests, and skills.
6. Support with financial planning is provided either directly, or indirectly.
7. The employment specialist actively seeks out employers and build relationships, rather than just responding to job vacancies or openings. This practice is referred to as job development.
8. Individualised supports are provided once someone commences employment. Support is provided to the employee and the employer (Becker & Drake, 2003, p. 22; Bond, 1998, 2004; Drake et al., 2012; Drake & Bond, 2014).

The core intervention within IPS involves a vocational rehabilitation specialist, referred to as an employment specialist. The employment specialist works with a capped caseload of people, usually between 20 and 25 active participants. Employment assistance can begin soon after the person receives treatment from a publicly funded mental health service. The employment specialist provides intensive, tailored, and continuous employment assistance to each program participant to assist them to secure and maintain competitive employment (Becker & Drake, 2003 p. 23) in line with the person’s employment preferences. In the USA the same agency typically employs both the mental health clinicians and the employment specialists (Cook et al., 2005). In New Zealand and Australia, the employment specialist is usually employed by a non-government organisation working in partnership with a public mental health service (Browne, Wright, Waghorn, & Stephenson, 2009; King et al., 2006). In the United Kingdom (UK) both types of staffing arrangements are found (Shepherd, Lockett, Bacon, & Grove, 2012). Whichever staffing method is in place the intention is to optimise the integration of employment assistance with mental health treatment and care services (Cook et al., 2005; Lockett & Bensemann, 2013; Shepherd et al., 2012).

**IPS Efficacy and Effectiveness**

A program of research beginning in the early 1990s, led by the Westat Center (formerly the Dartmouth Psychiatric Research Center) helped establish the efficacy and effectiveness of the IPS approach to vocational rehabilitation (Drake et al., 2012, p. 28). The distinction between efficacy and effectiveness is often artificial, depending on the extent that trials have restricted the implementation conditions (Glasgow, Lichtenstein, & Marcus, 2003). A commonly reported distinction is that efficacy is about the maximum effect size found under controlled, or optimum conditions, whereas effectiveness is more concerned with outcomes achieved in real-world contexts (Glasgow et al., 2003). The terms
efficacy and effectiveness have been used interchangeably in the IPS literature when referring to the studies involving randomised controlled trials (RCTs) (Drake et al., 2012, pp. 16, 46-64). Hence it remains unclear which of these are best described as efficacy studies or as effectiveness trials.

A quasi-experimental study in 1994 to establish proof of concept, trialled a comparison method and estimated the potential effect size (Drake et al., 1994). Within two years the results of the first RCT were published (Drake, McHugo, Becker, Anthony, & Clark, 1996). Replication RCTs followed, and by 2007 11 RCTs were reported. Seven in the USA, one in each of Canada, Hong Kong, and Australia, and a six-site trial in Europe (Bond et al., 2008a; Mueser, Drake, & Bond, 2016). By 2017, 24 RCTs were published in 14 countries. These were complemented by many other implementation studies (Drake et al., 2016; Drake & Bond, 2017). RCTs typically compared programs following IPS principles to traditional vocational rehabilitation services, skills training, or usual care (Drake et al., 2016). These RCTs provided consistent evidence of the effectiveness and efficacy, while trials outside of the USA demonstrated IPS generalisability to countries with similar developed market economies (see Table 1).

During this period, a 15-item scale for measuring adherence to the IPS principles was constructed and tested (Bond et al., 1997a). Evidence for the different practices making up the IPS approach were also examined (Bond, 1998, 2004) and an implementation resource kit was developed (Becker & Bond, 2004). These developments helped to describe and standardise the IPS approach and enabled researchers and practitioners outside the USA to replicate the practices, monitor adherence to program principles, and design and conduct further RCTs (Burns et al., 2007; Latimer et al., 2006).

In 2008, Bond, Drake, and Becker (2008) published a systematic review of IPS RCTs. This review brought together the findings from 11 RCTs, seven in the USA, one in Hong Kong, Australia, and Canada, and a multi-site trial in Europe. The review found the mean proportion of participants commencing competitive employment in the IPS interventions was 61.0%, compared to 23.0% in the controls. The review also found that two-thirds of participants who commenced employment were employed for 20 or more hours per week. Participants in the IPS intervention also had a shorter time to
commencing employment than those in the control groups, with IPS participants securing employment 10 weeks earlier (Bond et al., 2008a).

Four years later a second systematic review compared USA and non-USA findings (Bond et al., 2012b). Nine RCTs in the USA were compared with the results from six outside the USA. The USA sites outperformed the non-USA sites. The proportion of IPS participants commencing employment in the USA sites averaged 62.1%, compared to 47.3% in the non-USA sites. Non-USA RCTs were conducted in Canada, Europe, the UK, and Australia. The international variability in IPS performance was considered by the authors to be due to variation in IPS fidelity and differences in labour markets, social, and political policies (Bond et al., 2012b).

Other relevant research and practice developments included the establishment in 2001 of the IPS learning collaborative (Drake et al., 2012, pp. 119-123), the extension of IPS principles from six to eight (Evans & Bond, 2008), a revised IPS fidelity scale with 25-items (Bond, Peterson, Becker, & Drake, 2012c), and the publication of a revised IPS Manual (Swanson & Becker, 2013). These developments helped advance the operationalisation of IPS principles, standardise implementation, and reduce variability in IPS program performance particularly outside the USA (see Table 1).

Despite these advancements some important limitations of the IPS approach remained. The apparent reduced effectiveness of IPS programs outside the USA remained unexplained. In addition, on average around 40% of participants in IPS programs did not commence competitive employment, even though this was their primary goal at program entry. These concerns motivated researchers and practitioners to trial further enhancements and adaptations to the IPS approach (Boycott et al., 2012; Marino & Dixon; 2014).
Table 1. Overview of key publications in the evolution of IPS efficacy and effectiveness research

<table>
<thead>
<tr>
<th>Date</th>
<th>Authors</th>
<th>Study description</th>
<th>Study’s contribution to knowledge</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>Drake et al.</td>
<td>First RCT IPS vs. group skills training.</td>
<td>IPS efficacy over TVR.</td>
</tr>
<tr>
<td>1997a</td>
<td>Bond et al.</td>
<td>IPS-15 scale developed and validated using data from 27 sites in the USA.</td>
<td>IPS-15 scale has good interrater reliability, internal consistency, and discriminative validity.</td>
</tr>
<tr>
<td>1999b</td>
<td>Drake et al.</td>
<td>Second RCT IPS vs. TVR.</td>
<td>IPS efficacy over TVR, and IPS generalisability.</td>
</tr>
<tr>
<td>2003</td>
<td>Becker &amp; Drake</td>
<td>Published overview of IPS approach in A Working Life.</td>
<td>Outlined underpinning theory, six principles, and the practices of IPS.</td>
</tr>
<tr>
<td>2006</td>
<td>Latimer et al.</td>
<td>RCT in Canada comparing IPS with TVR.</td>
<td>Efficacy of IPS outside of the USA.</td>
</tr>
<tr>
<td>2008</td>
<td>Bond et al.</td>
<td>Systematic review of 11 RCTs. On average 61.0% participants in IPS intervention commence employment compared with 23.0% in TVR.</td>
<td>Efficacy of IPS over other forms of VR.</td>
</tr>
<tr>
<td>2008</td>
<td>Evans &amp; Bond</td>
<td>Delphi survey to identify critical ingredients of IPS.</td>
<td>Principles extended to eight (benefits counselling and job development added).</td>
</tr>
<tr>
<td>2012b</td>
<td>Bond et al.</td>
<td>Systematic review of 15 RCTs. On average USA IPS intervention sites outperform non-USA (62.1% USA compared to 47.3% non-USA sites).</td>
<td>Whilst IPS is more effective than TVR, there is large variation in performance of IPS programs particularly outside the USA.</td>
</tr>
<tr>
<td>2012c</td>
<td>Bond et al.</td>
<td>IPS-25 scale developed and validated using data from 79 sites in the USA.</td>
<td>IPS-25 has promising psychometric properties including internal consistency and predictive validity.</td>
</tr>
<tr>
<td>2012</td>
<td>Boycott et al.</td>
<td>A rapid evidence review of IPS enhancements: six studies included.</td>
<td>The evidence of the effectiveness of IPS enhancements is inconclusive.</td>
</tr>
<tr>
<td>2013</td>
<td>Kinoshita et al.</td>
<td>Cochrane systematic review and meta-analysis, 14 RCTs.</td>
<td>IPS outperforms TVR on time to first job, proportion of participants who commenced employment and on length of employment.</td>
</tr>
<tr>
<td>Year</td>
<td>Authors</td>
<td>Description</td>
<td>Findings</td>
</tr>
<tr>
<td>------</td>
<td>---------------</td>
<td>------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>2014</td>
<td>Marshall et al.</td>
<td>Systematic review included 12 systematic reviews and 17 RCTs.</td>
<td>IPS outperforms TVR on time to first job, proportion of participants who commenced employment, length of employment and earnings.</td>
</tr>
<tr>
<td>2014</td>
<td>Marino &amp; Dixon</td>
<td>Narrative review of the IPS literature, 20 RCTs included.</td>
<td>Efficacy and generalisability of IPS, but there is considerable variance in performance.</td>
</tr>
<tr>
<td>2015</td>
<td>Becker et al.</td>
<td>Supported employment fidelity review manual.</td>
<td>Standardisation of IPS implementation.</td>
</tr>
<tr>
<td>2016</td>
<td>Modini et al.</td>
<td>Meta-analysis of 17 RCTs as well as two follow-up studies.</td>
<td>Geographic area or local unemployment rate do not affect the effectiveness of IPS.</td>
</tr>
<tr>
<td>2017a</td>
<td>Metcalfe et al.</td>
<td>Meta-analysis of 21 RCTs. Examined the relative efficacy of IPS vs. TVR in 12 different countries.</td>
<td>The labour and disability regulations of a country affect the performance of the IPS intervention relative to the control.</td>
</tr>
<tr>
<td>2017</td>
<td>Suijkerbuijk et al.</td>
<td>Cochrane meta-analysis examined 48 RCTs to compare effectiveness of four VR interventions, this included 13 studies of IPS enhancements.</td>
<td>IPS is more effective than TVR and transitional employment. Enhancements to IPS were slightly better than IPS, but not significantly.</td>
</tr>
</tbody>
</table>

Notes. IPS=Individual Placement and Support. RCT=Randomised Controlled Trial. vs.=versus. SE=supported employment. TVR=Traditional Vocational Rehabilitation. VR=Vocational Rehabilitation. IPS-15=The 15-item IPS fidelity scale. USA=United States of America. IPS-25=The 25-item IPS fidelity scale.
**Enhancements to IPS programs**

To investigate and address the suboptimal performance of IPS programs, researchers and practitioners examined the underlying reasons why some participants either do not gain employment outcomes, or when they do, why their job tenure can be brief. This research has explored sociological, agency, and program barriers, as well as disease-based barriers, to successful employment (Loveland et al., 2007). For example, McGurk and Mueser (2004) reviewed research on the relationship between cognitive functioning, symptoms, and getting and keeping employment. The authors concluded that supported employment was compensating for some cognitive impairments in people with severe mental illnesses. The authors developed a model for how cognitive factors and symptoms interact with a person’s participation in vocational rehabilitation programs. This led to the design and delivery of standardised cognitive remediation training, in particular the Training Skills for Work Program (McGurk, Mueser, & Pascaris, 2005).

However, research exploring the influence of different types of cognitive remediation training has mixed findings (Boycott et al., 2012). The quality of trials is problematic, particularly the small sample sizes and the poor outcomes reported for the IPS control (Suijkerbuijk et al., 2017). These poor outcomes indicate there were implementation problems with both the IPS intervention and the control group (McGurk et al., 2005; McGurk, Mueser, Feldman, Wolfe, & Pascaris, 2007).

A recent RCT compared IPS enhanced by cognitive remediation training (the intervention group), to an IPS program plus cognitive remediation training provided to the employment specialists (the control group) (McGurk et al., 2015). The IPS control group achieved 36.0% (18 out of 50) participants commencing competitive employment. This proportion is still well below the expected USA average for IPS program of 62.1% (Bond et al., 2012b), despite the IPS control site having high adherence to the IPS principles as measured by the IPS-25 and IPS-15 fidelity scales (104/125 and 70/75 respectively). Participants were limited to those people who had not benefited from high-fidelity IPS programs, defined as having been in the program for three months or more without obtaining competitive employment, or having lost a job within three months (McGurk et al., 2015).
Other enhancements to IPS include social skills training (Tsang, Chan, Wong, & Liberman, 2009), behavioural feedback (Bell, Lysaker, & Bryson, 2003), and motivational interviewing (Craig et al., 2014). In these RCTs the IPS enhancements outperformed the standard IPS control sites. However, the design and implementation of these trials limits the interpretation of the results (Suijkerbuijk et al., 2017). For example, in the Craig et al. (2014) IPS enhancement study with young people with first episode psychosis, clinicians were trained in motivational interviewing. The two IPS control groups only succeeded in supporting 17.6% and 18.8% of participants into employment at 12 months, despite measuring exemplary fidelity (a score of 110 or more out of 125) on the IPS-25 fidelity scale. This is well below the average performance of 47.3% previously found in non-USA IPS programs (Bond et al., 2012b). This indicates implementation problems broader than adherence to IPS principles. When overall program delivery quality is low any additions or enhancements regardless of their efficacy, are likely to improve program performance (Durlak, 2016).

A 2017 Cochrane systematic review and meta-analysis examined four approaches to vocational rehabilitation: IPS enhancements, IPS programs, transitional employment, and traditional vocational rehabilitation. The review included 48 RCTs, 13 of which were studies of enhancements to IPS programs compared to an IPS intervention as the control (Suijkerbuijk et al., 2017). Whilst the enhancements seemed to improve employment outcomes and length of time in employment over the standard IPS programs, the overall difference was not statistically significant. The small number of IPS enhancement studies, along with limitations to study design, restricted the specific comparisons that could be examined (Suijkerbuijk et al., 2017).

**Addressing Variation in Performance**

Significant advancements in the science and practice of IPS vocational rehabilitation were made from the early 1990s to the present which established IPS efficacy and effectiveness internationally. However, the research on enhancements to IPS has not yet resulted in new evidence-based practices that make significant differences to the performance of IPS programs (Boycott et al., 2012;
Suijkerbuijk et al., 2017). Research investigating sources of variation in performance across programs has also had mixed results. A meta-analysis of 17 RCTs of IPS investigated IPS effectiveness in different countries and in different economic conditions (Modini et al., 2016). This analysis found that “neither geographic area, nor unemployment rates affected the overall effectiveness of IPS” (Modini et al., 2016, p. 14). In contrast, Metcalfe et al. (2017a) conducted a meta-analysis of 21 RCTs and found that labour market and welfare benefits regulations affected the relative performance of the IPS intervention over the control. These studies used different measures to investigate the influence of the wider socio-economic context. Future research involving these background influences would benefit from establishing consistent methods of measurement (Drebing et al., 2012).

The current state of research shows that the performance of IPS programs does not seem to be increasing and may even have reached a plateau. A possible explanation is the current conceptual framework for IPS implementation. The basis for this implementation framework is that variation in performance is predominantly attributed to variation in implementation fidelity, and that fidelity is the primary mediator between wider contextual factors and IPS performance (Drake et al., 2012, p. 112). This IPS conceptual implementation framework was developed from findings of the National Evidence-Based Practices project. This project examined the implementation of five evidence-based practices across eight states in the USA (Drake et al., 2001; Torrey, Bond, McHugo, & Swain, 2012). Contextual factors, such as organisational arrangements or state financing and leadership, were classified as either helping or hindering program implementation. In the Drake et al. (2012, p. 112) conceptual framework, the scope of the definition of the quality of program implementation is limited to adherence to evidence-based practices, i.e. implementation fidelity. In contrast, other authors from the prevention science field contend that the concept of implementation quality is much broader than program fidelity. These authors argue that a definition of implementation quality should encompass other dimensions of quality such as dosage, program delivery, program reach, and the responsiveness of participants (Dane & Schneider, 1998; Durlak & DuPre, 2008) (see Figure 1).
Figure 1. Broadening the conceptualisation of program implementation quality

Notes. In the science and practice of vocational rehabilitation, program implementation quality has been conceptualised as program fidelity (Drake, Bond, & Becker, 2012, p. 112). In this conceptualisation contextual factors such as training, policy and funding, staffing, and organisations factors, are hypothesised to influence program performance and this influence is considered largely mediated through program fidelity. The dotted line around the program fidelity box depicts a broader conceptualisation of program implementation quality suggested by researchers from the prevention science literature (Dane & Schneider, 1998; Durlak & DuPre, 2008).

Chapter Conclusions

The key issue emerging from the historical exploration of the science of vocational rehabilitation outlined in this chapter, is whether or not implementation fidelity is the main driver of vocational rehabilitation program performance. If this is the case, empirical studies are expected to show that the majority, if not all, of the variance in program performance is explained by program fidelity. On the other hand, if other broader aspects of implementation quality are important in addition to program fidelity, then empirical studies are expected to show that only a minority, or small proportion, of the total variance in program performance is explained by program fidelity. The next empirical question is therefore: What is the relationship between implementation fidelity and program performance internationally? This question is examined in detail in Chapter 3 which is a systematic review and meta-analysis of IPS implementation across international contexts.
Chapter 3: Program Fidelity and Program Performance

Prelude

Rationale for study 1: Systematic review and meta-analysis

As described in Chapter 2, whilst the efficacy, effectiveness, and generalisability of the Individual Placement and Support (IPS) approach to vocational rehabilitation is well established, performance remains suboptimal, and there continues to be large variations in performance across studies, and across countries. Recent research seeking to advance the science and practice of evidence-based practices in vocational rehabilitation by adding enhancements to IPS programs has identified some emerging and promising practices. However, evidence that these practices consistently and reliably improve performance, is needed (Suijkerbuijk et al., 2017).

This thesis aims to generate knowledge which will improve the performance of evidence-based practices in vocational rehabilitation. A starting point therefore, is a better understanding of variation in performance across and between programs. Reviewing the relationship between adherence to known evidence-based practices, also referred to as implementation fidelity (Carroll et al., 2007) and program performance across international contexts, could potentially help to identify reasons for variation in performance. The commonly accepted method for demonstrating adherence to evidence-based practices in vocational rehabilitation for people with severe mental illnesses is through measurement using one of the two validated IPS fidelity scales: IPS-15, the 15-item IPS fidelity scale (Bond et al., 1997a); and IPS-25, the 25-item IPS fidelity scale (Bond et al., 2012c). A widely accepted measure of the performance of vocational rehabilitation programs is the number of participants commencing competitive employment within a defined cohort over a defined period (Bond, Campbell, & Drake, 2012a; Drebing et al., 2012).

The following research question informed this investigation: What is the nature of the relationship between implementation fidelity of vocational rehabilitation programs and effectiveness in an international context? Two hypotheses were constructed.
1. Vocational rehabilitation programs which adhere more closely to the principles and practices of the IPS approach to vocational rehabilitation will have better outcomes.

2. The strength of the relationship between fidelity of implementation and employment outcomes will be affected by other influences. These influences were examined as covariates and the ones available to this analysis were: the diagnostic mix of program participants; participant’s initial vocational goals; study type (RCT or observational); region of study; and study duration.

To answer the two hypotheses a systematic review and meta-analysis of the international literature on IPS implementation was undertaken. As well as identifying studies for inclusion in the meta-analysis, the literature review aimed to bring together current knowledge on the association between total fidelity score and program outcomes and identify any other potential influences outside what is measured by the current fidelity scales that may be impacting on performance.


Following the published paper, the results of supplementary analysis conducted as part of the systematic review and meta-analysis are presented, these were not included in the published paper. The chapter finishes with concluding remarks to lead into Chapter 4.
Introduction

The quality of program implementation can influence program effectiveness (Durlak & DuPre, 2008) and this is well known in the context of supported employment for people with severe mental illnesses (Bond, 2007). There is a critical stage in developing any evidence-based psychosocial rehabilitation program where it becomes necessary to specify the conditions and requirements for good program implementation. Without such specifications, any previously demonstrated program efficacy could erode through poor implementations to the point where, poor outcomes are misconstrued as ineffectiveness (Bond, 2007). Sometimes implementation specifications evolve into a formal measure of quality of adherence to original program principles and practices. These are often described as measures of program fidelity (Mowbray et al., 2003).

A measure of program fidelity enables differences in program performance to be compared and contrasted while controlling for the quality of implementation. Fidelity measures are also useful for identifying ways in which programs can be developed further in order to be even more effective. To be useful for these purposes, fidelity measures need to demonstrate adequate reliability and validity. One particularly relevant property is predictive validity, which measures the extent that higher program fidelity scores are uniquely associated with good program outcomes and the extent that lower scores are uniquely associated with poor program outcomes (Mowbray et al., 2003).

Examples of fidelity measures are found in the individual placement and support (IPS) approach to supported employment. This is a specialised form of supported employment for people with severe mental illnesses. IPS has demonstrated efficacy for attaining competitive employment outcomes (Bond et al., 2008a, 2012b). It is a well-defined and extensively researched program, supported by findings from over 20 RCTs (Drake & Bond, 2014), 12 systematic reviews, including meta-analyses (Marshall et al., 2014), and a recent Cochrane review (Kinoshita et al., 2013). These studies consistently show that IPS is significantly more effective than alternative approaches to
vocational rehabilitation in terms of the proportion of participants commencing competitive employment.

Two IPS fidelity scales have been developed to date. Both scales, along with their supporting documents, aim to provide a clear specification of the intervention and how to implement its practices and principles. The first scale, a 15-item version (the IPS-15), was revised in 2008 to 25 items (IPS-25) also known as the supported employment fidelity scale. The latest 25-item version is supported by an instruction manual on how best to assess the program, administer, and code the scale (Swanson & Becker, 2013).

The IPS-15 and IPS-25 scales are sufficiently different to require separate examination. Both scales have a similar structure, and both have common and unique questions. In both scales each item has a five-point (1-5) rating with anchor descriptions for each. A score of 5 represents good implementation, and a score of 1 indicates poor implementation, or a failure to implement that practice. The score range on the IPS-15 is 15-75, while the score range on the IPS-25 is 25-125.

The psychometric properties of the IPS-15 include good inter-rater reliability, and good discriminant validity for distinguishing between IPS and non-IPS programs (Bond et al., 1997a). Predictive validity of the IPS-15 was examined in a meta-analysis of ten studies which reported employment outcomes and total fidelity scores. Six of ten studies reported an association between employment outcomes and fidelity. However, across all ten studies the proportion of variance in employment outcomes explained by fidelity score ranged from 5 to 58% (Bond, Becker, & Drake, 2011). This result is problematic and suggests that the predictive validity of the scale is either unstable or not sufficiently universal, through being too dependent on contextual factors.

It is possible that differences in how employment outcomes are measured account for the varied strength of prediction. For example, some studies used a point-in-time measure of participant employment status rather than tracking a defined cohort over time. The point estimates were from quarterly reports of the proportion employed among the current caseload (Becker, Smith, Tanzman,
Drake, & Tremblay 2001b; Becker, Xie, McHugo, Halliday, & Martinez, 2006). Another study used vocational rehabilitation case closure as a proxy measure for employment outcomes (Hepburn & Burns, 2007), again without defining the cohort. Yet another study measured the difference between the employment proportion in the IPS intervention site and the employment proportion among controls, as a measure of strength of effect (Burns et al., 2006, 2007; Catty et al., 2008).

The psychometric properties of IPS-25 were examined using a large data set collected from 79 ongoing IPS implementation sites across the USA. Because it was developed more recently to address the limitations of the IPS-15 scale, the IPS-25 is expected to have greater predictive validity than the IPS-15. Predictive validity was assessed using total fidelity scores and quarterly reports of the proportion of the caseload employed at each site. A monotonic positive relationship was found between total IPS-25 score and the point-in-time proportions employed ($r = .34, p = <.01$), (Bond et al., 2012c). Although the IPS-25 scale had strengths it did not show better predictive validity over the IPS-15 scale. Consequently, the authors recommended the use of both scales until further evidence on the psychometric properties of the IPS-25 became available (Bond et al., 2012c). A more recent study of 79 sites across the USA, where some sites differed from the earlier study, found a positive association between fidelity and point-in-time proportions employed ($r = .27, p = .02$), (Kim, Bond, Becker, Swanson, & Langfitt-Reese, 2015).

However, point-in-time proportions employed do not necessarily involve the same cohort, they do not account for attrition, or take into account new program commencements, or duration of assistance. In fact, it is possible to have a completely new cohort at each time point, with short average durations of participation in the program. It is also possible to correctly report 100% of the caseload as employed, with no new jobs commencing in the quarter, and no new participants commencing the program. This would occur if all non-workers dropped out of the program, which would also indicate high attrition. For these reasons the defined cohort is considered a more reliable method for reporting employment outcomes. This approach permits unbiased estimates of employment commencements.
because new program commencements, duration of assistance, and attrition can be taken into account. Despite these limitations point-in-time measures are often used outside controlled trials. It is also possible to supplement point-in-time monitoring of proportions employed with rolling averages for new program commencements and rolling averages for attrition.

Other measurement issues have emerged. A systematic review comparing the effectiveness of IPS in the USA to IPS outside the USA, found that competitive employment commencements in the USA programs were higher than in the non-USA studies (Bond et al., 2012b). Lack of fidelity to IPS was suggested as explaining the poor performance of the non-USA studies. The authors concluded that IPS is transportable outside the USA provided programs can attain high fidelity. Yet, differences in total fidelity scores across countries could also involve fidelity measurement error. This can arise through different methods such as using trained external assessors, using internal assessments, or measuring fidelity at different operational levels (Mowbray et al., 2003). In addition, factors other than fidelity that may confound this relationship can vary in strength across countries. For example, local background unemployment, or the extent to which labour market regulation inhibits access to entry-level jobs (Catty et al., 2008; Jonsdottir & Waghorn, 2015). This is important because the association between fidelity and employment outcomes often explains less than half of the variance in employment outcomes (Bond et al., 2011, 2012c; Kim et al., 2015).

This review explored these issues further by re-examining the association between fidelity and employment outcomes in a large international sample of published studies of IPS implementations. The review was restricted to studies which followed a defined cohort of participants, with mental illnesses, in a vocational rehabilitation program for at least six months, using a common outcome variable. This variable was the number of participants who entered the program who went on to commence one day or more of competitive employment, compared to the remainder who entered the program who did not commence employment. The primary research question concerned the predictive validity of the two
IPS fidelity scales, namely, the extent that total fidelity scores on each scale predicted new employment commencements.

Secondary research questions were specified to explore six cohort characteristics as potential influences on the strength of the relationship between fidelity and employment commencements. Since IPS was originally designed for those with the most severe forms of mental illnesses, a stronger relationship between fidelity and employment commencements was expected in studies with greater proportions of participants with psychotic disorders. In addition, the variance in employment outcomes was expected to be greater between unrelated studies than among cohorts within a single multi-site study. Cohorts from RCTs were expected to differ in some ways from those in observational studies, due to the different design conditions under which these study types are often conducted. The relationship between fidelity and employment commencements was expected to be stronger in USA cohorts, and in cohorts with clear goals of competitive employment, compared to those with less well-defined goals. A positive association was also expected between duration of assistance and employment outcomes because greater duration of assistance increases the opportunities for commencing employment.

**Methods**

The current review covered both RCTs and observational studies. The preferred reporting items for systematic review and meta-analyses were followed (Moher et al., 2009). In total, 14 essential and desirable criteria for study inclusion specified the participants, interventions, comparisons, outcomes, and study design (Moher et al., 2009) (see Table 2). The essential criteria ensured only studies using defined cohorts were considered, sampling similar populations who were not employed at baseline, exposed to a vocational intervention which was measured using one of the two IPS fidelity scales, and tracked with comparable outcome measures for a minimum period of six months. The included studies are shown in Table 3.
Search strategy

A literature search was conducted to identify published articles reporting cohort-based implementations of supported employment. Eight electronic databases were searched by author HL, who holds a master’s degree, and received training from a librarian in search methods. Relevant articles published in English were sought from January 1996 to June 2015. Earlier periods were not considered because the first report of an IPS fidelity scale did not emerge until 1996. The databases searched were: MEDLINE, PsychINFO, Scopus, EMBASE, Cinahl, Allied and Complementary Medicine, Campbell Collaborative Systematic review, and the Cochrane controlled trials library.

The following search terms were combined with the Boolean operators “OR” and “AND”: supported employment, vocational rehabilitation and IPS; psychosis, mental illness, psychiatric disability, schizophrenia and mental health; fidelity, scale and quality; vocation*, employ*, job*, and outcome*. Contents pages for key journals were examined from January 1996 to June 2015. These were the Journal of Vocational Rehabilitation, the American Journal of Psychiatry, the Psychiatric Rehabilitation Journal, and Psychiatric Services. In addition, the reference lists of candidate studies were examined. This step began by examining the included and excluded studies listed in a recent IPS Cochrane systematic review (Kinoshita et al., 2013). Finally, researchers who previously published this type of report were contacted by e-mail to assist in the identification of any new studies.

Data analysis

After applying the inclusion criteria, the articles retained were assessed by author HL for the following characteristics: First author, study date, study type (RCT or observational), type of vocational intervention, IPS scale, country of implementation, cohort size, duration of study, proportion of participants with a diagnosis of psychosis, primary goal of participants, days to first job, number commencing employment over the study period, study attrition, total fidelity score, frequency of fidelity measurement, method for assessing fidelity, and level of fidelity assessment (employment specialist or employment team).
Authors were contacted if any data were missing, incomplete or unclear with respect to the dependent variable employment commencements, or with respect to the focal independent variable, total fidelity score. When several fidelity scores were provided, the score closest to the end of the study follow-up period was used. Where a narrow fidelity score range was provided, the midpoint of that range was calculated. The binary dependent variable represented the actual number of participants in each cohort commencing competitive employment; and its complement the actual number of participants who were not known to have commenced employment during the follow-up period. Each cohort represented all those that commenced receiving employment assistance as the denominator for calculating the percent commencing employment. This method matches previous studies of supported employment that have applied this intention to treat assumption (Kinoshita et al., 2013). Analyses were conducted in R (version 3.2.2), a language and environment for statistical computing (R Core Team, 2015).

Study selection

In total 366 candidate articles were identified. A two-stage assessment process was then applied using the essential inclusion criteria (see Figure 2). The first step involved a review of titles and abstracts. From this step 221 studies were excluded. The second step consisted of carefully reading each article with respect to each inclusion criterion. From this step 111 published reports were excluded.

The 111 excluded studies included five reports that were included in previous reviews. These five reports failed to meet our essential criterion, that all reports use a common dependent variable to define employment outcomes. The most common definition of competitive employment in this literature is one day or more of employment during the follow-up period. In contrast, Hoffmann, Jäckel, Glauser, and Kupper (2012) and Hoffmann, Jäckel, Glauser, Mueser, and Kupper (2014) defined employment as two weeks or more. Heslin et al. (2011) and Howard et al. (2010) defined employment as at least 30 days. Tsang et al. (2009) defined employment as continuous work for two
months or more for at least 20 hours per week. Hence, all five excluded reports involved more conservative definitions of employment that introduce aspects of employment duration or hours worked.

In total 34 studies met the essential inclusion criteria. From these four were removed through duplicate reporting of the same cohort, leaving 30 studies in the review. These 30 studies defined 69 distinct cohorts, where each cohort was exposed to a defined vocational intervention, with a known total fidelity score, and with outcomes reported in terms of commencing employment, tracked over a minimum period of six months (see Table 2). Thus, control groups for which the fidelity of the intervention to IPS was not measured were excluded. In the second stage, each included study was assessed according to how it met six additional desirable inclusion criteria. The aim of this step was to assess the included studies for optimal quality, including the identification of any potential sources of bias and their relevance to this purpose. No studies met all of the desirable criteria.

The combined data consisted of 69 cohorts reported in 30 studies published between 1999 and 2015, and conducted in 12 countries: USA, Canada, England, Germany, Italy, Switzerland, Bulgaria, the Netherlands, China (Hong Kong), Japan, Australia, and New Zealand. In total, 52 cohorts reported total IPS fidelity score using the IPS-15, while 17 reported fidelity measured by the IPS-25. Two (McGurk et al., 2015) cohorts reported total fidelity with both IPS-15 and IPS-25 and were included in both analyses. Across all 69 cohorts there were 8,392 program participants (7,216 participants in IPS-15 cohorts and 1,176 participants in IPS-25 cohorts).
Table 2. Study inclusion criteria and rationale for each

<table>
<thead>
<tr>
<th>Essential criteria</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>A A suitable cohort is reported in English language in a peer reviewed published report.</td>
<td>Reports published as monographs, dissertations, and unpublished studies were not included as these are less accessible to readers and of less consistent quality through not being peer reviewed.</td>
</tr>
<tr>
<td>B A cohort of not currently employed adults (aged 16 to 64 years) with a diagnosed mental illness is described and exposed to a defined vocational intervention.</td>
<td>The cohort needs to be sufficiently defined to enable comparisons between studies. It is also important that only those not employed at baseline are included in the cohort.</td>
</tr>
<tr>
<td>C The cohort is defined in terms of Intention To Treat (ITT) assumptions or sufficient information is available to calculate ITT.</td>
<td>These assumptions enable a common denominator for the outcome measure to be specified as those who entered the program and commenced to receive assistance. Cohorts defined by non-ITT assumptions may specify a different denominator, such as those who remained in the program at follow-up.</td>
</tr>
<tr>
<td>D The primary employment outcome (numerator) is defined as one day or more of competitive employment in the open labour market, or being self-employed, in jobs not reserved for people with disabilities, and paid at national minimum wages or above.</td>
<td>A common definition of competitive employment is needed to enable outcomes from different studies to be compared. Some studies report similar definitions of competitive employment but can diverge by the time period defining commencing employment. For instance, some studies include all jobs of one day or more, whereas others require a job to be held for at least a month for that job to be counted.</td>
</tr>
<tr>
<td>E The vocational intervention was assessed at least once over the study period using either the 15-item or 25-item IPS fidelity scale.</td>
<td>This is necessary because the purpose of the study is to investigate the relationship between evidence-based practices as reflected in Individual Placement and Support (IPS) fidelity scores, and competitive employment outcomes.</td>
</tr>
<tr>
<td>F The number of participants gaining competitive employment was reported.</td>
<td>This is the primary outcome variable (numerator) of interest. Raw numbers and proportions are needed. Some studies fail to report these essential frequency and descriptive statistics.</td>
</tr>
<tr>
<td>G Each participant in the defined cohort was followed for at least six months.</td>
<td>The longer the period that a person receives assistance, the more likely they are to be successful. Hence it is important to control exposure to any variable expected to contribute to outcomes.</td>
</tr>
<tr>
<td>H The diagnostic mix of participants in each cohort is reported.</td>
<td>Since those with more severe psychiatric disabilities are likely to need more intensive assistance (Jonsrud and Waghorn, 2015) it is considered important to measure and control for diagnostic mix in each cohort.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Desirable criteria</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Attrition is adequately defined and reported.</td>
<td>Attrition can be an alternative explanation for results and should be reported. Some studies report attrition as those that left the program or ceased providing information to the research team before obtaining a vocational benefit. This is the preferred definition, but others will also be considered.</td>
</tr>
<tr>
<td>B The cohort was part of a single site or multi-site randomised controlled trial, or multi-site observational study.</td>
<td>Cohorts that are part of a larger study design often have common features (method of implementation, participant characteristics, background variables, outcome measurement) that reduce noise when included in a meta-analysis.</td>
</tr>
<tr>
<td>C The nature of the vocational intervention was assessed according to the IPS-25 fidelity scale.</td>
<td>The IPS-25 is the latest version of the IPS fidelity scale and is expected to be more predictive of competitive employment outcomes than the previous version, the IPS-15.</td>
</tr>
<tr>
<td>D The majority of participants had a diagnosis of a severe mental illness, or the diagnostic mix of participants is well described.</td>
<td>People diagnosed with some mental illnesses, on average, need less intensive forms of employment assistance than others. Evidence-based supported employment is intended for people with a diagnosis of severe mental illness. Hence it will be necessary to control for diagnostic mix in the meta-analysis.</td>
</tr>
<tr>
<td>E Each participant in each cohort had an initial goal of competitive employment.</td>
<td>IPS is designed for people who want to be employed and discourages expectations for less than competitive employment.</td>
</tr>
<tr>
<td>F Item-level fidelity scores were reported as well as total fidelity scores.</td>
<td>Item-level fidelity scores enable the contribution of individual fidelity items to be assessed.</td>
</tr>
</tbody>
</table>

Notes. IPS-15=The 15-item IPS fidelity scale. IPS-25=The 25-item IPS fidelity scale.
Figure 2. Flowchart of article screening and assessment against the inclusion criteria.

Notes. n=number of articles. IPS-15=The Individual Placement and Support 15-item fidelity scale. IPS-25=The Individual Placement and Support 25-item fidelity scale.
Table 3. *Studies included in the systematic review and meta-analysis*

<table>
<thead>
<tr>
<th>First author, date published</th>
<th>Location</th>
<th>Total sample size in study</th>
<th>Follow up period (months)</th>
<th>Number of included cohorts</th>
<th>Cohort sizes</th>
<th>Brief description of study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drake et al., 1999b</td>
<td>Washington D.C. USA</td>
<td>152</td>
<td>18</td>
<td>1</td>
<td>76</td>
<td>RCT comparing IPS vs. step-wise vocational rehabilitation services. Participants had a goal of competitive employment.</td>
</tr>
<tr>
<td>Lehman et al., 2002</td>
<td>Baltimore, USA</td>
<td>219</td>
<td>24</td>
<td>1</td>
<td>113</td>
<td>RCT comparing IPS vs. step-wise vocational rehabilitation services. Participants did not need to have a desire to work.</td>
</tr>
<tr>
<td>Mueser et al., 2004</td>
<td>Hartford, USA</td>
<td>204</td>
<td>24</td>
<td>3</td>
<td>67, 68, 69</td>
<td>An RCT with three cohorts which compared IPS with psychosocial treatment and with non-integrated supported employment services. The study was conducted in an inner city, participants were African American or Latino and had a goal of competitive employment.</td>
</tr>
<tr>
<td>McGurk et al., 2005</td>
<td>New York, USA</td>
<td>44</td>
<td>12</td>
<td>2</td>
<td>21, 23</td>
<td>RCT with IPS as the control vs. IPS with cognitive training. As part of the inclusion criteria people had to have a history of at least one unsatisfactory job ending. Participants had a goal of competitive employment.</td>
</tr>
<tr>
<td>Gold et al., 2006</td>
<td>South Carolina, USA</td>
<td>143</td>
<td>24</td>
<td>1</td>
<td>66</td>
<td>RCT comparing IPS plus assertive community treatment vs. step-wise vocational rehabilitation services. 77% of the participants were African American and all had a goal of competitive employment and lived in rural areas.</td>
</tr>
<tr>
<td>Latimer et al., 2006</td>
<td>Montreal, Canada</td>
<td>150</td>
<td>12</td>
<td>1</td>
<td>75</td>
<td>RCT comparing IPS vs. step-wise vocational rehabilitation services. Participants had a goal of competitive employment.</td>
</tr>
<tr>
<td>Burns et al., 2006, 2007</td>
<td>Six cities in Europe</td>
<td>312</td>
<td>18</td>
<td>12</td>
<td>25-27</td>
<td>Multi-site RCT comparing IPS vs. step-wise vocational rehabilitation services. Participants had a goal of competitive employment.</td>
</tr>
<tr>
<td>Bond et al., 2007</td>
<td>Chicago, USA</td>
<td>187</td>
<td>24</td>
<td>1</td>
<td>96</td>
<td>RCT comparing IPS vs. step-wise vocational rehabilitation services. Participants had a goal of any paid work not limited to competitive employment.</td>
</tr>
<tr>
<td>Wong et al., 2008</td>
<td>Hong Kong, China</td>
<td>92</td>
<td>18</td>
<td>1</td>
<td>46</td>
<td>RCT comparing IPS vs. step-wise vocational rehabilitation services. Participants had a goal of competitive employment.</td>
</tr>
<tr>
<td>First author, date published</td>
<td>Location</td>
<td>Total sample size in study</td>
<td>Follow up period (months)</td>
<td>Number of included cohorts</td>
<td>Cohort sizes</td>
<td>Brief description of study</td>
</tr>
<tr>
<td>------------------------------</td>
<td>----------</td>
<td>---------------------------</td>
<td>--------------------------</td>
<td>---------------------------</td>
<td>--------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>Killackey et al., 2008</td>
<td>Melbourne, Australia</td>
<td>41</td>
<td>6</td>
<td>1</td>
<td>20</td>
<td>RCT comparing IPS vs. treatment as usual (referred by the mental health team to an external vocational service and/or to a vocationally orientated group program run by the mental health team). All participants were young people aged 15-25 with a diagnosis of first episode psychosis and had an initial vocational goal (which could include voluntary work and education).</td>
</tr>
<tr>
<td>Davis et al., 2012</td>
<td>Tuscaloosa, USA</td>
<td>85</td>
<td>12</td>
<td>2</td>
<td>42, 43</td>
<td>RCT comparing IPS vs. vocational rehabilitation services. All participants were veterans with post-traumatic stress disorder with an initial goal of competitive employment. People with a diagnosis of psychosis or bi-polar disorder were excluded from this study.</td>
</tr>
<tr>
<td>Drake et al., 2013; Frey et al., 2011</td>
<td>23 cities across the USA</td>
<td>2059</td>
<td>24</td>
<td>1</td>
<td>1121</td>
<td>RCT comparing a multi-faceted intervention which consisted of IPS, medical management and health services vs. usual mental health treatment services. Participants could have a goal of any paid work (not limited to competitive employment).</td>
</tr>
<tr>
<td>Oshima et al., 2014</td>
<td>Tokyo, Japan</td>
<td>37</td>
<td>6</td>
<td>1</td>
<td>18</td>
<td>RCT comparing IPS vs. step-wise vocational rehabilitation services. Participants did not need to have a desire to work.</td>
</tr>
<tr>
<td>Waghorn et al., 2014</td>
<td>Queensland, Australia</td>
<td>208</td>
<td>12</td>
<td>8</td>
<td>20-42</td>
<td>A four-site RCT with IPS as the experimental intervention and a control intervention at each site. The control group used enhanced treatment as usual where the mental health team was assisted in making referrals to external disability employment services. Participants all had a goal of competitive employment.</td>
</tr>
<tr>
<td>Lucca et al., 2004</td>
<td>Massachusetts, USA</td>
<td>90</td>
<td>54</td>
<td>1</td>
<td>90</td>
<td>A retrospective evaluation over four and a half years examining the employment outcomes achieved by people participating in supported employment services. The initial goal of participants is not specified but they had all voluntarily enrolled in an employment service.</td>
</tr>
<tr>
<td>First author, date published</td>
<td>Location</td>
<td>Total sample size in study</td>
<td>Follow up period (months)</td>
<td>Number of included cohorts</td>
<td>Cohort sizes</td>
<td>Brief description of study</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>----------</td>
<td>---------------------------</td>
<td>---------------------------</td>
<td>---------------------------</td>
<td>--------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>Porteous &amp; Waghorn, 2007</td>
<td>Wellington, New Zealand</td>
<td>225</td>
<td>24</td>
<td>2</td>
<td>100, 125</td>
<td>A prospective study of two cohorts of young people aged 14 to 27 with a diagnosis of first episode psychosis and a goal of any work (which could include voluntary work).</td>
</tr>
<tr>
<td>Van Erp et al., 2007</td>
<td>Four sites across the Netherlands</td>
<td>316</td>
<td>24</td>
<td>4</td>
<td>53-125</td>
<td>A prospective multi-site study of the implementation of IPS in 4 regions. Participants all had a goal of competitive employment.</td>
</tr>
<tr>
<td>Browne et al., 2009</td>
<td>Hawkes Bay, New Zealand</td>
<td>125</td>
<td>48</td>
<td>1</td>
<td>125</td>
<td>A retrospective case study examining employment outcomes achieved by a non-integrated supported employment service. Participants had a goal of any paid work (not limited to competitive employment).</td>
</tr>
<tr>
<td>Browne et al., 2010</td>
<td>Christchurch, New Zealand</td>
<td>49</td>
<td>24</td>
<td>1</td>
<td>49</td>
<td>A retrospective case study examining employment outcomes achieved by a non-integrated supported employment service working with young people aged 16-25. Participants had a goal of any paid work (not limited to competitive employment).</td>
</tr>
<tr>
<td>Waghorn et al., 2011b</td>
<td>Christchurch, New Zealand</td>
<td>270</td>
<td>24</td>
<td>1</td>
<td>270</td>
<td>A retrospective study examining the employment outcomes achieved by a non-integrated supported employment service. Participants had a goal of any paid work (not limited to competitive employment).</td>
</tr>
<tr>
<td>Henry et al., 2014</td>
<td>Massachusetts, USA</td>
<td>3474</td>
<td>12</td>
<td>1</td>
<td>3474</td>
<td>A retrospective evaluation over nine years of participants receiving supported employment services. Participants had a goal of any paid work (not limited to competitive employment).</td>
</tr>
<tr>
<td>Williams et al., 2015</td>
<td>Queensland, Australia</td>
<td>114</td>
<td>13 to 26</td>
<td>3</td>
<td>28-52</td>
<td>A prospective study examining the employment outcomes across three sites implementing IPS services. Participants in one site all had a diagnosis of first episode psychosis. Participants all had a goal of competitive employment.</td>
</tr>
<tr>
<td>IPS-25</td>
<td>California, USA</td>
<td>58</td>
<td>12</td>
<td>1</td>
<td>30</td>
<td>RCT comparing IPS vs. step-wise vocational rehabilitation services. All participants were over 45 years with a diagnosis of schizophrenia and had an initial goal of any work (which could include voluntary work).</td>
</tr>
</tbody>
</table>

This article is © Emerald Publishing Limited and permission has been granted for this version to appear here https://researchspace.auckland.ac.nz/handle/2292/2. Emerald does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Emerald Publishing Limited.
<table>
<thead>
<tr>
<th>First author, date published</th>
<th>Location</th>
<th>Total sample size in study</th>
<th>Follow up period (months)</th>
<th>Number of included cohorts</th>
<th>Cohort sizes</th>
<th>Brief description of study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Craig et al., 2014</td>
<td>Midlands and London, England</td>
<td>159</td>
<td>12</td>
<td>4</td>
<td>38-42</td>
<td>RCT with IPS as the control vs. IPS plus motivational interviewing as the intervention arm. Participants were young people aged 18 to 35 with a diagnosis of first episode psychosis. Participants did not need to have a desire to work.</td>
</tr>
<tr>
<td>Bond et al., 2015</td>
<td>Chicago, USA</td>
<td>87</td>
<td>12</td>
<td>1</td>
<td>43</td>
<td>RCT comparing IPS to a comparison group offered a job club approach. Participants all had severe mental illnesses and a history of offending. All participants had a goal of competitive employment.</td>
</tr>
<tr>
<td>McGurk et al., 2015</td>
<td>New Hampshire and Chicago, USA</td>
<td>107</td>
<td>24</td>
<td>2</td>
<td>50, 57</td>
<td>RCT comparing enhanced IPS (with specialised cognitive training for employment specialists) as the control with IPS plus a cognitive enhancement program 'Thinking Skills for Work' as the intervention. Participants had not gained employment in other supported employment programs. All participants had a goal of competitive employment.</td>
</tr>
<tr>
<td>Kern et al., 2013</td>
<td>Los Angeles, USA</td>
<td>60</td>
<td>24</td>
<td>1</td>
<td>60</td>
<td>A prospective study examining the performance of three employment specialists with personal experience of mental illness implementing IPS; 25% of participants had education outcomes as their primary goal.</td>
</tr>
<tr>
<td>Morris et al., 2014</td>
<td>New South Wales, Australia</td>
<td>95</td>
<td>12</td>
<td>4</td>
<td>18-28</td>
<td>A prospective study over four sites implementing IPS through a formal partnership between mental health and disability employment services. All participants had a goal of competitive employment.</td>
</tr>
<tr>
<td>Marwaha et al., 2014</td>
<td>West Midlands, England</td>
<td>106</td>
<td>12</td>
<td>2</td>
<td>39, 57</td>
<td>A prospective study of the implementation of an employment intervention in two community mental health teams through either retraining an existing staff member or through employing a dedicated employment specialist. Participants had a goal of any work (which could include voluntary work).</td>
</tr>
<tr>
<td>van Vegge et al., 2014</td>
<td>Sussex, England</td>
<td>586</td>
<td>12</td>
<td>2</td>
<td>140, 446</td>
<td>A parallel group observational design. One cohort was followed prior to IPS implementation, one cohort post IPS implementation. Participants had mixed vocational goals which included education.</td>
</tr>
</tbody>
</table>

Total units in the meta-analysis: 69


This article is © Emerald Publishing Limited and permission has been granted for this version to appear here https://researchspace.auckland.ac.nz/handle/2292/2. Emerald does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Emerald Publishing Limited.
**Meta-analysis**

A meta-analysis explored the relationship between total fidelity score and the number of participants in all cohorts commencing employment, using summary data summarised by participants within a cohort, with cohorts as the unit of observation. Hence, each cohort provided a single observation for the independent variable (total fidelity score) and multiple participant-level observations for the binary dependent variable, commencing employment. Two separate analyses were performed, one for studies reporting fidelity using the IPS-15 scale, and another smaller set reporting fidelity using the IPS-25 scale.

Cohorts reporting IPS-15 fidelity scores were first classified as recommended by the scale authors into three categories: not supported employment, 15-56; fair fidelity, 56-65; and good fidelity, 66-75. A precision analysis and an analysis of negative predictive value were conducted. This approach was preferred over a sensitivity and specificity analysis for the main reason that low-fidelity implementations were under-represented, while high-fidelity implementations were overrepresented, in this sample. This is because low-fidelity cohorts were often control groups, and some of these were excluded if fidelity was not measured. In addition, IPS researchers usually strive to attain high fidelity therefore sampling for low fidelity requires a broader search outside the purpose of this investigation. The under-representativeness of the low-fidelity category favoured comparisons involving precision and negative predictive value over sensitivity and specificity.

The precision analysis began by examining a scatter plot of all cohorts by plotting both total fidelity score and the proportion of participants commencing competitive employment (see Figure 3). From this plot empirical thresholds for binary classification of fidelity and employment outcomes were determined. Values for precision and negative predictive value were calculated for three levels of fidelity score, and from counts of employment commencements, classified into three levels.

Multivariate analysis of deviance for summary data was then applied, as recommended for systematic reviews of this type. For a full discussion of the common issues and problems involved in...
This kind of meta-analysis, see Harvard and Lau (1993) and Petitti (1999). One property of our analysis of deviance is that it assumes all independent variables are random effects, modelling study heterogeneity as an important source of uncertainty. The dependent variable common to all studies was the frequency count of participants commencing employment complemented by the count of those not commencing employment. This was expressed as a log-odds ratio for commencing employment.

**Analysis of IPS-25 cohorts**

The cohorts were classified by fidelity score into the categories recommended by the scale authors for reporting fidelity, namely: not supported employment, 25-73; fair fidelity, 74-99; good fidelity, 100-114; and exemplary fidelity, 115-125. A precision and negative prediction analysis was conducted as for IPS-15 cohorts, to determine how well fidelity score at two levels predicted employment commencements at three levels. The small number of IPS-25 cohorts restricted any further statistical modelling of this data set.

**Results**

Total fidelity scores on the IPS-15 ranged from 24 to 74, in a possible range of 15-75 (Mean 61, SD 12.4). Total fidelity score on the IPS-15 was positively associated with employment commencements ($r=.41, r^2=.17$). Total fidelity scores on the IPS-25 ranged from 65 to 116, in a maximum range of 25-125 (Mean 96, SD 16.4).

We examined precision or positive prediction, and negative prediction or negative predictive value. Cohorts were initially categorised into whether they had a fair to good fidelity score on IPS-15 (a score of 56 or more), and good fidelity or better on IPS-25 (a score of 100 or more). Employment commencements for each cohort were classified into high (44% or more commencing employment) or low (43% or less commencing employment). This criterion was based on the visual cut off point from the scatter plot where scores below this all had low fidelity (see Figure 3).
Further categorisation of cohorts by different levels of fidelity and employment commencements enabled comparisons of different thresholds for precision and negative prediction (see Tables 4 and 5). Two additional fidelity thresholds (66 or more on IPS-15, and 115 or more on IPS-25) were added because these matched the scale authors’ definitions of good and exemplary fidelity, respectively. Further information on threshold selection is given in the footnotes to Table 5. Precision or positive prediction is the probability that high-fidelity programs obtain good employment outcomes. Precision values approaching 100% indicate accurate prediction of good employment outcomes from high fidelity. Negative prediction is the probability that low fidelity programs will have poor employment outcomes. Values approaching 100% indicate accurate prediction of poor employment outcomes from low-fidelity scores.

The threshold values which appear to optimise both precision and negative prediction on the IPS-15 are a fidelity score of 61 or more and 44% or more employment commencements (see Table 4). For IPS-25 the threshold for maximizing precision and negative prediction appears to be a fidelity score of 100 or more and employment commencements of either 44% or more or 47% or more (see Table 5). Both these employment outcome thresholds have the same properties, so either can be used. But if the threshold for employment outcomes is raised to 62% or more, the precision of the IPS-25 drops to 10 percent, while negative prediction is maintained at 85.7%. Since the IPS-25 results are based on only 17 observations, these parameters are likely to change in a larger sample of IPS-25 cohorts. It is worth noting however, that fidelity scores of 61 on IPS-15 (15 x 4 = 60), and 100 (25 x 4 = 100) on IPS-25, both represent mean item-level scores of four out of a maximum score of five for each fidelity practice item.
Figure 3. The proportion of employment commencements by fidelity score for cohorts measuring fidelity with the IPS-15 scale (n=52)\textsuperscript{1,2}

Notes. IPS-15=The Individual Placement and Support 15-item fidelity scale. \textit{n}=number of cohorts measuring fidelity with the IPS-15 scale. 1. An apparent minimum threshold for classification into good employment outcomes is shown as a horizontal dotted line and represents 44\% of participants or more commencing employment. 2. An apparent minimum threshold for fidelity is a score of 61, because no scores of 60 or less were associated with good employment outcomes and this is shown as a vertical dotted line. 3. The gradient represents the estimated relationship ($r=.41$) between fidelity and employment outcomes derived from the analysis of deviance calculated in log-odds ratios, refitted to match proportion of employment commencements.
Multivariate analysis

The heterogeneity of the IPS-15 studies are shown in Figure 3. Here all cohorts are ranked by fidelity score, with the highest in the first row. The forest plot in Figure 4 shows the relative odds of participants commencing employment in each cohort. The forest plot is consistent with the visual pattern in Figure 3, and with the results of the precision and negative prediction analyses. The cluster of cohorts with good fidelity yet low proportions commencing employment is also visible in Figure 4.

Multivariate analysis of deviance for the IPS-15 studies, using log-odds ratios, enabled the following cohort characteristics to be adjusted for fidelity score: study type (RCT or observational study); region of study (USA and Canada, Europe, or Australasia); author group (18 unique author groups) to represent non-random associations among cohorts due to being part of the same trial, or due to being reported by the same author; diagnostic mix, the proportion of participants with psychotic disorders (50% or more or less than 50%); participants’ employment goals (whether participants had competitive employment goals or mixed vocational goals including education, voluntary work or no employment goals); and study duration (durations of more than 12 months or 12 months or less).

Once each independent variable was adjusted for fidelity score, only author group ($\chi^2=5.01$, df=17, $p<.01$) and fidelity score ($\chi^2=15.31$, df=1, $p<.001$) were associated with commencing employment (see Table 6). The significance of author group suggests there are other unknown variables associated with differences between cohorts that can also account for differences in employment commencements. These are likely to include important country differences such as health and welfare policies, availability of entry-level jobs, and background unemployment. Yet, even with that caveat, the multivariate analysis highlights the importance of fidelity score as a robust predictor of employment commencements in this analysis after controlling for fidelity in six study characteristics.
Figure 4. The relative odds of participants commencing employment in each cohort ordered by fidelity measured with the IPS-15 scale (n=52)\(^1\)

**Notes.** IPS-15=The Individual Placement and Support 15-item fidelity scale. \(n=\)number of cohorts measuring fidelity with the IPS-15 scale. 1. The cohorts are ranked by fidelity score; the cohort with the highest score (Lucca, 2004, fidelity score=74) is at the top of the forest plot and the cohort with the lowest fidelity score (Burns et al., 2006 Rimini Control (RC), fidelity score=24) is at the bottom of the plot. Two horizontal axes are aligned to show how log-odds ratios correspond to employment commencements. A value of zero on the log-odds ratio corresponds to 50:50 odds of commencing employment (50% commenced employment).
Table 4. Precision analysis of the Individual Placement and Support 15-item fidelity scale (IPS-15)

<table>
<thead>
<tr>
<th>Fidelity score</th>
<th>44% or more</th>
<th>47% or more</th>
<th>62% or more</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Precision (%)</td>
<td>NPV (%)</td>
<td>Precision (%)</td>
</tr>
<tr>
<td>56 or more</td>
<td>58.1</td>
<td>100.0</td>
<td>53.5</td>
</tr>
<tr>
<td>61 or more</td>
<td>62.5</td>
<td>100.0</td>
<td>57.5</td>
</tr>
<tr>
<td>66 or more</td>
<td>70.8</td>
<td>71.4</td>
<td>62.5</td>
</tr>
</tbody>
</table>

Table 5. Precision analysis of the Individual Placement and Support 25-item fidelity scale (IPS-25)

<table>
<thead>
<tr>
<th>Fidelity score</th>
<th>44% or more</th>
<th>47% or more</th>
<th>62% or more</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Precision (%)</td>
<td>NPV (%)</td>
<td>Precision (%)</td>
</tr>
<tr>
<td>100 or more</td>
<td>40.0</td>
<td>85.7</td>
<td>40.0</td>
</tr>
<tr>
<td>115 or more</td>
<td>0.0</td>
<td>68.8</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Notes. 1. Only Precision or positive prediction, and Negative prediction or Negative predictive value (NPV) are reported. This approach was preferred over the alternative, a Sensitivity and Specificity analysis because cohorts with low fidelity scores are under-represented in this sample. This is because control group cohorts that did not have a fidelity score were excluded from this sample and these were most likely to have low fidelity scores. Precision is important because it is the probability that high-fidelity programs also have good employment outcomes. The formula for this is: TP/(TP+FP). Negative predictive value (NPV) is the probability that low fidelity programs have poor employment outcomes. The formula for this is: TN/(TN+FN). True positives (TP) represented high fidelity scores and good employment outcomes. False positives (FP) represented high fidelity scores and low employment outcomes. True negatives (TN) represented low fidelity scores and low employment outcomes. False negatives (FN) represented low fidelity scores and good employment outcomes. 2. The threshold of 44% or more commencing employment was set empirically by visual inspection of the scatter plot (see Figure 2). 3. The threshold of 47.3% or more commencing employment was suggested from previous studies, from the mean proportion commencing employment in USA cohorts (Bond et al., 2012b). 4. The threshold of 62.1% or more commencing employment was suggested by the mean proportion commencing employment in USA cohorts (Bond et al., 2012b). 5. The IPS-15 threshold fidelity score of 56 or more was selected from the scale authors’ definition of Fair fidelity. 6. The IPS-15 threshold fidelity score of 61 or more was based on visual inspection of the scatter plot (see Figure 3). 7. The IPS-15 fidelity threshold of 66 or more was selected from the scale authors’ definition of Good fidelity. 8. The IPS-25 threshold of 100 or more was based on visual inspection of the scatter plot (see Figure 3). 9. The IPS-25 threshold of 115 or more was selected from the scale authors’ definition of Exemplary fidelity.
Table 6. *Multivariate analysis of correlates of commencing competitive employment*¹

<table>
<thead>
<tr>
<th>Cohort characteristics</th>
<th>Did not commence employment</th>
<th>Commenced employment</th>
<th>Mean fidelity scores</th>
<th>$\chi^2$</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total fidelity score (c=51)</strong></td>
<td>3678</td>
<td>51.1</td>
<td>3520</td>
<td>48.9</td>
<td>61.2</td>
<td>15.31</td>
</tr>
<tr>
<td><strong>Study type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCT (c=37)</td>
<td>1388</td>
<td>53.9</td>
<td>1185</td>
<td>46.1</td>
<td>59.7</td>
<td>0.08</td>
</tr>
<tr>
<td>Observational study (c=14) (Ref)</td>
<td>2290</td>
<td>49.5</td>
<td>2335</td>
<td>50.5</td>
<td>65.1</td>
<td></td>
</tr>
<tr>
<td><strong>Region of study</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USA and Canada (c=17)</td>
<td>2758</td>
<td>49.7</td>
<td>2793</td>
<td>50.3</td>
<td>65.1</td>
<td>0.47</td>
</tr>
<tr>
<td>Europe (c=16)</td>
<td>441</td>
<td>70.6</td>
<td>184</td>
<td>29.4</td>
<td>51.5</td>
<td></td>
</tr>
<tr>
<td>Australasia (c=18) (Ref)</td>
<td>479</td>
<td>46.9</td>
<td>543</td>
<td>53.1</td>
<td>66.2</td>
<td></td>
</tr>
<tr>
<td><strong>Author group by first author</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond (c=1) (Ref)</td>
<td>27</td>
<td>28.1</td>
<td>69</td>
<td>71.9</td>
<td>72.0</td>
<td>35.01</td>
</tr>
<tr>
<td>Browne (c=2)</td>
<td>60</td>
<td>34.5</td>
<td>114</td>
<td>65.5</td>
<td>65.5</td>
<td></td>
</tr>
<tr>
<td>Burns (c=12)</td>
<td>184</td>
<td>59.0</td>
<td>128</td>
<td>41.0</td>
<td>48.3</td>
<td></td>
</tr>
<tr>
<td>Davis (c=2)</td>
<td>41</td>
<td>48.2</td>
<td>44</td>
<td>51.8</td>
<td>50.5</td>
<td></td>
</tr>
<tr>
<td>Drake (c=1)</td>
<td>31</td>
<td>40.8</td>
<td>45</td>
<td>59.2</td>
<td>70.0</td>
<td></td>
</tr>
<tr>
<td>Frey Drake (c=1)</td>
<td>595</td>
<td>53.1</td>
<td>526</td>
<td>46.9</td>
<td>67.0</td>
<td></td>
</tr>
<tr>
<td>Gold (c=1)</td>
<td>24</td>
<td>36.4</td>
<td>42</td>
<td>63.6</td>
<td>72.0</td>
<td></td>
</tr>
<tr>
<td>Henry (c=1)</td>
<td>1698</td>
<td>48.9</td>
<td>1776</td>
<td>51.1</td>
<td>63.0</td>
<td></td>
</tr>
<tr>
<td>Killackey (c=1)</td>
<td>7</td>
<td>35.0</td>
<td>13</td>
<td>65.0</td>
<td>68.0</td>
<td></td>
</tr>
<tr>
<td>Latimer (c=1)</td>
<td>40</td>
<td>53.3</td>
<td>35</td>
<td>46.7</td>
<td>72.0</td>
<td></td>
</tr>
<tr>
<td>Lehman (c=1)</td>
<td>82</td>
<td>72.6</td>
<td>31</td>
<td>27.4</td>
<td>70.0</td>
<td></td>
</tr>
<tr>
<td>Lucca (c=1)</td>
<td>16</td>
<td>17.8</td>
<td>74</td>
<td>82.2</td>
<td>74.0</td>
<td></td>
</tr>
<tr>
<td>McMurc (c=4)</td>
<td>82</td>
<td>54.3</td>
<td>69</td>
<td>45.7</td>
<td>66.5</td>
<td></td>
</tr>
<tr>
<td>Mueser (c=3)</td>
<td>122</td>
<td>59.8</td>
<td>82</td>
<td>40.2</td>
<td>59.7</td>
<td></td>
</tr>
<tr>
<td>Porteous (c=2)</td>
<td>131</td>
<td>58.2</td>
<td>94</td>
<td>41.8</td>
<td>71.0</td>
<td></td>
</tr>
<tr>
<td>Van Erp (c=4)</td>
<td>257</td>
<td>82.1</td>
<td>56</td>
<td>17.9</td>
<td>61.0</td>
<td></td>
</tr>
<tr>
<td>Waghorn (c=12)</td>
<td>267</td>
<td>47.9</td>
<td>290</td>
<td>52.1</td>
<td>65.2</td>
<td></td>
</tr>
<tr>
<td>Weng (c=1)</td>
<td>14</td>
<td>30.4</td>
<td>32</td>
<td>69.6</td>
<td>69.0</td>
<td></td>
</tr>
</tbody>
</table>

**Participants with psychotic disorders**³

<table>
<thead>
<tr>
<th></th>
<th>Did not commence employment</th>
<th>Commenced employment</th>
<th>Mean fidelity scores</th>
<th>$\chi^2$</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>50% or more (c=42)</td>
<td>2839</td>
<td>52.4</td>
<td>2575</td>
<td>47.6</td>
<td>60.8</td>
<td>0.88</td>
</tr>
<tr>
<td>Less than 50% (c=9) (Ref)</td>
<td>839</td>
<td>47.0</td>
<td>945</td>
<td>53.0</td>
<td>63.3</td>
<td></td>
</tr>
</tbody>
</table>

**Participants’ employment goals**

<table>
<thead>
<tr>
<th></th>
<th>Did not commence employment</th>
<th>Commenced employment</th>
<th>Mean fidelity scores</th>
<th>$\chi^2$</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Competitive employment goal (c=40)</td>
<td>991</td>
<td>61.4</td>
<td>624</td>
<td>38.6</td>
<td>59.3</td>
<td>3.30</td>
</tr>
<tr>
<td>Mixed vocational goals⁴ (c=11) (Ref)</td>
<td>2687</td>
<td>48.1</td>
<td>2896</td>
<td>51.9</td>
<td>68.3</td>
<td></td>
</tr>
</tbody>
</table>

**Study duration**

<table>
<thead>
<tr>
<th></th>
<th>Did not commence employment</th>
<th>Commenced employment</th>
<th>Mean fidelity scores</th>
<th>$\chi^2$</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>12 months or less (c=15)</td>
<td>1952</td>
<td>50.0</td>
<td>1954</td>
<td>50.0</td>
<td>63.6</td>
<td>0.34</td>
</tr>
<tr>
<td>More than 12 months (c=36) (Ref)</td>
<td>1726</td>
<td>52.4</td>
<td>1566</td>
<td>47.6</td>
<td>60.2</td>
<td></td>
</tr>
</tbody>
</table>

¡Notes. c=number of cohorts in the analysis of deviance. Ref=Reference group. 1. Analysis of deviance using random effects model (Harvard & Lau, 1993). Fifty-one cohorts were included in the analysis of deviance. One cohort (Oshima et al., 2014) was excluded because the proportion of participants with a psychotic disorder could not be determined. 2. Row percents add to 100% within rows. 3. Psychotic disorders included Schizoaffective disorder, Schizophrenia and First episode psychosis. Non-psychotic disorders included Bipolar affective disorder, Major depression, Post-traumatic stress disorder, and Other anxiety disorders. 4. Mixed vocational goals included those with education goals, voluntary work goals, or no employment goals. χ²=chi-square. df=degrees of freedom. p=p-value. p-values less than 0.05 are considered statistically significant.

This article is © Emerald Publishing Limited and permission has been granted for this version to appear here https://researchspace.auckland.ac.nz/handle/2292/2. Emerald does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Emerald Publishing Limited.
Discussion

This systematic review and meta-analysis examined the predictive validity of two measures of fidelity to the IPS approach to supported employment for people with severe mental illnesses. In total, 30 studies met the essential inclusion criteria. All 30 followed a defined cohort of eligible participants, who were exposed to a defined vocational intervention, for at least six months. The dependent variable was common to all included studies and was defined as the frequency of participants commencing competitive employment, relative to the complement, all other participants in the defined cohort were not known to or did not commence employment at any time during the follow-up period.

A positive relationship emerged between IPS-15 program fidelity score and the proportion commencing employment. Fidelity score on the IPS-15 was the strongest predictor of the proportion commencing employment, apart from the unique author group, which cannot at this time be investigated further without a larger sample of studies reporting additional characteristics not examined here. Fidelity score on the IPS-15 explained 16.8% of the variance in employment commencements. This is at the conservative end of the range of variance explained previously by the IPS-15 (12-58%, Bond et al., 2011), but just above the mean value of 15% (Bond et al., 2011). Since this review and meta-analysis used data on defined cohorts followed for at least six months, we expected some divergence from previous reviews using point-in-time measures of employment status. Nevertheless, these results are consistent with and support previous findings that the IPS-15 has moderate, yet important, validity for predicting employment commencements.

The precision analysis examined several IPS-15 fidelity thresholds and found that both precision and negative prediction are optimal when good fidelity is defined as 61 or more on the IPS-15, and when good employment commencements are defined as 44% or more of the cohort commencing employment (see Figure 3 and Table 4). The new threshold of 61 on the IPS-15 updates previous recommendations from the scale authors for fair fidelity (56-65) to be attained. The new benchmark provides a more precise goal for program implementers. Lower IPS-15 fidelity scores of
56-60, in the fair range, can no longer be considered sufficient for expecting good employment commencements. Whereas fidelity scores of 66 improve prospects of good employment commencements over an IPS-15 score of 61-65 they do not ensure good employment commencements will be attained.

It is possible that higher fidelity scores have a positive impact on other aspects of program performance, such as longer job tenure or reduced time to the first job. To understand this further, an analysis is needed which maps cohort fidelity scores to multiple employment outcome variables. Such an investigation was not possible with this data set but may be possible in the future when enough suitable studies are available for analysis.

High fidelity scores on the IPS-15 scale (61 or more) predicted good employment outcomes of 44% or more in 63% of cohorts. Increasing the fidelity threshold to 66 or more increased the prediction of good employment commencements among 71% of the cohorts. This means that IPS implementations achieving good fidelity or higher as defined by the higher threshold (66 or more) remains insufficient to ensure attainment of good employment commencements. Once good or higher fidelity is reached, ongoing program evaluation appears to be needed to check whether the expected good performance, as measured by the proportion of participants commencing employment, is actually being attained or not. Supplementing fidelity assessment with the form of cohort-based evaluation reported by these studies does not seem to be needed until this fidelity threshold is reached.

It is also important to note that the threshold for the proportion commencing competitive employment was defined in this analysis as 44% or more. This differs from the benchmarks of 33, 45, 57% previously published by Becker, Drake, and Bond (2011a). However, these benchmarks are not comparable because Becker et al. (2011a) used a different measure of employment status. Their employment variable captured percentages employed from quarterly progress reports where both the cohort and the proportion employed could vary, whereas our threshold of 44% is based on a defined cohort (fixed denominator) of participants, followed for at least six months. In our opinion, the cohort-
based measure is less prone to bias because it takes into account new program commencements (these are held constant), and program attrition because all dropouts are retained as part of the denominator. Program leaders could either define cohorts for prospective follow-up or consider ways to supplement quarterly monitoring of employment outcomes. One way would be to calculate rolling averages of new program commencements that obtained employment in the past six or 12 months, and record rolling averages for attrition over the same period.

Limitations

The smaller sample of IPS-25 studies limited the analysis that could be conducted. However, a similar pattern emerged from the precision analysis suggesting that the predictive validity of the IPS-25 is likely to be quite similar to the IPS-15. Although limited, the results for the IPS-25 support previous predictive validity findings (Bond et al., 2012c; Kim et al., 2015) and suggest that the IPS-25 may not have improved predictive validity compared to the IPS-15. If so, there may be no clear advantages in using the IPS-25 particularly since the IPS-15 is a shorter scale and less costly to administer. Further investigation of the predictive validity of the IPS-25 is urgently needed. The increasing use of the IPS-25 since 2008 also means that larger samples may soon be available for a similar meta-analysis.

The use of a single employment outcome variable is a limitation. This meta-analysis could only use one measure of program effectiveness, namely, the proportion commencing competitive employment during the study period. This is the measure most frequently used by studies of IPS implementations, and no other outcome measure was more widely or commonly reported. Alternatives include: time to first job, duration of longest job, duration of all jobs in weeks per year, hours worked, and earnings. These are sometimes reported by implementation studies, but not often enough at present for use in a meta-analysis.

Another limitation is that overall associations between the total fidelity score and the proportion commencing employment may obscure strong contributions of particular items or groups of items, in either direction. The scales also assume all items are weighted equally and provide no basis to infer
predictive strength at an item level. Future studies of IPS implementation could report item-level fidelity scores to enable researchers to explore predictive validity with respect to both groups of items and individual items. This knowledge would be helpful in terms of identifying the practices and subgroups of practices, most associated with particular employment outcomes.

Precision and negative prediction rather than sensitivity and specificity were the focus of this analysis because the search strategy was not designed to collect equal and representative samples of both high-fidelity and low-fidelity cohorts. For several reasons outlined previously, low-fidelity cohorts were under-represented compared to high-fidelity cohorts. The impact of this was greatest on specificity and sensitivity. Whereas precision and negative prediction, were less adversely affected. This differential influence can be seen by reconstructing the two cell by two cell contingency table from the dotted lines in Figure 3 and from the information in the footnotes to Table 5. In future meta-analyses, a more balanced sample of cohorts across the range of fidelity scores will enable all four characteristics of sensitivity, specificity, precision, and negative prediction, to be examined.

One cohort characteristic that could not be explored was the type and frequency of fidelity assessment. In some studies fidelity assessments were conducted by a research team independently of staff delivering the program. In others it was an external assessor trained by the team that developed the IPS program. Since fidelity assessment can involve synthesizing information from several sources over two days or more, and assessors can vary in experience and training, it would be useful to include the method of fidelity assessment when sufficient data permit. This could be important because inflation of fidelity score by internal assessment, or by using untrained assessors, could weaken the relationship between total fidelity score and employment commencements (Mowbray et al., 2003).

Frequency and timing of fidelity assessment can also be important. Some studies reported measuring fidelity regularly at six monthly intervals, whereas others reported one fidelity assessment at the end of the study period. Controlling the timing of fidelity assessments can be important because fidelity may be improving steadily over time and dated assessments may introduce noise into the
relationship between fidelity and employment outcomes. This is also an issue when providing feedback to practitioners because their interest in the results will be greater if both the performance measure and the fidelity assessment are based on the latest available information.

Another limitation concerned the availability of cohort-level information. Two large multi-site trials (Drake et al., 2013; Frey et al., 2011; Henry, Hashemi, & Zhang, 2014) did not provide site-level data about cohorts, so each was treated as a single cohort represented by an average fidelity score. In some RCTs, fidelity was not measured in the control group cohorts. This means these cohorts cannot contribute to future meta-analyses, and the nature of the control group intervention, in terms of similarity to the IPS approach, remains unknown.

The significance of the unique author group variable in this meta-analysis flags the presence of other influences on the relationship between fidelity score and employment commencements that could not be examined in this analysis. Candidates for other influences have been identified previously. For instance, low attrition in a moderately effective program can lead to better outcomes overall compared to a more effective program with high attrition (Morris, Waghorn, Robson, Moore, & Edwards, 2014). There was insufficient information reported by the included studies about attrition for this to be included in the analysis. Another potentially important influence is employment specialist expertise (Corbière et al., 2013). This can sometimes explain high outcomes for a cohort exposed to only a fair fidelity program. A possible example of this can be seen in Morris et al. (2014) where at one site (Peel), the employment specialist had the lowest fidelity (78/125) on the IPS-25 scale yet had unusually low attrition (0%) and the highest employment commencements (72.0%).

Country factors, such as health services, welfare policies, and labour market characteristics may also influence the employment outcomes attained by IPS programs. In this analysis, one multi-site study conducted in Australia (Waghorn, Dias, Gladman, Harris, & Saha, 2014) provided eight cohorts (four IPS cohorts and four control cohorts) where the effects of these country-specific factors were controlled by holding these constant. Not surprisingly, with these variables held constant, the
proportion of variance in employment commencements explained by IPS-15 fidelity score reached 72.3% ($p=.008$).

The significance of the author group variable shows that there are other unidentified cohort differences that explain variation in employment commencements other than the ones investigated here. Yet this systematic review and meta-analysis also shows how after controlling fidelity in six cohort characteristics, the IPS-15 total fidelity score remained a robust predictor of commencing employment across diverse international and research contexts. Further studies are now needed to explore this link at an item level, and with respect to several other candidate employment outcome variables.

**Conclusions**

The multivariate analysis shows the IPS-15 scale has moderate, yet important, predictive validity across diverse international research contexts. Low fidelity on both scales accurately predicted low-employment commencements but fidelity above the thresholds (61/75 in IPS-15, and 100/125 in IPS-25) did not accurately predict good employment commencements. Therefore, attaining these fidelity thresholds appears necessary but not sufficient for ensuring 44% or more of participants commence employment. On the other hand, not attaining the fidelity thresholds of 61 or more seems to guarantee failure represented by less than 44% of participants commencing competitive employment. Once this fidelity threshold is attained, cohort-based program evaluation can help determine whether expected performance is being achieved or not. Cohort-based evaluation does not appear needed until these fidelity thresholds are reached. Once more studies are available this analysis can be extended to the IPS-25 using a similar approach. Future research can also investigate the potential value of cohort-based evaluation, with regular feedback to service providers, for strengthening the translation of evidence-based practices into good employment outcomes.
Supplementary Analysis

Supplementary analysis was conducted to identify other factors that could help to explain differences between the cohorts that had high fidelity and high performance (hi-hi cohorts) and the cohorts that had high fidelity but low performance (hi-low cohorts). There were 25 hi-hi cohorts, each had a score of 61/75 or more on IPS-15 and each achieved 44% or more participants commencing competitive employment. There were 15 hi-low cohorts, each had a score of 61/75 or more on IPS-15 but were classified as hi-low because each had 43% or less participants commencing competitive employment. The supplementary analysis addressed two questions.

1. How do the hi-hi cohorts differ from the hi-low cohorts with the same total fidelity score on IPS-15? Can any of the available covariates account for the differences?
2. Are there other factors identified in the included research studies, but not identified in the meta-analysis, that could account for the poor performance of the high-fidelity, low performing (hi-low) cohorts?

Methods

The first question was addressed in three steps. Firstly, descriptive statistics were examined for each of two groups to compare the groups in relation to the independent and dependent variables, as well as the covariates. Chi-square tests of association were conducted for categorical variables, and Student T-tests for continuous variables. Next a scatter plot visually represented the hi-hi cohorts (n=25) and the hi-low cohorts (n=15) by total fidelity score and by the log-odds ratio of commencing employment. Thirdly, the multivariate analysis of deviance was re-run for the subgroups of hi-hi and hi-low cohorts (n=40).

To address the second question, a qualitative analysis was undertaken of the hi-low cohorts measuring fidelity with IPS-15 (n=15) and the cohorts measuring fidelity with IPS-25 (n=6). This analysis aimed to identify possible candidate variables for explaining the unexplained variance. This involved re-examining each study with a hi-low cohort to identify data not extracted for the meta-analysis which could help to further explain the low performance.
Results

Quantitative subgroup analysis

The characteristics of the two groups (hi-hi and hi-low) classified by each co-variate and the dependent and independent variables, are shown in Table 7. The two groups did not differ significantly on any of the covariates. The difference in participants’ primary goal, and the difference in the diagnostic mix approached, but did not reach, the 95% confidence level. As expected, there was a significant difference between the two groups on the dependent variable, the proportion of participants that commenced competitive employment during the study.

Table 7. A comparison of the hi-hi and the hi-low cohort (IPS-15)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Hi-hi (n=25)</th>
<th>Hi-low (n=15)</th>
<th>Test of significance^1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td>USA, n (%)</td>
<td>10 (40.0)</td>
<td>4 (26.7)</td>
</tr>
<tr>
<td>Study type</td>
<td>RCT, n (%)</td>
<td>19 (76.0)</td>
<td>9 (60.0)</td>
</tr>
<tr>
<td>Participants’ primary goal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Competitive employment, n (%)</td>
<td></td>
<td>15 (60.0)</td>
<td>13 (86.7)</td>
</tr>
<tr>
<td>Study duration</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Months of follow-up, mean (SD)</td>
<td></td>
<td>20.3 (11.3)</td>
<td>18.7 (5.9)</td>
</tr>
<tr>
<td>Diagnostic mix</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% of cohort with a diagnosis of psychosis, mean (SD)</td>
<td></td>
<td>63.7 (28.1)</td>
<td>77.4 (12.6)</td>
</tr>
<tr>
<td>Total fidelity score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fidelity score IPS-15, mean (SD)</td>
<td></td>
<td>67.6 (3.9)</td>
<td>65.9 (3.2)</td>
</tr>
<tr>
<td>Program performance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion commenced CE, mean (SD)</td>
<td></td>
<td>60.8 (12.4)</td>
<td>27.5 (10.3)</td>
</tr>
</tbody>
</table>

Notes. Hi-hi cohorts=cohorts with high fidelity score and high performance. Hi-low cohorts=cohorts with high fidelity score and low performance. There were 25 hi-hi cohorts, each had a score of 61/75 or more on IPS-15 and achieved 44% or more participants commencing competitive employment. There were 15 hi-low cohorts, each had a score of 61/75 or more on IPS-15 but had 43% or less participants commencing competitive employment. IPS-15=The 15-item IPS fidelity scale. n=number of cohorts. SD=standard deviation. CE=competitive employment. p-values less than 0.05 are considered statistically significant. 1. T-tests were unpaired, two-tailed.
Examination of the scatter plot (see Figure 5), shows that for any given value of fidelity score, all hi-hi cohorts exceeded all hi-low cohorts in terms of the performance in log-odds ratios. Although there was a consistent separation of groups, fewer hi-low sites and more high-high sites had very high fidelity scores. Multivariate analysis of deviance using log-odds ratios enabled the 40 cohorts to be adjusted for fidelity score. They were also adjusted for the same six co-variates extracted in the meta-analysis. 1. Study type. 2. Region of study. 3. Author group. 4. The proportion of participants with a diagnosis of psychosis. 5. Participants’ employment goals. 6. Duration of study. The analysis of deviance confirms that the only effect accounting for any significant difference is being classified as a hi-hi cohort. No other covariates account for the difference in performance between groups (see Table 8).

Table 8. Multivariate analysis of the correlates of commencing employment within the high-fidelity cohorts (n=40)

<table>
<thead>
<tr>
<th>Cohort characteristic</th>
<th>x</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hi-hi category</td>
<td>8.23</td>
<td>1</td>
<td>0.004</td>
</tr>
<tr>
<td>Total fidelity score</td>
<td>0.40</td>
<td>1</td>
<td>0.53</td>
</tr>
<tr>
<td>Study type</td>
<td>0.08</td>
<td>1</td>
<td>0.78</td>
</tr>
<tr>
<td>Region of study</td>
<td>0.74</td>
<td>2</td>
<td>0.69</td>
</tr>
<tr>
<td>Author group</td>
<td>9.93</td>
<td>15</td>
<td>0.82</td>
</tr>
<tr>
<td>Proportion with a psychosis diagnosis</td>
<td>0.43</td>
<td>1</td>
<td>0.51</td>
</tr>
<tr>
<td>Study duration</td>
<td>0.12</td>
<td>1</td>
<td>0.73</td>
</tr>
<tr>
<td>Participants’ employment goals</td>
<td>0.32</td>
<td>1</td>
<td>0.57</td>
</tr>
</tbody>
</table>

Notes. n=the number of cohorts in the analysis of deviance. p-values less than 0.05 are considered statistically significant at the 95% confidence level.
Figure 5. The performance of the high performing cohorts compared with the low performing cohorts of the same fidelity score (IPS-15)

Notes. IPS-15=The Individual Placement and Support 15-item fidelity scale. There were 25 high performing cohorts, each had a score of 61/75 or more on IPS-15 and achieved 44% or more participants commencing competitive employment. There were 15 low performing cohorts, each had a score of 61/75 or more on IPS-15 but had 43% or less participants commencing competitive employment.

Qualitative analysis of the studies with hi-low cohorts.

The results from the qualitative analysis of the studies with the hi-low cohorts are outlined in Table 9. Nine possible candidates for explaining low performance were identified. 1. Individual characteristics of the cohort of participants e.g. nature of vocational goal (education versus immediate employment), and offending history. 2. The absence of technical assistance or on-site leadership. 3. A low score, 3 or less out of 5, on the fidelity item: Integration with the clinical team. 4. The presence of
non-evidence-based practices. 5. Restrictions imposed by contracts, policy or labour market regulations. 6. Low frequency of contact between the employment specialist and each program participant. 7. The skills of vocational staff. 8. High study or intervention attrition. 9. The cohort was the control group in an RCT which was investigating an enhancement to IPS.

Chapter Conclusions

This chapter presented a peer-reviewed report that investigated the association between implementation fidelity, measured through one of two validated fidelity scales, and program performance, measured by the proportion of participants in the cohort commencing competitive employment. Fidelity to original program principles and practices accounted for a 16.8% of the variance in program performance. This means that the majority of the variance in performance remains unexplained. The implication is that an opportunity exists to further improve the effectiveness of evidence-based practices by finding other factors that account for the unexplained variance. Some candidates for potential explanatory variables emerged from a qualitative analysis of the research reports included in the meta-analysis. However, a broader and more systematic approach to identifying and prioritising potential causal influences is needed. Chapter 4 attempts this by examining the wider theoretical and empirical literature on psychosocial program implementation to begin this search for candidates for the unexplained variance in program performance.
Table 9. Qualitative analysis of the hi-low cohorts\(^\d\) (IPS-15, \(n=15\); IPS-25, \(n=6\))

<table>
<thead>
<tr>
<th>Authors, date</th>
<th>Description of study and the hi-low cohort(s)</th>
<th>Cohort identifier(^2)</th>
<th>Cohort size</th>
<th>Total fidelity score</th>
<th>Proportion commenced CE (%)</th>
<th>Possible influences on program performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>McGurk et al., 2015</td>
<td>The hi-low cohort was the control arm of an RCT comparing IPS enhanced with specialised cognitive training for ES (the control arm) with IPS plus a cognitive enhancement program 'Thinking Skills for Work' (the intervention arm).</td>
<td>Cohort A</td>
<td>50</td>
<td>70</td>
<td>36.0</td>
<td>The hi-low cohort was in the control arm of an IPS enhancement trial. The population of participants in the study were people for whom SE had not previously been effective, defined as the participant had not commenced CE or they had not managed to hold down a competitive job (i.e. left the job within three months).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cohort B</td>
<td>28</td>
<td>67</td>
<td>37.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cohort C</td>
<td>34</td>
<td>63</td>
<td>12.0</td>
<td></td>
</tr>
<tr>
<td>Williams et al., 2015</td>
<td>A prospective study examining the employment outcomes of three sites implementing IPS services. Participants in one site all had a diagnosis of first episode psychosis. Participants all had a goal of competitive employment. All three sites were hi-low cohorts.</td>
<td>Cohort A</td>
<td>52</td>
<td>64</td>
<td>33.3</td>
<td>It is possible that terms in the contract affected implementation, the sites were contracted to provide generic disability employment services.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cohort B</td>
<td>28</td>
<td>67</td>
<td>37.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cohort C</td>
<td>34</td>
<td>63</td>
<td>12.0</td>
<td></td>
</tr>
<tr>
<td>Waghorn et al., 2014</td>
<td>A four-site RCT with IPS as the experimental intervention and a control intervention at each site. The control group used enhanced treatment as usual where the mental health team was assisted in making referrals to external disability employment services. Participants all had a goal of competitive employment. Four cohorts were hi-low cohorts.</td>
<td>PA IPS</td>
<td>27</td>
<td>67</td>
<td>40.7</td>
<td>Access criterion was that participants needed to be available to work more than eight hours a week. Clinicians identified study volunteers. There was high study attrition. The two control sites were located in the same employment service as the intervention site, as there were no other disability employment services in the area to refer to. The control site scored 3 on fidelity item: integration of services. The authors noted a lack of technical assistance and on-site leadership.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T IPS</td>
<td>17</td>
<td>69</td>
<td>29.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>T Con</td>
<td>18</td>
<td>65</td>
<td>33.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cairns Con</td>
<td>25</td>
<td>61</td>
<td>16.0</td>
<td></td>
</tr>
<tr>
<td>Porteous &amp; Waghorn, 2007</td>
<td>A prospective study of two cohorts of young people aged 14 to 27 with a diagnosis of first episode psychosis and a goal of any work (which could include voluntary work). One cohort a was hi-low cohort.</td>
<td>First cohort</td>
<td>100</td>
<td>71</td>
<td>36.0</td>
<td>Intervention attrition was high across all sites. Participants had educational goals as well as employment goals.</td>
</tr>
<tr>
<td>Authors, date</td>
<td>Description of study and the hi-low cohort(s)</td>
<td>Cohort identifier(^2)</td>
<td>Cohort size</td>
<td>Total fidelity score</td>
<td>Proportion commenced CE (%)</td>
<td>Possible influences on program performance</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------------------------------------</td>
<td>-------------------------</td>
<td>-------------</td>
<td>---------------------</td>
<td>----------------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Van Erp et al., 2007</td>
<td>A prospective multi-site study of the implementation of IPS in four regions of the Netherlands. Participants all had a goal of competitive employment. All cohorts in the study had low outcomes, but cohorts one and four also had low fidelity (60 and 54, out of 75).</td>
<td>Cohort 2</td>
<td>64</td>
<td>65</td>
<td>25.0</td>
<td>An ES was re-assigned from working in TVR to working in the mental health team, training in IPS was provided. One ES covered four clinical teams. Intervention attrition was high (38.6%) across all sites. High ES turnover. The authors reported a lack of project and program management. Social security system offers lots of protection to individuals, employers are also responsible to pay the first 12-months of sick leave. There were complex funding arrangements for the IPS programs. All four sites had low scores on fidelity items: finding permanent jobs and providing services in the community. The ES did not make face-to-face meetings with employers.</td>
</tr>
<tr>
<td>Burns et al., 2006, 2007</td>
<td>Multi-site RCT comparing IPS vs. TVR. Participants had a goal of competitive employment.</td>
<td>IPS site Groningen</td>
<td>26</td>
<td>67</td>
<td>34.6</td>
<td>The Groningen sites scored low on fidelity items: integration with mental health team and ongoing work assessment.</td>
</tr>
<tr>
<td>McGurk et al., 2005</td>
<td>RCT with IPS as the control vs. IPS with cognitive training. The inclusion criteria for participants included: history of at least one unsatisfactory job ending. Participants had a goal of competitive employment.</td>
<td>Control site (IPS) in RCT</td>
<td>21</td>
<td>63</td>
<td>4.8</td>
<td>The hi-low cohort was in the control arm of an IPS enhancement study. Participants were people for whom SE had not been effective or who had not managed to hold down a competitive job (i.e. left within three months). Evidence of prior job failure was also part of inclusion criteria.</td>
</tr>
<tr>
<td>Mueser et al., 2004</td>
<td>An RCT with three cohorts which compared IPS with psychosocial treatment and with non-integrated supported employment services. The study was conducted in an inner city. Participants were African American or Latino and had a goal of competitive employment.</td>
<td>Standard vocational service in RCT</td>
<td>69</td>
<td>61</td>
<td>27.5</td>
<td>Intervention attrition was much higher for the hi-low cohort compared to the other two cohorts in this study. This cohort were in a standard vocational service that would not have been integrated with mental health services. There were no Spanish speaking staff in the program serving the hi-low cohort, so it had limited capacity to serve Latino participants.</td>
</tr>
<tr>
<td>Authors, date</td>
<td>Description of study and the hi-low cohort(s)</td>
<td>Cohort identifier</td>
<td>Cohort size</td>
<td>Total fidelity score</td>
<td>Proportion commenced CE (%)</td>
<td>Possible influences on program performance</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------------------------</td>
<td>------------------</td>
<td>-------------</td>
<td>----------------------</td>
<td>-----------------------------</td>
<td>--------------------------------------------</td>
</tr>
<tr>
<td>Lehman et al., 2002</td>
<td>RCT comparing IPS vs. step-wise vocational rehabilitation services.</td>
<td>IPS site in an RCT</td>
<td>113</td>
<td>70</td>
<td>27.4</td>
<td>Participants did not need to have a desire to work and did not have to attend two research interviews, although participants were paid for attending interviews. Large numbers of participants had co-existing substance use disorders and this subgroup had poor employment outcomes. The sites supported participants into non-competitive jobs as well.</td>
</tr>
<tr>
<td>IPS-25</td>
<td>Bond et al., 2015</td>
<td>IPS group</td>
<td>43</td>
<td>110</td>
<td>30.3</td>
<td>All participants had offending histories. Participants reported having more pressing needs than employment i.e. housing, finance issues.</td>
</tr>
<tr>
<td></td>
<td>Mcgurk et al., 2015</td>
<td>IPS enhanced control group</td>
<td>50</td>
<td>104</td>
<td>36.0</td>
<td>The hi-low cohort was the control arm of an IPS enhancement trial. Participants had failed to benefit previously from SE.</td>
</tr>
<tr>
<td>Craig et al., 2014</td>
<td>RCT with IPS as the control vs. IPS plus motivational interviewing as the intervention arm. Participants were young people aged 18 to 35 with a diagnosis of first episode psychosis. Participants did not need to have a desire to work.</td>
<td>Midlands</td>
<td>42</td>
<td>111</td>
<td>38.1</td>
<td>No requirement that participants needed to have a desire to work or study. Participants had educational goals as well as employment goals.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>London</td>
<td>39</td>
<td>111</td>
<td>33.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mid Con</td>
<td>38</td>
<td>116</td>
<td>15.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lon Con</td>
<td>40</td>
<td>111</td>
<td>15.0</td>
<td></td>
</tr>
</tbody>
</table>

Notes. IPS=Individual Placement and Support. IPS-15=The 15-item IPS fidelity scale. IPS-25=The 25-item IPS fidelity scale. ES=Employment Specialist. CE=Competitive Employment. RCT=Randomised Controlled Trial. SE=Supported Employment. TVR=Traditional Vocational Rehabilitation. PA=Princes Alexandra Hospital Mental Health Service. T=Townsville Mental Health Service. Con=Control site. Mid=Midlands. Lon=London. 1. Hi-low cohorts achieved fidelity scores of 61 or more out of 75 on IPS-15, or 100 or more out of 125 on IPS-25, but less than 44% of participants in the cohort commenced competitive employment. 2. The cohort identifier is from the meta-analysis coding of cohorts used in the study by Lockett, Waghorn, Kydd, and Chant (2016).
Chapter 4: A New Implementation Framework for Improving the Effectiveness of Evidence-Based Practices

Prelude

Rationale for study 2: A review of the science of psychosocial program implementation

Chapter 3 investigated the relationship between implementation fidelity and program performance through a systematic review and meta-analysis of vocational rehabilitation implementation studies internationally. Low fidelity scores were, with one exception (Morris et al., 2014), associated with low performance. Of the 40 cohorts in vocational rehabilitation programs categorised as high fidelity, 37.5% also reported low performance. Some of these low performers also had very high fidelity scores (more than 90% of total fidelity score). This implies that program fidelity is necessary but not sufficient to expect acceptable program performance. Total scores on the IPS-15 scale only explained 16.8% of the variance in program performance. This means that whilst measuring program fidelity is important, program leaders cannot rely on fidelity score as a measure of good program outcomes. Program leaders should be regularly measuring program outcomes using individual-level data from defined cohorts, to check the performance of the program.

The findings reported in Chapter 3 also highlight the importance of attaining a total score of 80% or more on the current fidelity scales. Until this 80% fidelity threshold is attained, program leaders should focus on improving fidelity to evidence-based practices. Once the threshold is attained program leaders need to decide where to focus their resources to improve performance.

To improve the performance of evidence-based practices sources of unexplained variance need identifying. Supplementary analysis reported in Chapter 3, found that the co-variates in the analysis of deviance did not explain the difference between the performance of the high-fidelity, high performance cohorts and the high-fidelity, low performance cohorts. However, the qualitative analysis did identify some candidate influences worth further exploration as potentially sources for the remaining unexplained low performance of the high-fidelity cohorts.
Previous research has identified some possible candidates for improving program performance. These are: (a) employment specialists’ competencies; (b) participant characteristics, particularly motivation to work; (c) environmental factors e.g. societal, cultural, labour laws, health and disability policies, government regulations; (d) presence of non-evidence-based practices; (e) improving the integration between employment and clinical services; (f) enhancing employer engagement; and (g) greater involvement of family and clinicians (Corbière et al., 2017; Drake & Bond, 2011; Drebing et al., 2012; Kukla & Bond, 2012; Marshall et al., 2014). However there appears to be an absence of a logical framework to systematically prioritise and measure potential candidates which could explain the variation in program performance once good fidelity has been reached.

The study presented in this chapter aims to address this current research and implementation gap. The following research question guided a review of the conceptual and empirical literature: Should a wider theoretical implementation framework be applied to improve vocational rehabilitation for people with severe mental illnesses?

The following material on pages 65 to 85 is reprinted from the *Journal of Vocational Rehabilitation*, 49(1), 15-31. Lockett, H., Waghorn, G., & Kydd, R., A framework for improving the effectiveness of evidence-based practices in vocational rehabilitation. Copyright (2018), with permission from IOS Press”. The publication is available at IOS Press through doi:10.3233/JVR-180951. Permission received on 8 June 2018.

The chapter finishes with supplementary analysis which followed the published paper, including a description of the included studies in the first phase of the literature review, a change in the descriptor of one dimension of implementation quality, and the naming of the new framework to the Implementation Framework for Vocational Rehabilitation (IFVR), before concluding comments lead into Chapter 5.
Introduction

The science of vocational rehabilitation for people with severe mental illnesses has advanced significantly over the past thirty years and has reached a critical stage in the development of evidence-based practices. There is now a widely agreed set of principles and practices, known collectively as Individual Placement and Support (IPS) which has proven to be more effective than other approaches to vocational rehabilitation (Drake et al., 2012, p. 59). This IPS approach has been adopted and implemented across four continents and for different populations (Drake & Bond, 2014). These different populations have included young adults and older adults, people with a homelessness history, post-traumatic stress disorder, substance-use disorders, and involvement with the criminal justice system (Drake & Bond, 2014).

The evidence of IPS effectiveness includes 24 RCTs (Drake & Bond, 2017), 12 systematic reviews (Marshall et al., 2014) and a Cochrane review (Kinoshita et al., 2013). This body of evidence has consistently found IPS is more effective than alternative approaches to vocational rehabilitation in terms of the numbers of participants commencing competitive employment. A recent systematic review and meta-analysis of 17 RCTs comparing IPS to alternative types of vocational rehabilitation, found IPS increased the odds of participants commencing employment by more than two times (pooled risk ratio 2.40, 95% CI 1.99-2.90) (Modini et al., 2016). This review covered IPS implementation in 10 countries.

Despite this body of evidence there is still room for improvement (Drake et al., 2012, p. 4). For instance, a loss of effectiveness has been observed for IPS implementations outside the USA. This loss of effectiveness has been attributed to reduced adherence to the original principles and practices of the program (Bond et al., 2012b). Faithful adherence to the way a program was originally delivered in research trials, is referred to as program fidelity (Stith et al., 2006). Currently the main strategy advocated to improve outcomes from IPS program implementations is to increase fidelity as measured by the latest standardised fidelity scale (Drake et al., 2012, p. 112). However, evidence from a recent systematic review and meta-analysis suggests that this strategy may not be adequate to improve intended program outcomes. This is because good program fidelity is necessary, but not sufficient, for
good employment outcomes to be achieved (Lockett, Waghorn, Kydd, & Chant, 2016). Hence it is possible that other aspects of program implementation quality, other than fidelity, influence program performance. If these other aspects can be identified, and if causally linked to program outcomes, attention to these could further develop vocational rehabilitation practices for people with severe mental illnesses.

Identifying factors which influence the quality of program implementation is a problem that is not unique to vocational rehabilitation. A research to practice effectiveness gap has been noted in the literature concerned with improving the effectiveness of a wide range of psychosocial programs. In fact, the discipline of implementation science emerged in response to this widely reported type of problem (Nilsen, 2015). The premise underpinning implementation science is that improving program performance first requires improving the quality of program implementation (Durlak & DuPre, 2008). Durlak (2015) has argued that, “It is not evidence-based programs that are effective, but it is well-implemented evidence-based programs that are effective” (p. 1124). This implies the need for a measure of program implementation quality, coupled with an understanding of the wider ecological factors which influence the quality of implementation in a particular context (Damschroder et al., 2009). Such a wider conceptual framework would assist researchers, practitioners, and policymakers to systematically map factors outside fidelity that might influence program outcomes (Menear & Briand, 2014). It could also help to identify which implementation factors appear to be influential across different settings, or ones that are unique to cultures or contexts (Roe, 2017).

The purpose of this literature review and evidence synthesis was to construct a broader conceptual framework for improving the implementation of evidence-based practices in vocational rehabilitation for people with severe mental illnesses. This involved examining the broader implementation science literature to identify factors besides fidelity, influencing the quality of health and social science program implementation. The intention was to highlight the most promising of these as potential candidates for enhancing the effectiveness of psychiatric vocational rehabilitation programs.
Methods

The literature review and evidence synthesis were conducted in three stages. The first stage aimed to identify the common dimensions of implementation quality and contextual factors known to influence program performance, in order to develop a generic conceptual framework which could then be applied to vocational rehabilitation. In stage two, the literature was revisited, along with literature on IPS implementation, to explore the strength of the evidence for all the identified dimensions of implementation and contextual factors. The third stage was to use the synthesised evidence to construct a conceptual implementation framework for vocational rehabilitation, for people with severe mental illnesses in particular.

Literature review: stage one

In stage one a broad approach was taken exploring the implementation science literature from across the health and social sciences. To manage the volume, papers were included in the search if they were published in English in a peer reviewed journal between 1997 and 2016.

The first author undertook searches of Medline, PubMed, PsycINFO and the Cochrane library, using the following search terms combined with the Boolean operators ‘OR’ and ‘AND’; implement*, evidence-based practice*, evidence-based treatment, evidence-based medicine, vocational rehabilitation, mental illness, mental health, review. Reference lists of selected articles were also searched for other articles which appeared relevant to the review. In addition, contents pages of two key journals were searched: the Journal of Vocational Rehabilitation, and Implementation Science.

Papers of any study design were permitted (e.g. cross-sectional survey, longitudinal studies, RCTs) and were not restricted by quantitative or qualitative emphasis. The focus of this first stage was primarily on systematic reviews or meta-analyses. Single studies were included if particularly relevant, or if they were missed by reviews. The literature search was considered to be complete when data saturation was reached, and additional publications did not contribute new information to the dimensions and contextual factors identified.

This first stage of the review aimed to broaden out the scope of existing implementation frameworks for vocational rehabilitation. The included papers were classified as being conceptual, or
based on empirical evidence, or both. The papers were examined for all the factors identified as influencing implementation quality and intervention performance. Factors were included if they were found at least twice in different studies, no other criteria were applied. The literature examined in this first stage identified a set of common dimensions of quality and contextual factors in a generic conceptual framework for program implementation.

**Literature review: stage two**

In stage two, the literature was revisited to explore the strength of the evidence for each of the identified dimensions of implementation and contextual factors. The focus in this second stage was on empirical papers identifying influences on program performance and the performance of evidence-based practices in vocational rehabilitation, for people with severe mental illnesses in particular. For this reason, published papers from IPS implementation studies were also examined for possible influences on program performance, outside those measured by the current fidelity scales. To be included in this second phase of the review, studies needed to report empirical evidence for a direct relationship between a dimension of quality, or contextual factor in the implementation framework, and a measure of program performance. Given the focus of this review to identify influences on program performance other than fidelity, studies which only examined the influence of fidelity were excluded.

Each of the included empirical papers was re-examined and the evidence in each paper mapped against one or more dimension of quality or contextual factor from the generic conceptual framework. Each paper was also rated according to its target relevance to vocational rehabilitation for people with severe mental illnesses. Papers from vocational rehabilitation or general psychiatry were rated A. Papers from the general health sector but outside psychiatry, were rated B. Papers outside the health sector were rated C. For example, a paper showing evidence of an association between staff experience and program performance in illness-prevention programs, was mapped against the contextual factor in the generic framework staff expertise, and given a relevance rating of B.

Mapping the evidence against the generic conceptual framework was an iterative process. It involved a gradual refinement and re-naming of some of the dimensions of quality and contextual
factors for clarity, and applicability to vocational rehabilitation. Re-naming was informed by established evidence-based practices in vocational rehabilitation.

Once this mapping of all the included papers was completed, each dimension of quality and each contextual factor was classified according to whether it was a direct or an indirect influence. Where there was no evidence, or there was only one study with evidence of a direct influence on program performance, the relationship was classified as indirect. Where evidence from two or more studies converged, this was classified as a direct influence.

**Literature review: stage three**

The third stage involved rating each dimension of quality for strength of evidence and feasibility to measure. This rating was based on the identified evidence, examples of measures of the dimension or contextual factor used in the empirical papers, and the first author’s judgements about how the phenomenon could be measured. The highest rating was two: indicating both good evidence of a direct influence on program performance, and the dimension was considered sufficiently operationalised to be feasible to measure. Good evidence was defined as three or more studies from the vocational rehabilitation or general psychiatry literature providing evidence of a direct influence. The next highest rating was one: limited evidence of a direct influence, but the dimension was considered sufficiently operationalised to be feasible to measure. Limited evidence was defined as less than three studies from the vocational rehabilitation or general psychiatry literature providing evidence of a direct influence. The lowest rating was zero: limited evidence for a direct influence, and the dimension appeared insufficiently operationalised and therefore difficult to measure.

Once rated, the direct influences were grouped together as new dimensions of quality, and the indirect influences were grouped as contextual factors in the conceptual framework for vocational rehabilitation.
Results

A generic conceptual implementation framework to enhance program performance

Twenty-four studies were retained in the first phase of the literature review, twenty reviews and four single studies (see Table 12 in the supplementary analysis). These studies represented published reports from the domains of public health, general health, health promotion and prevention, education, communications, and marketing. The studies included interventions focused on children, adolescents and adults. The largest review, Fixsen, Naoom, Blasé, Friedman, and Wallace (2005) covered 743 studies, of which 22 were empirical. The review covered multiple research areas including health, mental health, social services, justice, business, and engineering. The largest empirically-based review examined 581 studies on prevention and health promotion programs for children and adolescents (Durlak & DuPre 2008).

The concepts and constructs influencing implementation quality and intervention performance were extracted from each included review or study. Although different terminology was used across these studies, it was possible to identify a common set of influences on program performance and to categorise these influences. These were used to develop a generic conceptual implementation framework for improving program performance (see Table 10, column 1). This framework consists of the core dimensions of implementation quality including program fidelity measurement and four groups of contextual factors. 1. The socio-economic and political context. 2. The implementation process. 3. Characteristics of the participating organisations. 4. Characteristics of the participating individuals. The characteristics of the intervention were also identified in the reviewed literature as a common influence on implementation quality. This was not included in the framework because the focus of this review was on a single type of intervention, namely evidence-based vocational rehabilitation.

A conceptual framework to enhance performance of vocational rehabilitation programs

Sixteen of the included studies from phase one were excluded from the second review phase. This was because they focused on intermediate goals of improving program implementation, for example program fidelity, rather than on the end goal of improving program outcomes attained by
program participants. The eight remaining studies that examined direct influences on program outcomes were retained for examination in the second stage of the literature review. These were the reviews by Berkel, Mauricio, Schoenfelder, and Sandler, 2011; Dane and Schneider, 1998; Dorr et al., 2007, Durlak and DuPre, 2008; Fixsen et al., 2005; Kitson, Harvey, and McCormack, 1998; Schell et al., 2013; and Woltmann et al., 2012. In addition to these eight reviews, a further 19 studies (one review and 18 single studies) were identified as meeting the inclusion criterion (see Table 11).

Data from the included studies were synthesised to build a new conceptual framework for improving evidence-based vocational rehabilitation. This involved mapping the empirical evidence into the generic implementation framework to adapt and refine the dimensions of quality and the relevant contextual factors. For example, the dimension of quality, monitoring of comparison conditions in the generic framework, was further refined to quality of mental health treatment. Program differentiation was refined into the removal of non-evidence-based practices. The contextual factor described as government policy in the generic framework was divided into two policy forms: Health policy and social security policy. Each dimension of quality and each contextual factor were then classified as direct or indirect, depending on the evidence for their influence on vocational rehabilitation program performance (see Table 10).
Table 10. The common dimensions in a generic implementation framework mapped to vocational rehabilitation

<table>
<thead>
<tr>
<th>Generic implementation framework&lt;sup&gt;1&lt;/sup&gt;</th>
<th>Implementation framework for improving evidence-based vocational rehabilitation</th>
<th>Type of influence&lt;sup&gt;2&lt;/sup&gt;</th>
<th>Measurement rating&lt;sup&gt;3&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimensions of implementation quality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fidelity</td>
<td>Fidelity</td>
<td>Direct</td>
<td>**</td>
</tr>
<tr>
<td>Dosage</td>
<td>Program intensity</td>
<td>Direct</td>
<td>**</td>
</tr>
<tr>
<td>Quality of program delivery</td>
<td>Quality of program delivery</td>
<td>Direct</td>
<td>**</td>
</tr>
<tr>
<td>Participant responsiveness</td>
<td>Participant responsiveness</td>
<td>Direct</td>
<td>*</td>
</tr>
<tr>
<td>Program differentiation</td>
<td>Removal of non-evidence-based practices</td>
<td>Direct</td>
<td>*</td>
</tr>
<tr>
<td>Monitoring of comparison conditions</td>
<td>Quality of mental health treatment</td>
<td>Direct</td>
<td>*</td>
</tr>
<tr>
<td>Program reach</td>
<td>Program reach</td>
<td>Indirect</td>
<td></td>
</tr>
<tr>
<td>Program adaptation</td>
<td>Program adaptation</td>
<td>Indirect</td>
<td></td>
</tr>
<tr>
<td><strong>Contextual factors</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>The socio-economic and political context</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Government policy</td>
<td>Health policy &amp; social security policy</td>
<td>Indirect</td>
<td></td>
</tr>
<tr>
<td>Politics</td>
<td>Politics</td>
<td>Indirect</td>
<td></td>
</tr>
<tr>
<td>Economic</td>
<td>Local labour markets</td>
<td>Indirect</td>
<td></td>
</tr>
<tr>
<td>Funding levels, resources</td>
<td>Funding levels and resources</td>
<td>Indirect</td>
<td></td>
</tr>
<tr>
<td>Contract terms and conditions</td>
<td>Contract terms and conditions</td>
<td>Indirect</td>
<td></td>
</tr>
<tr>
<td>Legislation and regulation, incentives</td>
<td>Legislation and regulation</td>
<td>Indirect</td>
<td></td>
</tr>
<tr>
<td>and reward structures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Geography</td>
<td>Geography (rural/urban)</td>
<td>Indirect</td>
<td></td>
</tr>
<tr>
<td>The implementation process</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leadership and program champions</td>
<td>Leadership and program champions</td>
<td>Indirect</td>
<td></td>
</tr>
<tr>
<td>Training and education</td>
<td>Training and education</td>
<td>Indirect</td>
<td></td>
</tr>
<tr>
<td>Technical assistance and coaching</td>
<td>Technical assistance</td>
<td>Direct</td>
<td>*</td>
</tr>
<tr>
<td>Administrative support and information</td>
<td>Administrative support and information systems</td>
<td>Indirect</td>
<td></td>
</tr>
<tr>
<td>systems</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Program evaluation and feedback</td>
<td>Program evaluation and feedback</td>
<td>Direct</td>
<td>*</td>
</tr>
<tr>
<td>Communication levels and channels between stakeholders</td>
<td>The communication between stakeholders</td>
<td>Indirect</td>
<td></td>
</tr>
<tr>
<td>Management and supervision</td>
<td>Management and supervision</td>
<td>Indirect</td>
<td></td>
</tr>
<tr>
<td><strong>The characteristics of the organisations involved</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Culture and values</td>
<td>Culture and values of the employment service. Culture and values of the clinical service</td>
<td>Indirect</td>
<td></td>
</tr>
<tr>
<td>Readiness for innovation and approach to change</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strategies, workflows and practices</td>
<td>Strategies and practices of the clinical service. Strategies and practices of the employment service. Employer organisations human resources practices</td>
<td>Indirect</td>
<td></td>
</tr>
<tr>
<td><strong>The characteristics of the individuals involved</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Staff experience: education, knowledge, attitudes, skills, and self-efficacy.</td>
<td>Employment specialist expertise</td>
<td>Direct</td>
<td>**</td>
</tr>
<tr>
<td>Staff perceived need for the program: benefits of the program.</td>
<td>Clinicians’ experience, beliefs and attitudes. The attitudes and beliefs of line managers and work colleagues.</td>
<td>Indirect</td>
<td></td>
</tr>
<tr>
<td>Service users and families</td>
<td>Service users and families’ attitudes and beliefs.</td>
<td>Indirect</td>
<td></td>
</tr>
</tbody>
</table>

Notes. 1. Twenty reviews and four single studies identified the common dimensions of a generic implementation framework. 2. Each dimension of quality and each contextual factor was classified according to whether it was a direct or indirect influence on program performance. Where there was no evidence or there was only one study with evidence of a direct influence the relationship was classified as indirect. Where two studies or more cited evidence this was classified as a direct influence. 3. Each influence was rated based on the strength of evidence and feasibility to measure it. **=good evidence (three or more studies from the vocational rehabilitation or general psychiatry literature providing evidence of a direct influence on program performance), and the dimension was considered sufficiently operationalised to be feasible to measure; *=limited evidence (less than three studies from the vocational rehabilitation or general psychiatry literature providing evidence of a direct influence), but the dimension was considered sufficiently operationalised to be feasible to measure. No asterisk=limited evidence for a direct influence and the dimension appeared insufficiently operationalised and therefore difficult to measure.
Table 11. The empirical studies used to develop the conceptual framework for evidence-based practices in vocational rehabilitation (n=27)

<table>
<thead>
<tr>
<th>Included paper</th>
<th>Brief description of paper</th>
<th>Factors found to be directly influencing program performance</th>
<th>Empirical findings which were mapped to the conceptual framework</th>
<th>Relevance rating</th>
</tr>
</thead>
</table>
| Areberg et al., 2013 | A qualitative study interviewing 17 IPS program participants, 15 of whom got a job through the program. | • Individual needs and wishes valued by ES.  
 • Benefitting from the skills of ES and from being two people instead of one.  
 • Bringing the individual’s own qualifications to the process.  
 • Participants having a clear vision of the benefits for them through their participation.  
 • Having others who also believe (including the clinical team). | ES expertise  
The quality of program delivery  
Participant responsiveness  
Clinicians’ beliefs | A |
| Bejerholm et al., 2015 | A Swedish RCT of IPS versus TVR with 120 participants. | Welfare regulations required all participants to go through a period of internship and assessment prior to searching for competitive employment. | Removal of non-evidence-based practices  
Social security policy  
Quality of delivery  
Adaptation  
Participant responsiveness | A |
| Berkel et al., 2011 | A review of empirical papers which explored the association between dimensions of implementation and outcomes in prevention programs. | Fidelity, quality of delivery, adaptation and participant responsiveness. | Program adaptation | B |
| Boycott et al., 2012 | A rapid evidence assessment of interventions designed to supplement IPS. It included six empirical papers which looked at enhancing IPS with either skills training or cognitive remediation. | It was difficult to measure the contribution that the IPS adaptation was having on program performance. | The quality of program delivery | A |
| Catty et al., 2011 | Analysis of data from a six-site European RCT of IPS versus TVR to identify factors affecting outcomes. | • Patient and worker rated therapeutic relationship as measured by the Helping Alliance Scale.  
 • Patients’ previous work history. | Program evaluation and feedback | B |
| Coleman et al., 2009 | A review and meta-analysis of the chronic illness management literature included 56 empirical articles. The aim was to examine the association between practice re-design and outcomes. | Studies which focused on practice re-design, such as re-organising the way care was delivered, and which used data and guidelines to engage patients as well as focusing on modifying patient behaviours, were more likely to have better patient outcomes. |  |

73
<table>
<thead>
<tr>
<th>Included paper</th>
<th>Brief description of paper</th>
<th>Factors found to be directly influencing program performance</th>
<th>Empirical findings which were mapped to the conceptual framework</th>
<th>Relevance rating²</th>
</tr>
</thead>
</table>
| Corbière et al., 2014 | The behaviours, attitudes and competencies of 153 ES were measured to explore the associations with employment outcomes. | The BAKES scale was developed and tested. The subscales of 'relationships with employers and supervisors' and 'support and client-centred approach' were associated with employment commencements and six-month job tenure. | ES expertise  
Quality of program delivery i.e. relationships between the ES and employers                                                                                                                   | A                |
Quality of program delivery                                                                                                                                            | B                |
| Dorr et al., 2007     | Review of chronic illness literature to examine the association between information systems, quality and outcomes. Included 109 empirical articles, 23% of which had a focus on patients with mental illness. | Factors correlated with positive outcomes were: electronic medical record, computerised prompts, reports and feedback, decision-support aids, electronic scheduling and personal health records. Those factors which were barriers to outcomes were: costs, data privacy, security concerns and a lack of focus on workflows. | Program evaluation and feedback  
Administration and information system                                                                                                                             | B                |
| Drake et al., 2013     | A RCT testing the chronic care model of wrap-around services including supported employment, medication management and other mental health treatments against usual services, on employment outcomes. | The intervention group had a greater proportion of employment commencements (60% versus 40%) and had improvements in mental health and quality of life. Additional services offered to the intervention group were: training, experts, insurance cover and fidelity monitoring. | The quality of mental health treatment  
Technical assistance  
Legislation and regulation  
Program evaluation and feedback                                                                                                               | A                |
<p>| Durflak &amp; DuPre, 2008 | A review of the prevention and health promotion literature from 1976-2006 to examine the influence of implementation on outcomes (542 studies) and the factors affecting implementation quality (81 studies). | Identified seven dimensions of implementation in addition to fidelity and twenty-three ecological factors.                                                                                                                                            | Eight dimensions of implementation quality                                                                                                                                          | B                |</p>
<table>
<thead>
<tr>
<th>Included paper</th>
<th>Brief description of paper</th>
<th>Factors found to be directly influencing program performance</th>
<th>Empirical findings which were mapped to the conceptual framework</th>
<th>Relevance rating</th>
</tr>
</thead>
</table>
| Fixsen et al., 2005 | A systematic review covering multiple research fields including mental health. The findings from 22 empirical papers were reported separately. | • The role of experts in improving implementation, quality of care and patient outcomes.  
• Multilevel implementation strategies needed.  
• The competency and capacity of staff, organisational commitment and leadership.  
• Access to medication support and psychotherapy improved mental health and employment outcomes in patients with depression. | Technical assistance  
ES expertise  
Quality of mental health treatment | B |
| Glover & Froundelker, 2011 | An exploratory study to identify factors which differentiated successful from less successful employment specialists. | The more successful employment specialists worked efficiently, developed egalitarian relationships with program participants and collaborated effectively with partners. | ES expertise  
The quality of program delivery | A |
| Gold et al., 2006 | An RCT comparing assertive community treatment combined with IPS versus TVR in a rural area, 143 participants. | It was difficult to recruit staff in a rural area. The intervention outperformed the control on proportion of employment commencements and IPS participants also had more mental health services than control. | Role of geography  
Quality of mental health treatment  
Program intensity | A |
| Gowdy et al., 2004 | A qualitative study exploring the differences between high-performing and low-performing supported employment programs. | • Leaders emphasised the value of work and the belief that if people want to work they can work.  
• High-performing sites used data to track and feedback outcomes.  
• ES recognised that participants are motivated to find employment. | ES expertise  
Administration and information systems  
Leadership and program champions  
Program evaluation and feedback | A |
| Guydish et al., 2014 | An examination of the relationship between treatment fidelity, competence and empathy and six outcomes (number of days of drug, employment, medical, social and psychological outcomes) using an addiction treatment (twelve step facilitation) with 151 patients in 871 therapy sessions. | All three dimensions of quality were associated with improved employment stans at three-month follow-up. Higher therapist competence was associated with lower self-report drug use. | Staff expertise  
The quality of delivery i.e. empathy | A |
<table>
<thead>
<tr>
<th>Included paper</th>
<th>Brief description of paper</th>
<th>Factors found to be directly influencing program performance</th>
<th>Empirical findings which were mapped to the conceptual framework</th>
<th>Relevance rating²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heslin et al., 2011</td>
<td>An RCT comparing IPS to TVR in England with 219 participants followed for two years.</td>
<td>Fidelity was measured and found to be good, but outcomes were lower than other RCTs. The average number of contacts with the ES were low (15 over two-year period).</td>
<td>Program intensity</td>
<td>A</td>
</tr>
<tr>
<td>Johnson et al., 2009</td>
<td>A qualitative study which interviewed 182 participants to understand their perceptions of the effective ingredients of supported employment.</td>
<td>Service participants highlighted the importance of the quality of support that was offered, particularly emotional support and practical assistance, combined with a person-centred approach.</td>
<td>The quality of delivery, particularly through interpersonal dynamics</td>
<td>A</td>
</tr>
<tr>
<td>Kitson et al., 1998</td>
<td>Empirical testing of the PARIHS framework using four case studies in health care settings.</td>
<td>For successful implementation, the evidence for the intervention, the context and the method of facilitation should be equally weighted. For example, strength of evidence does not warrant more attention than the process of implementation. In one case study, the facilitator’s role overcame poor contextual conditions.</td>
<td>Technical assistance</td>
<td>B</td>
</tr>
<tr>
<td>Latimer et al., 2006</td>
<td>A Canadian RCT of IPS vs. TVR with 150 participants.</td>
<td>IPS program outcomes were lower than in the USA studies. This was attributed to aspects of social security policy. For example: a lack of financial incentives to earn more than Canadian $100 per week, free public transport passes only for those in TVR, and the availability of non-evidence-based programs which paid minimum wage or above. In addition, when attrition was controlled for the difference between the outcomes in the IPS sites and the controls was no longer significant.</td>
<td>Social security policy, Removal of non-evidence-based practices, Participant responsiveness i.e. attrition</td>
<td>A</td>
</tr>
<tr>
<td>Lockett et al., 2016</td>
<td>A systematic review and meta-analysis of IPS implementation studies to examine the predictive validity of the IPS fidelity scales.</td>
<td>Fidelity, whilst an important predictor of outcomes in a range of countries and contexts, only explains 16.8% of the variance in employment commencements.</td>
<td>There are factors in addition to fidelity which influence the outcomes of IPS programs.</td>
<td>A</td>
</tr>
<tr>
<td>Included paper</td>
<td>Brief description of paper</td>
<td>Factors found to be directly influencing program performance</td>
<td>Empirical findings which were mapped to the conceptual framework</td>
<td>Relevance rating $^2$</td>
</tr>
<tr>
<td>----------------</td>
<td>-----------------------------</td>
<td>----------------------------------------------------------</td>
<td>-----------------------------------------------------------------</td>
<td>-----------------</td>
</tr>
</tbody>
</table>
| MacDonald-Wilson et al., 2003 | Surveyed 539 service users, 120 service providers and 174 family members to examine the effects of social security work incentives on return to work. | • The concern over loss of welfare and health benefits  
• Lack of awareness of financial incentives that were available. | The role of social security policy  
Service users and families' attitudes and knowledge | A |
| McGuire et al., 2011 | Examined data from studies of 91 people who had received IPS to explore the relationship between service intensity and outcomes. | The level of service intensity predicted job tenure although where participants had high levels of cognitive symptoms this weakened the relationship. | Program intensity | A |
| Schell et al., 2013 | A literature review of 85 papers, combined with input from an expert panel, to identify the core constructs for achieving sustainability in public health programs. | Funding stability, political support, partnerships, organisational capacity, program adaptation, program evaluation, communications, public health impacts, strategic planning | Adaptation  
Program evaluation and feedback | B |
| Taylor & Bond, 2014 | A cross-sectional survey of 57 ES and their supervisors in high fidelity IPS programs to examine predictors of employment outcomes. | The amount of contact with an ES and the supervisor rated ES job performance and efficacy were related to job commencements and job tenure. | Program intensity  
ES expertise | A |
| Topor & Ljungberg, 2016 | Interviews with nine participants in an IPS program in Sweden to examine the importance of the relationship with the ES. | Participants described the ‘helpful relationship’ with the ES where they were ‘treated like a human being’. | Quality of delivery | A |
| Woitek et al., 2012 | A systematic review and meta-analysis of RCTs comparing collaborative chronic care models with other care models on patient outcomes; mental health symptoms and quality of life. Seventy-eight articles with 161 analyses from 57 trials. | Chronic care models improve mental health outcomes as well as physical health. | Administration and information systems  
Program evaluation and feedback  
Organisational and leadership support  
Patient support, including self-management tools, and coaching  
Networks and community links | B |

Notes. IPS=Individual Placement and Support. ES=Employment Specialist. TVR=Traditional Vocational Rehabilitation. BAKES scale=Behaviour Attitudes Knowledge of Employment Specialists scale (Corbière et al., 2007). PARIHS=Promoting Action on Research Implementation in Health Services framework. USA=United States of America. RCT=Randomised Controlled Trial. 1. $n=$ number of included studies. 2. Each empirical paper was rated according to its relevance to vocational rehabilitation for people with severe mental illnesses. A=Papers from vocational rehabilitation or psychiatry. B=Papers from the health sector but outside psychiatry.
Five additional dimensions of implementation quality, other than program fidelity, in the generic implementation framework were classified as having a direct influence. Three contextual factors were also considered to have a direct influence: Technical assistance, program evaluation and feedback, and staff expertise (see Table 10). Together these eight dimensions of implementation quality were expected to have the strongest direct influence on improving practices in vocational rehabilitation.

The new conceptual framework shows both direct and indirect influences (see Figure 6). The direct influences are shown within the dimensions of implementation quality ellipse (see Figure 6) and represent new candidates for dimensions of program implementation quality relevant to psychiatric vocational rehabilitation. They are: program intensity, quality of delivery, employment specialist expertise, program evaluation and feedback, technical assistance, quality of mental health treatment, removal of non-evidence-based practices, and participant responsiveness. These dimensions were determined from evidence of program performance based on the proportion of program participants commencing competitive employment.

Twenty-three contextual factors were identified and are shown in the quadrants outside the ellipse in Figure 6. These were classified using the four categories in the generic implementation framework: the socio-economic and political context, the implementation process, characteristics of the participating organisations, and characteristics of the participating individuals. These contextual factors are expected to indirectly influence program effectiveness, mediated through one or more of the dimensions of implementation quality. For example, funding levels and resourcing is likely to indirectly influence the proportion of program participants commencing employment, mediated through potential impacts on program fidelity, program intensity, quality of mental health treatment, the availability of technical assistance, and employment specialist expertise.
Figure 6. A conceptual framework for improving the effectiveness of evidence-based practices in vocational rehabilitation

Notes. There are eight dimensions of implementation quality in the ellipse which along with fidelity, are hypothesised to have the strongest direct influence on achieving high program performance. There are 23 contextual factors in the quadrants, these are hypothesised as having a mostly indirect influence mediated through fidelity and the other dimensions of implementation quality in the ellipse. These contextual factors may also have a direct influence, but this is considered to be either weaker or harder to measure. Any possible direct influence of these contextual factors is not discussed in this review. There are also likely to be multiple inter-relationships between all the dimensions of quality and contextual factors in the framework. These are not depicted in Figure 6. Each direct influence was rated based on the strength of evidence and feasibility to measure it. **=Good evidence (three or more studies from the vocational rehabilitation or general psychiatry literature providing evidence of a direct influence on program performance), and the dimension was considered sufficiently operationalised to be feasible to measure. *=Limited evidence (less than three studies from the vocational rehabilitation or general psychiatry literature providing evidence of a direct influence), but the dimension was considered sufficiently operationalised to be feasible to measure. No asterisk=Limited evidence for a direct influence and the dimension appeared insufficiently operationalised and therefore difficult to measure.
Discussion

This conceptual framework for implementing evidence-based vocational rehabilitation considers program fidelity along with other aspects of implementation quality. Eight additional dimensions of implementation quality are identified which are expected to have a direct influence on improving program performance. This framework is anticipated to be useful to program funders, practitioners, administrators, evaluators, and researchers to consider a broader range of direct and indirect influences on program performance.

Measuring program fidelity is an important part of this framework. A recent systematic review has shown that a focus on program fidelity is necessary during program implementation for developing program performance (Lockett et al., 2016). However, it is also clear from this review that it is possible for a program to have high fidelity yet be a poor performer in terms of the primary program outcome. Of the 52 cohorts reporting fidelity according to the IPS-15 scale, 40 were classified as high-fidelity, a total score of 61 out or more, out of 75, on the IPS-15 scale. Of these 15/40 (37.5%) were low performing (Lockett et al., 2016, p. 268). If high fidelity is achieved during program implementation, yet program performance remains below expectations, attention can shift to other dimensions of implementation quality that may be limiting program performance. For programs that achieve high fidelity and high performance, it is possible that attention to these other dimensions of implementation quality may further improve performance.

New dimensions of implementation quality in vocational rehabilitation

The structure of the IPS fidelity scales (IPS-15 and IPS-25) recognise the organisational context in which practices are examined. Principles of good practice are considered in both scales with respect to: staffing, services, and organisation. However, most other dimensions of program implementation quality are not covered by these scales (Bond, Drake, McHugo, Rapp, & Whitley, 2009). For example, fidelity assessors are instructed to observe employment specialists in meetings with program participants to observe the presence or absence of particular practices, such as the provision of financial guidance, or how disclosure issues are managed. The fidelity assessors are not required to observe other employment specialist attributes and behaviours, such as goal focus, knowledge of employers, or
sensitivity and empathy shown to participants. This is a notable limitation of the current fidelity scales because employment specialist attitudes and knowledge have been found to be directly associated with program outcomes (Glover & Frounfelker, 2011; Gowdy, Carlson, & Rapp, 2004).

**Employment specialist expertise**

Staff expertise describes the skills, attitudes, knowledge, and competencies of expert, or particularly successful staff, implementing evidence-based practices. This expertise has been identified as an important dimension of implementation quality (Berkel et al., 2011; Dusenbury, Brannigan, Hansen, Walsh, & Falco, 2005). This is particularly salient in the IPS approach to vocational rehabilitation, because the employment specialist working with individual program participants, can directly influence their motivation, behaviour, engagement with the program, and subsequent program outcomes.

Employment specialist expertise can be measured in terms of role experience, and through observational ratings of their behaviours and skills against pre-determined competency criteria (Guydish et al., 2014). The behaviours, attitudes, and knowledge of employment specialists have been measured using a standardised scale (Corbière, Brouwers, Lanctôt, & van Weeghel, 2014). This measure, named the BAKES (Behaviours, Attitudes, Knowledge in Employment Specialists) scale was tested with 153 employment specialists working in a range of supported employment services. A positive relationship was found between employment specialist attributes and program performance. Those who took a supportive and individually-focused approach, and who had relationships with employers and supervisors, had better program performance as measured by employment commencements and job tenure. Studies of other aspects of employment specialist expertise have found that knowledge of IPS principles is not as important as an ability to apply those principles (Glover & Frounfelker, 2011; Taylor & Bond, 2014). This is a promising research direction where further studies are needed to identify other aspects of employment specialist expertise which may further improve program outcomes.
Program intensity

Program intensity is the concentration of a program and contributes directly to program performance (Durlak & DuPre, 2008). It can also be measured in vocational rehabilitation programs (McGuire, Bond, Clendenning, & Kukla, 2011). Measurable aspects of intensity include the length of time participants are in the program, the numbers of face-to-face contacts between the participants and the employment specialist, the frequency of contact between the employment specialist and the clinical teams, and between the employment specialists and employers (Drake et al., 2012, p. 42).

Low program intensity represented by contact frequency between participants and the employment specialist, explains the unusually low proportion of participants commencing employment in a UK RCT of IPS compared with other RCTs. The UK IPS intervention achieved more than 80% adherence to fidelity (67/75 and 69/75) yet only 22% of participants obtained competitive employment (Heslin et al., 2011). A mean of 15 contacts between program participants and an employment specialist was recorded over a two-year period (Heslin et al., 2011). This is 3.2 times lower than the mean of 48 contacts over the two-year study period found by McGuire et al. (2011) in the USA. This USA IPS intervention achieved high levels of fidelity (72/75) and encouragingly, 75.0% of participants commenced competitive employment (Bond et al., 2007).

The quality of program delivery

The quality of program delivery refers to how well the different aspects of the program have been delivered to participants. This has been found to be positively associated with program outcomes (Berkel et al., 2011; Dane & Schneider, 1998; Dusenbury et al., 2005). In vocational rehabilitation, measures of the quality of program delivery include how well the employment specialist engages with participants about disclosure of participants’ mental health history, or how empathetic and responsive the employment specialist is to participants' preferences (Areberg, Björkman, & Bejerholm, 2013; Johnson et al., 2009; Topor & Ljungberg, 2016). The quality of relationships between the employment specialist and participants, were found to be positively associated with primary program outcomes in a six-site international RCT (Catty et al., 2011). Other authors have also acknowledged that the quality of
program delivery is likely to be an important influence on IPS program performance (Drake et al., 2012, p. 44).

**The quality of mental health treatment**

Some forms of vocational rehabilitation for people with severe mental illnesses such as IPS involve close coordination with a complementary program, namely continuing mental health treatment and care. Complementary programs are considered a core dimension of implementation quality which also needs to be taken into account (Durlak & DuPre, 2008). Monitoring these is important because comparison services and complementary program conditions can influence primary program outcomes. This implies a need to measure the amount and type of complementary program services received by participants. In the case of evidence-based vocational rehabilitation for people with severe mental illnesses, this could involve measuring the type and extent of mental health services along with any other wrap-around health or social services such as: general health services, housing, income support and welfare benefits, financial counselling, disability support, and legal assistance.

The provision of wrap-around services including mental health treatment and medication management, in addition to employment services were an integral part of the IPS intervention in a USA multi-site RCT (Drake et al., 2013). The intervention outperformed the control in terms of the proportion of participants commencing employment (60% versus 40%). In another IPS RCT, participants in the IPS program were found not only to receive more vocational services, but also more case management and medication management services. They also had more contacts with mental health services than the control group (140 contacts versus 83 contacts) (Gold et al., 2006). It remains unclear whether the amount of mental health treatment confounded the results or were an unintended, yet beneficial, consequence of program integration. Further investigations are needed to understand how the quality of mental health treatment, as a complementary program, contributes to the performance of the primary program, namely evidence-based vocational rehabilitation.

**Other dimensions of quality in the new framework**

The conceptual framework identifies four other dimensions of quality in addition to the four discussed above. These are: program evaluation and feedback, technical assistance, removal of non-
evidence-based practices, and participant responsiveness. Participant responsiveness is how the program holds the involvement and interest of program participants. These four dimensions were identified as direct influences on program performance in this review, because evidence from two or more studies converged to indicate a direct relationship. For example, program evaluation and feedback were found to distinguish high performing from low performing IPS sites (Gowdy et al., 2004). Technical assistance was provided across all the IPS sites in the large-scale USA trial (Drake et al., 2013), and continues to be a core feature of IPS implementation across the USA, (Drake & Bond, 2017) because it is believed by program leaders to support higher program fidelity and higher program outcomes.

The presence of pre-vocational training, an example of a non-evidence-based practice, was considered to have lowered the performance of a Swedish IPS program (Bejerholm, Areberg, Hofgren, Sandlund, & Rinaldi, 2015) and a Canadian IPS program (Latimer et al., 2006). Participant responsiveness may also influence program outcomes. For example, Latimer et al. (2006) found that the difference between the performance of the IPS program and the control group was no longer significant once participant responsiveness, measured as program attrition, was controlled.

Evidence for these other dimensions of quality may not be as strong in the vocational rehabilitation or general psychiatry literature as for the previously discussed four dimensions. However, when considered among the evidence from the broader implementation science literature (Coleman, Mattke, Perrault, & Wagner, 2009; Dorr et al., 2007; Fixsen et al., 2005; Kitson et al., 1998; Woltmann et al., 2012), these other dimensions appear relevant and applicable to vocational rehabilitation. For these reasons, they are also included as dimensions of implementation quality likely having a direct influence on program performance, in the conceptual framework for vocational rehabilitation (see Figure 6).

Quality assurance and control

This review also noted the extensive business management literature on quality assurance and quality control, and particularly the concept of total quality management. This approach attempts to ensure that an organisation delivers a consistent quality of goods of services (Schröder, Schmitt,
Schmitt, 2015). Recommended practices in quality assurance and quality management can and should underpin the delivery of all psychosocial programs. However, a focus on this alone is unlikely to advance our understanding of how to bring about higher program performance. This is because the focus of quality management is typically on compliance and the prevention of poor performance (Schröder et al., 2015), rather than generating high performance.

**Limitations**

It is possible that some relevant literature was missed because the wider implementation science literature is extensive and distributed across a large number of journals, and electronic databases. Mapping empirical evidence to dimensions of implementation quality and contextual factors revealed that the evidence appeared to become weaker when mapped against specific dimensions of quality and contextual factors. Perhaps this was because it was difficult to isolate the effects of individual dimensions or factors on program performance. There are also likely to be multiple dynamic relationships among the dimensions of quality and contextual factors, yet these are not depicted in the framework.

**Conclusions**

This literature review and synthesis generated a new conceptual framework to guide the implementation of evidence-based vocational rehabilitation program for people with severe mental illnesses. This framework extends current thinking by widening the definition of implementation quality beyond program fidelity, which considers only a pre-defined set of established principles and practices. This broader framework can guide investigations of other candidates that may be impeding high-fidelity implementations of evidence-based programs from achieving expected levels of program performance, as well as to further enhance the outcomes achieved by high-fidelity, high performing programs. It may also have relevance to vocational rehabilitation for individuals other than those with severe mental illnesses. This conceptual framework is intended to assist program funders, researchers, practitioners and policymakers to consider a wider range of potential influences on program
performance, in order to advance the science and practice of vocational rehabilitation for people with severe mental illnesses.

**Supplementary Analysis**

**The papers included in phase one of the literature review**

The first phase of the literature review reported in this chapter was based on 24 papers, 20 review studies and four single studies. These papers were used to identify the common dimensions of implementation quality and contextual factors in a generic conceptual framework. The description of these studies was not included in the published paper, Lockett, Waghorn, and Kydd, 2018a. Table 12 presents the 24 papers and provides a brief description of each, including how the paper contributed to the development of the generic framework.

**Expanding one of the dimensions of implementation quality**

In the generic implementation framework (see Table 10, column 1) the monitoring of comparison, or complementary conditions, was identified as a core dimensions of implementation quality. However, this description was narrowed to “quality of mental health treatment” in the published paper (Lockett et al., 2018a). During the process of reviewing the literature to operationalise the framework the dimension, “quality of mental health treatment” was broadened back out to “complementary programs” (see Figure 7).

The reason behind this change was that although mental health treatment is a critical ingredient of evidence-based practices in vocational rehabilitation (Drake, Becker, Bond, & Mueser, 2003), broadening the description opens the possibility that other important complementary programs could be identified. If identified, important complementary programs could assist to improve program performance. The concept, “quality of mental health treatment” is a subset of the dimension of implementation quality “complementary programs”.

86
Naming the new conceptual implementation framework

During the process of operationalising the concepts in the newly developed implementation framework for vocational rehabilitation programs for people with severe mental illnesses, the author made a decision to name it the Implementation Framework for Vocational Rehabilitation (IFVR) (see Figure 7).

Chapter Conclusions

This chapter presented a peer-reviewed report that investigated a current research and implementation gap in vocational rehabilitation by answering the research question: Should a wider theoretical implementation framework be applied to improve vocational rehabilitation for people with severe mental illnesses? The paper outlines a new implementation framework for improving the effectiveness of vocational rehabilitation programs, the Implementation Framework for Vocational Rehabilitation (IFVR).

The IFVR identifies twenty-three contextual factors and eight dimensions of implementation quality, in addition to implementation fidelity. The IFVR is theoretical, it needs testing. Program implementation is country specific, so before the IFVR can be operationalised an examination of the economic and policy context where the framework is to be applied needs conducting. Chapter 5 provides a case study of how to conduct such an analysis, using New Zealand as the example country.
<table>
<thead>
<tr>
<th>Study</th>
<th>Type of paper</th>
<th>Brief description of review/study</th>
<th>Main findings</th>
<th>Influences on program performance</th>
<th>Phase two review¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berkel et al., 2011</td>
<td>Empirical</td>
<td>A review of the empirical papers from three systematic reviews to develop an integrated model of program implementation.</td>
<td>Three dimensions of implementation (fidelity, quality and adaptation) influence participant responsiveness and program outcomes.</td>
<td>Quality of delivery, adaptation participant responsiveness.</td>
<td>Y</td>
</tr>
<tr>
<td>Bond et al., 2009</td>
<td>Empirical</td>
<td>A model for improving the fidelity of implementation to EBPs was tested in 53 sites across eight USA states, over two years.</td>
<td>The study identified ecological factors which influence fidelity (a) the specific features of the evidence-based practice, e.g. ease of implementation; (b) governmental factors; (c) leadership; (d) features of the fidelity review process.</td>
<td>Funding and incentives, features of the EBPs/innovation, leadership commitment to EBP implementation, removal of non-EBPs and conflicting practices, fidelity monitoring, feedback and utilisation of feedback.</td>
<td>N</td>
</tr>
<tr>
<td>Carroll et al., 2007</td>
<td>Conceptual</td>
<td>A review of conceptual papers between 2002 and 2007 to develop a conceptual framework for understanding and measuring implementation quality.</td>
<td>The role of facilitation strategies including manuals, training, guidelines, monitoring and feedback, capacity building, and incentives on the quality of implementation.</td>
<td>Participant responsiveness, program intensity and program reach, intervention complexity, facilitation strategies.</td>
<td>N</td>
</tr>
<tr>
<td>Dunschloroder et al., 2009</td>
<td>Conceptual</td>
<td>Review of 19 published theories on implementation, knowledge translation, innovation, and dissemination. The review standardised terminology, and combined the constructs to develop the CFIR for advancing implementation science.</td>
<td>There are five domains to implementation: intervention characteristics, outer setting, inner setting, characteristics of the individuals involved, and the process of implementation.</td>
<td>Intervention characteristics, the outer setting, the inner setting, characteristics of individuals, the implementation process itself.</td>
<td>N</td>
</tr>
<tr>
<td>Dane &amp; Schneider, 1998</td>
<td>Empirical</td>
<td>The review included 162 articles published between 1980 and 1994 which examined prevention interventions in children. Included studies needed to have measured the effects of fidelity and dosage on program outcomes.</td>
<td>Three dimensions of program fidelity were identified as being associated with program outcomes: 1. Exposure/dosage. 2. Adherence (fidelity). 3. Quality of delivery.</td>
<td>Adherence, exposure, quality of delivery, participant responsiveness, program differentiation.</td>
<td>Y</td>
</tr>
<tr>
<td>Dorr et al., 2007</td>
<td>Empirical</td>
<td>The review included 104 experimental studies. It examined the role of</td>
<td>Two-thirds of the reviewed experiments found that information systems had a</td>
<td>Information systems.</td>
<td>Y</td>
</tr>
</tbody>
</table>

¹ Y: Yes, N: No
<table>
<thead>
<tr>
<th>Study</th>
<th>Type of paper</th>
<th>Brief description of review/study</th>
<th>Main findings</th>
<th>Influences on program performance</th>
<th>Phase two review</th>
</tr>
</thead>
<tbody>
<tr>
<td>Durlak &amp; DuPre, 2008</td>
<td>Empirical</td>
<td>A review of the prevention and health promotion literature between 1976 and 2006 to look at: (a) the influence of implementation quality on outcomes (542 quantitative studies); and (b) the factors affecting implementation quality (qualitative and quantitative studies).</td>
<td>Positive influence on the process and outcomes of health care.</td>
<td>Dimensions of implementation quality: dosage, quality of delivery, participant responsiveness, monitoring of comparison conditions, program differentiation, program reach, and program adaption and 23 ecological factors.</td>
<td>Y</td>
</tr>
<tr>
<td>Dusenbury et al., 2005</td>
<td>Empirical</td>
<td>The implementation of a drug use prevention program delivered in seven schools in Baltimore was evaluated to look at the interaction between mediating variables.</td>
<td>The review found strong evidence to show that the quality of implementation affects program outcomes, and identified seven dimensions of implementation quality. The review also identified 23 ecological factors which influence the quality of implementation.</td>
<td>Dosage, adherence, quality of the process, adaptations, the attitudes of staff towards the program, staff understanding of program, staff prior experience.</td>
<td>N</td>
</tr>
<tr>
<td>Finnerty et al., 2009</td>
<td>Empirical</td>
<td>The construction and validation of an instrument (SHAY) to measure the effect of state conditions in facilitating EBP implementation. Data from the national implementing EBPs project were used (Bond et al., 2009).</td>
<td>Teacher’s understanding of the program positively correlated with program adherence, as did length of time teaching. Teachers that made positive adaptations were also more likely to adhere to program. Length of time teaching was correlated to positive adaptations.</td>
<td>Adequacy of finance, training - ongoing consultation and technical support, leadership (senior and operational), policies and regulations, fidelity monitoring, stakeholder support (consumers, families, providers).</td>
<td>N</td>
</tr>
<tr>
<td>Fixsen et al., 2005</td>
<td>Conceptual and empirical</td>
<td>This review of the implementation evaluation literature covered multiple research areas including psychiatry and covered all ages. The review included 743 studies, including 22 empirical studies.</td>
<td>Developed a conceptual framework for implementation.</td>
<td>Expert technical assistance, communication, feedback mechanisms, the sphere of influence, i.e. the socio-economic political context.</td>
<td>Y</td>
</tr>
<tr>
<td>Study</td>
<td>Type of paper</td>
<td>Brief description of review/study</td>
<td>Main findings</td>
<td>Influences on program performance</td>
<td>Phase two review¹</td>
</tr>
<tr>
<td>-------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>---------------</td>
<td>-----------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Greenhalgh et al., 2004</td>
<td>Conceptual and empirical</td>
<td>A meta-narrative review of 495 studies (including 213 empirical studies) which covered multiple disciplines including sociology, communications, medicine, marketing, and health promotion in the implementation of health care treatment and prevention for people of all ages.</td>
<td>Developed an evidence-based conceptual model for diffusion, dissemination and implementation of innovations in health services. Differentiated between ‘letting it happen’, ‘helping it happen’, and ‘making it happen’.</td>
<td>The system’s readiness for the innovation. Characteristics of the adopter. The outer socio-political context. The characteristics of the innovation and how implementation is communicated (unplanned, intentional). The implementation process including feedback.</td>
<td>N</td>
</tr>
<tr>
<td>Helfrich et al., 2010</td>
<td>Conceptual and empirical</td>
<td>A review of the literature on the PARIHS framework (24 articles, including 18 empirical papers).</td>
<td>Concept clarity is important in the development of implementation frameworks.</td>
<td>The importance of clear concept definition and what is meant by successful implementation.</td>
<td>N</td>
</tr>
<tr>
<td>Kitson et al., 1998</td>
<td>Conceptual and empirical</td>
<td>Based on experiences of authors working with nurses a three-dimensional conceptual framework for successful implementation was developed which considers: the evidence, the context and the method of facilitation. The framework was tested on four case studies in health care settings.</td>
<td>For successful implementation, all factors should be equally weighted i.e. strength of evidence doesn’t warrant more attention than the process of implementation. For example, in one case study, the facilitator’s role overcame poor contextual conditions.</td>
<td>Characteristics of the evidence-based practice. The environment (culture, leadership, measurement) The method / process of facilitating the implementation. All three aspects of the implementation framework are equally important.</td>
<td>Y</td>
</tr>
<tr>
<td>McGuire et al., 2015</td>
<td>Empirical single study</td>
<td>Compared implementation of illness management and recovery programs with sites that didn’t manage to implement it.</td>
<td>Successful sites provided: awareness of the evidence-based intervention, had importer-champions, autonomy-supporting leadership, veteran-centred care, sensitive period, psychosocial recovery centre.</td>
<td>Stuff with working knowledge of the evidence-based practices, expert and innovation champion, supportive leadership, complementary programs, openness to innovation, person-centred.</td>
<td>N</td>
</tr>
<tr>
<td>Nilsen, 2015</td>
<td>Conceptual</td>
<td>A review of selective literature to develop a taxonomy of implementation theories, models and frameworks.</td>
<td>Identified five categories: process models, determinant frameworks, classic theories, implementation theories, evaluation frameworks.</td>
<td>Characteristics of the implementation object, the user/adopter, the end users, the context, and the strategy for facilitating implementation.</td>
<td>N</td>
</tr>
<tr>
<td>Study</td>
<td>Type of paper</td>
<td>Brief description of review/study</td>
<td>Main findings</td>
<td>Influences on program performance</td>
<td>Phase two review</td>
</tr>
<tr>
<td>-----------------------</td>
<td>---------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>O’Donnell, 2008</td>
<td>Conceptual and empirical</td>
<td>Systematic review (five studies) to define and conceptualise fidelity of implementation (across the health and social sciences), and to measure the relationship between fidelity of implementation and outcomes in a teaching curriculum.</td>
<td>There is a lack of consistency on how fidelity is defined and measured. There is a need to distinguish between process and structural aspects of fidelity. Process aspects of fidelity are more strongly related to outcomes than structural aspects.</td>
<td>Adaptation and fidelity of implementation are different constructs. Fidelity has two aspects: process and structural.</td>
<td>N</td>
</tr>
<tr>
<td>Powell et al., 2014</td>
<td>Empirical</td>
<td>A systematic review of 11 studies which tested implementation strategies for evidence-based treatments using a controlled clinical trial design. Studies covered children, adolescents and adult mental health settings.</td>
<td>Passive strategies such as clinical guidelines, treatment manuals and opinion leaders are not sufficient in isolation Most studies used multi-faceted implementation approaches.</td>
<td>On-going training, access to supervision, expert consultation, peer support, fidelity monitoring, characteristics of the intervention, policy ecology.</td>
<td>N</td>
</tr>
<tr>
<td>Powell et al., 2011</td>
<td>Conceptual and empirical</td>
<td>Reviewed 205 articles on implementation strategies between 1995 and 2011. Used the CIFR to guide the review process.</td>
<td>68 implementation strategies grouped into six key implementation processes.</td>
<td>Planning, educating, restructuring, financing, managing quality, attending to the policy context.</td>
<td>N</td>
</tr>
<tr>
<td>Proctor et al., 2009</td>
<td>Conceptual</td>
<td>A review of conceptual frameworks.</td>
<td>Developed a conceptual model of implementation research. Identified different levels of ‘implementation outcomes’ including feasibility, fidelity, penetration, acceptability, sustainability, update and costs as well as ‘intervention outcomes’.</td>
<td>Reimbursement, legal and regulatory policy, organisational structure and strategy, cooperation and coordination, knowledge, skills, and expertise of staff implementation the intervention.</td>
<td>N</td>
</tr>
<tr>
<td>Sandström et al., 2011</td>
<td>Empirical</td>
<td>Systematic review to explore the role of nursing leadership in the implementation of EBPs. Seven included studies.</td>
<td>Characteristics of the leader (accessible, visible, credible, communicative, driven, commitment to research). Characteristics of the organisation (administrative support, policy revisions, resource allocation, human and material support, library access).</td>
<td>Characteristics of the leader and organisation, and the organisational culture.</td>
<td>N</td>
</tr>
<tr>
<td>Study</td>
<td>Type of paper</td>
<td>Brief description of review/study</td>
<td>Main findings</td>
<td>Influences on program performance</td>
<td>Phase two review&lt;sup&gt;1&lt;/sup&gt;</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------------</td>
<td>---------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>Schell et al., 2013</td>
<td>Conceptual and empirical</td>
<td>A literature review (85 papers, 53 with empirical component) and an expert panel (39 stakeholders) were used to develop the core domains of a conceptual framework for public health program’s capacity for sustainability.</td>
<td>Sustainability is an important program outcome and there are a number of factors which affect a program's ability to sustain its activities and outcomes over time.</td>
<td>Funding stability, political support, partnerships, organisational capacity, program adaptation, program evaluation, communications, strategic planning.</td>
<td>Y</td>
</tr>
<tr>
<td>Shojaeia &amp; Grimshaw, 2005</td>
<td>Conceptual</td>
<td>A critical review of the quality improvement literature.</td>
<td>The importance of multi-faceted strategies for successful implementation which include training, education and using data to enhance performance, field based supervision, systems reengineering.</td>
<td>Training, education and using data to enhance performance, field-based supervision, systems reengineering.</td>
<td>N</td>
</tr>
<tr>
<td>Stith et al., 2006</td>
<td>Conceptual and empirical</td>
<td>A review of 40 studies implementing prevention programs. Recommendations were weighted according to the strength of evidence of the source article.</td>
<td>The review identified five recommendations for implementing community-based prevention programs.</td>
<td>Community readiness, community fit, community coalitions, deliver program with fidelity, resources, training, technical assistance and evaluation.</td>
<td>N</td>
</tr>
<tr>
<td>Woltemann et al., 2012</td>
<td>Empirical</td>
<td>Systematic review and meta-analysis to assess comparative effectiveness of chronic care models for people with mental health conditions. Sample of 78 patients.</td>
<td>The main components improving physical and mental health outcomes are: patient skills and confidence; delivery system redesign; use of clinical information systems; provider decision support (e.g. reminders); linkage to community resources; health care organisational support.</td>
<td>Admin support and information systems, feedback systems, features of the organisation including support, leadership support, system redesign, patient support, including self-management tools, coaching, workforce redefinition of roles, provision of expert information, wider networks / community links.</td>
<td>Y</td>
</tr>
</tbody>
</table>

<sup>1</sup> Note. EBP=Evidence-Based Practices. CFIR=Consolidated Framework for Implementation Research. 3. SHAY=State Health Authority Yardstick. 4. PARIHS=Promoting action on research implementation in health services. Y=Yes. N=No.
The IFVR is adapted from the conceptual implementation framework in the published paper (Lockett, Waghorn, & Kydd, 2018a). A change was made during the operationalisation of one of the dimensions of implementation quality, ‘quality of mental health treatment’. The dimension was broadened out to the dimension of quality, complementary programs. The original dimension, quality of mental health treatment is a subset of this wider dimension of complementary programs. No other changes to the framework were made. There are eight dimensions of implementation quality in the ellipse which along with fidelity, are hypothesised to have the strongest direct influence on achieving high program performance. There are 23 contextual factors in the quadrants, these are hypothesised as having a mostly indirect influence mediated through fidelity and the other dimensions of implementation quality in the ellipse. These contextual factors may also have a direct influence, but this is considered to be either weaker or harder to measure. Any possible direct influence of these contextual factors is not discussed in this review. There are also likely to be multiple inter-relationships between all the dimensions of quality and contextual factors in the framework. These are not depicted in Figure 7. Each direct influence was rated based on the strength of evidence and feasibility to measure it. **= Good evidence (three or more studies from the vocational rehabilitation or general psychiatry literature providing evidence of a direct influence on program performance), and the dimension was considered sufficiently operationalised to be feasible to measure. *= Limited evidence (less than three studies from the vocational rehabilitation or general psychiatry literature providing evidence of a direct influence), and the dimension was considered sufficiently operationalised to be feasible to measure. No asterisk= Limited evidence for a direct influence and the dimension appeared insufficiently operationalised and therefore difficult to measure.

Figure 7. The Implementation Framework for Vocational Rehabilitation (IFVR)
Chapter 5: Applying the Implementation Framework for Vocational Rehabilitation (IFVR)

Prelude

Rationale for study 3: An analysis of New Zealand’s economic and social policy context

The findings from the systematic review and meta-analysis in Chapter 3 support the rationale for this third investigation, that to improve program performance it is not sufficient to focus only on improving fidelity to evidence-based practices (Lockett et al., 2016). Chapter 4 systematically brought together the conceptual and empirical evidence from the broader science of psychosocial program implementation. Through this review and synthesis of the literature, the most promising candidates for researchers, policymakers, funders, and practitioners seeking to improve the performance of vocational rehabilitation programs for people with severe mental illnesses, were identified. The most promising candidates are outlined in a new Implementation Framework for Vocational Rehabilitation (IFVR). The IFVR consists of eight dimensions of implementation quality in addition to implementation fidelity, and 23 contextual factors (Lockett et al., 2018a).

The application of the IFVR is country specific. This fifth chapter therefore seeks to address the third research question of this investigation: How does the socio-economic policy context influence the implementation of evidence-based practices in vocational rehabilitation? This involves a systematic examination of documents and national surveys relating to the economic and social policy context in New Zealand to identify if, and how the wider policy environment is affecting the implementation of evidence-based practices in vocational rehabilitation for people with severe mental illnesses.

In conducting a policy analysis, this chapter targets a knowledge gap in the science of implementing evidence-based practices that is seldom discussed or reported. This knowledge gap consists of the social and economic policy settings that can facilitate, hinder, or preclude the implementation of known evidence-based practices in vocational rehabilitation. The economic and social context of a country represents an important influence on the performance of evidence-based practices. Exactly how this influence is affecting implementation and how this wider contextual...
influence should be measured, is less understood (Drebing et al., 2012; Metcalfe, et al., 2017a; Metcalfe et al., 2018b, Modini et al., 2016).

New Zealand was chosen as it the country because this is where the author of this investigation resides. New Zealand also has a relatively small population, with a single national government and no state or federal governments. This limits the applicable policy documents and national surveys. In addition compared to other countries, New Zealand has a relatively simple service delivery and policy structure which lends itself to a targeted policy analysis. Although the focus is on New Zealand policy settings the methods of analysis applied in this chapter could be applied to other countries. In addition, the implications of the policies included in the analysis have application to other countries through parallels with other country's policy settings.

The following material on pages 96 to 118 is reprinted from WORK: A Journal of Prevention, Assessment and Rehabilitation, 60(3), 421-435. Lockett, H., Waghorn, G., & Kydd, R., Policy barriers to evidence-based practices in vocational rehabilitation for people with psychiatric disabilities in New Zealand. Copyright (2018), with permission from IOS Press. The publication is available at IOS Press through doi:10.3233/WOR-182752. Permission received on 21 August 2018.

Following the presentation of the peer-reviewed paper, the chapter finishes with concluding comments to lead into Chapter 6. There is no supplementary analysis reported in Chapter 5.
Introduction

Addressing the science to practice gap in the rehabilitation sciences is a ubiquitous challenge for the health and rehabilitation sciences. It is not enough to discover or develop more effective practices, because if new evidence-based practices are not adopted or implemented, and existing ineffective practices are not displaced, no community benefits will be realised (Fixsen et al., 2013). Even when there is rigorous and overwhelming scientific evidence for the efficacy of an intervention or set of practices, they often do not quickly become widely available in routine service settings (Drake et al., 2016). It is possible that such failures in the translation of evidence into practice are linked to out of date government policy and regulatory frameworks that can facilitate, hinder or preclude the adoption of evidence-based practices (Drake et al., 2016; King et al., 2006). Despite the potential importance of government policy settings for supporting the translation of evidence into practice (Lockett et al., 2018a), the background policy context is rarely investigated (Goldman et al., 2001).

The term psychiatric disability is often used to represent mental illnesses such as schizophrenia, schizoaffective disorder, first episode psychosis (excluding drug-induced psychosis), bipolar affective disorders, affective disorders with psychotic features, and other severe psychiatric disorders as defined in the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2013). These conditions represent the psychiatric disorders most likely to produce high levels of psychiatric disability which is defined as significant impairments to social and occupational functioning.

The extent of disability associated with a particular psychiatric disorder varies across individuals but can range from none to severe, while often fluctuating in cycles by days, weeks or months, depending on: disorder type, methods of treatment, course patterns of illness, and the social context. For the majority of working-age adults living with psychiatric disabilities, obtaining a first job or returning to employment after acute episodes of illness remains an important personal recovery goal. For example, returning to employment was one of the top three recovery priorities identified by all respondents in the Australian national survey of high impact psychosis. Among those aged 18-34 years, employment was rated as most important ahead of financial matters, loneliness and social isolation...
Increasing opportunities for employment and increasing community integration are considered important steps in a recovery-oriented approach to mental health treatment and care (Leamy, Bird, Le Boutillier, Williams, & Slade, 2011). There is also promising evidence that once established, employment contributes to ongoing recovery and wellness (Hoffmann et al., 2014; Peterson, et al., 2017). However, these recovery goals remain largely unattained in this population according to a recent international review of labour force activity (Jonsdottir & Waghorn, 2015).

The science of vocational rehabilitation for people with psychiatric disabilities has advanced significantly over the past thirty years. There is now a widely agreed set of evidence-based principles and practices developed for those with psychiatric disabilities who are also typically involved with public funded mental health services. Known as the Individual Placement and Support (IPS) approach to supported employment, it has several distinguishing features (see Table 13) and has proven more effective than alternative approaches in terms of the proportion commencing employment (Drake et al., 2012, p. 55).

Table 13. Five distinguishing features of Individual Placement and Support

<table>
<thead>
<tr>
<th>Distinguishing features¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
</tbody>
</table>


A recent meta-analysis of 17 RCTs found that the odds of commencing employment using IPS programs were 2.4 times greater than in programs utilizing more traditional vocational rehabilitation practices (Modini et al., 2016). IPS programs have also outperformed alternative programs on other outcome measures including time to first job and the proportion working more than 20 hours per week.
(Bond et al., 2012b). Participants in IPS employment programs can also have a significant reduction in admissions to psychiatric hospital and days spent in hospital compared to participants in other vocational programs (Hoffman et al., 2014). Yet despite being more effective than other forms of vocational rehabilitation, IPS is not yet routinely available in mental health treatment services within the USA (Drake et al., 2016). Nor is it routinely available in mental health services outside the USA in countries such as Canada, UK, Australia, and New Zealand (Boardman & Rinaldi, 2013; Lockett & Bensemann, 2013; Waghorn & Hielscher, 2015). For countries to make available and target, these evidence-based practices to people who need intensive employment assistance, it is important to understand the labour force activity of people with psychiatric disabilities, and how this profile differs from that of the wider community (King et al., 2006).

The aim of this analysis was to identify whether, and how, the availability of evidence-based practices in vocational rehabilitation is linked to government social and economic policy settings. This involved systematically examining the New Zealand social and economic policy context for conflicts, barriers, opportunities and challenges, in relation to evidence-based practices in vocational rehabilitation for people with psychiatric disabilities. The Individual Placement and Support approach to vocational rehabilitation is the focus of this analysis through currently being the leading evidence-based form of vocational rehabilitation designed for people with psychiatric disabilities. The analysis was driven by three questions (see Figure 8).

1. What is the extent of unmet need for evidence-based practices in vocational rehabilitation in New Zealand for people with psychiatric disabilities?

2. To what extent are evidence-based practices in vocational rehabilitation available for people with psychiatric disabilities in New Zealand?

3. How do New Zealand social and economic policy settings facilitate, hinder, or preclude the provision of evidence-based practices in vocational rehabilitation for people with psychiatric disabilities?
Methods

Data search

Data searching began by reviewing summary reports from January to December 2016 from well-known national surveys conducted by Statistics New Zealand (Statistics NZ) and two other government departments. These surveys were: the New Zealand Mental Health Survey (Oakley Browne, Wells, & Scott, 2006); the New Zealand Health Survey (Ministry of Health (MoH), 2016a); and the New Zealand Disability Labour Force Survey (Statistics NZ, 2014). The publicly accessible data files from the New Zealand Household Labour Force (Ministry of Business, Innovation and Employment (MBIE), 2014) and the Disability Labour Force surveys (Statistics NZ, 2013) were also examined. In addition, the Google scholar database was searched using the terms labour force, employment, unemployment; combined with psychiatric disability, disability, or mental illness; and New Zealand, to identify any relevant published peer-reviewed literature.

Next, information was sought about the availability of evidence-based practices in vocational rehabilitation for people with psychiatric disabilities, and the different ways these practices are purchased in New Zealand. This step was facilitated by an unpublished survey of employment support providers in the New Zealand mental health sector conducted in 2014 (Te Pou o te Whakaaro Nui (Te Pou), 2015). The 20 government providers of mental health services who participated in the 2014 survey were contacted. Each organisation was asked about current evidence-based practices in vocational rehabilitation, and if relevant, asked to provide copies of documents such as vocational rehabilitation funding contracts, employment service evaluation reports, and staff training policies.

The wider social and economic policy context was systematically explored to assess whether it supports or hinders the availability of evidence-based practices in vocational rehabilitation. Over a two-year period from March 2015 to March 2017, the first author consulted with leaders in the field using a snowball sampling method (Bryman, 2016, p. 415). This step also involved contacting government advisors from the health, welfare, and disability agencies, and staff of Statistics NZ responsible for the disability labour force survey. Government department websites were searched for relevant population level social and economic statistics; health, disability, and welfare strategies, or policy documents.
during the period January 2016 to January 2017. The most applicable government websites were: The Ministries of Health; Social Development; Business, Innovation and Employment; the Office for Disability Issues; the New Zealand Treasury; and Statistics NZ.

To obtain copies of vocational rehabilitation contracts, contract managers from the Ministry of Health (MoH), the Ministry of Social Development (MSD), and the non-government providers of IPS were contacted. Specific policy and contractual documents were sought that were relevant to the delivery of vocational rehabilitation services for people with psychiatric disabilities.

**Data analysis**

Labour force surveys were examined for information about relevant subpopulations, such as all New Zealand residents of working age, working age adults using mental health services, and disabled residents versus non-disabled residents. Information about labour force activity was considered relevant if based on internationally agreed definitions of labour force activity where three mutually exclusive categories apply: not in the labour force (either not available for or not seeking employment); unemployed (both available for and seeking employment); and currently employed according to a specific definition of employment (International Labour Office, 1982).

Evidence-based practices were considered available if one or more of the following criteria were met by the evidence sourced (a) the funding contract required that IPS programs were being purchased, (b) an assessment of IPS practices conducted in the last five years had found evidence-based practices were being provided, or (c) the employment provider demonstrated a commitment to training staff to deliver evidence-based practices in vocational rehabilitation for people with psychiatric disabilities. Evidence sourced from all 20 health regions of New Zealand were mapped against these criteria.

To assess the impact of New Zealand's social and economic policy settings an analysis was conducted in three parts.

1. A description of the relevant features of the New Zealand social and economic context including profiles of adults receiving welfare benefits by payment types.
2. An examination of health, welfare and disability policies to identify references to evidence-based practices in vocational rehabilitation for people with psychiatric disabilities.

3. An evaluation of the strengths and limitations of vocational rehabilitation contracts for providers targeting people with psychiatric disabilities while intending to apply evidence-based principles.

Relevant government policies were considered current if published in the past 10 years and if they remained the most recent version available. The vocational rehabilitation contracts were evaluated in relation to whether they helped or hindered the implementation of the eight IPS principles (Drake et al., 2012, pp. 33-39).

**Data synthesis**

The aim of data synthesis was to collect the general and specific implications for the availability of evidence-based practices in vocational rehabilitation for people with psychiatric disabilities (see Figure 8). The need for such a synthesis is supported by the “Recommendations of the Council of Integrated Mental Health, Skills and Work Policy” (OECD, 2015b). These recommendations encourage the New Zealand government to develop policy in relation to employment and mental health.
Figure 8. The steps taken in data search, data analysis, and the synthesis of findings

Notes. EBPs=Evidence-Based Practices. VR=Vocational Rehabilitation. NZ=New Zealand. 1. Labour force surveys (LFS) were examined for relevant sub-populations. Information was considered relevant if based on internationally agreed definitions of labour force activity (International Labour Office, 1982). 2. Documents searched were: the NZ Mental Health Survey (Oakley Brownell, Wells, & Scott, 2006); the NZ Health Survey (Ministry of Health, 2016a); the NZ Disability LFS (Statistics NZ, 2014), and a doctoral thesis (Welsh, 2010). 3. EBPs were considered available if one or more of the following criteria were met: (a) the funding contract required that EBPs were being purchased, (b) an assessment of program practices conducted in the last five years found EBPs were being provided, or (c) the employment provider demonstrated a commitment to training staff to deliver EBPs. 4. Documents sourced: the NZ Disability Strategy (Office for Disability Issues, 2016); the report of the Welfare Working Group (2011); the State Services Commission’s Better Public Services targets (State Services Commission, 2012); the Ministry of Social Development’s annual performance report 2013 (Ministry of Social Development, 2014); the NZ Mental Health Strategy (Ministry of Health, 2012); and the NZ Health Strategy (Ministry of Health, 2016b,c). 5. Evaluated against the principles of IPS (Drake, Bond, & Becker, 2012, pp. 33-39). 6. Informed by the Recommendations of the Organisation for Economic Co-operation and Development (OECD), 2015b.
The latest recommendations highlight the need for concerted action in health policy, labour market policy and social policy. New policies are considered needed for: a more integrated approach to mental health and people of working-age; introducing employment outcomes into the health system’s quality and outcomes frameworks; and for more responsive social welfare systems that stimulate co-operation between employment services and the health sector (OECD, 2015b).

**Results**

**The need for evidence-based practices in vocational rehabilitation**

Limited national information was found about labour force activity by diagnostic groups, by extent and severity of disabilities, or by use of public mental health services. This was because health data collections and surveys, either do not measure or do not report labour force activity. Welfare surveys and data sets also had low utility because they did not use internationally agreed terms to describe and classify psychiatric disorders and psychiatric disabilities. Internationally agreed terms are reported in the regularly updated series, the International Statistical Classification of Diseases and Related Health Problems (WHO, 1992). While New Zealand has a national disability labour force survey, mental illnesses are not differentiated from other chronic health conditions which can also cause difficulties with social and occupational functioning (see Table 14). These are described in the labour force survey in more general ways as difficulties with communicating, socialising or in everyday activities such as employment (Statistics NZ, 2014).

These limitations hinder estimating the unmet need for evidence-based practices in vocational rehabilitation in New Zealand. Table 14 summarises the relevant information extracted during the data and literature search. Until more New Zealand-specific information is available, targeting those most in need of evidence-based practices in vocational rehabilitation can be informed by findings from Australia and other comparable overseas studies (see Table 14).
Table 14. *The labour force activity of working-age population*¹ groups in New Zealand (NZ) compared to international benchmarks

<table>
<thead>
<tr>
<th>Population group</th>
<th>Not in the labour force²</th>
<th>Unemployed³</th>
<th>Employed⁴</th>
<th>Row totals</th>
<th>Source⁵</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (000)</td>
<td>Row %⁶</td>
<td>n (000)</td>
<td>Row %</td>
<td>n (000)</td>
</tr>
<tr>
<td>All household residents aged 15 years and over</td>
<td>1114</td>
<td>31.1</td>
<td>137</td>
<td>3.8</td>
<td>2328</td>
</tr>
<tr>
<td>Residents without disabilities aged 15 years and over</td>
<td>612</td>
<td>24.2</td>
<td>102</td>
<td>4.0</td>
<td>1816</td>
</tr>
<tr>
<td>Residents with disabilities aged 15 years and over</td>
<td>459</td>
<td>50.1</td>
<td>42</td>
<td>4.6</td>
<td>416</td>
</tr>
<tr>
<td>Māori without disabilities aged 15 years and over</td>
<td>72</td>
<td>24.2</td>
<td>24</td>
<td>8.1</td>
<td>202</td>
</tr>
<tr>
<td>Māori with disabilities aged 15 years and over</td>
<td>64</td>
<td>46.7</td>
<td>12</td>
<td>8.8</td>
<td>61</td>
</tr>
<tr>
<td>People with psychological or psychiatric disabilities⁶ aged 15 years and over</td>
<td>90</td>
<td>47.6</td>
<td>13</td>
<td>6.9</td>
<td>85</td>
</tr>
<tr>
<td>Long-term⁸ service users of public mental health services aged 18-64</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>People with affective disorders</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>People with bi-polar disorders</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>People with schizophrenia spectrum disorders</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Notes.
1. The working-age population is NZ residents aged 15 years and over who are either not in the labour force, unemployed, or employed (full-time or part-time). The legal age for leaving school in NZ is 16 years, so this definition of the working-age population includes people aged 15 who are still in school. Students are classified as not in the labour force.
2. The proportions not in the labour force are the number of people who are neither employed nor seeking employment (unemployed), divided by the row total in scope. Each frequency unit (n) represents 1000 people.
3. Unemployed is defined as being available for and actively looking for employment. The proportion is calculated with the row total in scope as the denominator.
4. The proportion employed is the number of employed people divided by the row total in scope as the denominator.
5. Sources of information: (a)=Ministry of Business, Innovation and Employment, August 2014, p. 5; (b)=Statistics NZ, March, 2013: Table 1: Labour force status by gender and age group; Table 3: Labour force status by ethnic group; Table 6: Labour force status by impairment type; (c)=Welsh, 2010. Data period: July 2003 to July 2005; (d)=International review by Jonsdottir and Waghorn, 2015.
6. Row percents add to 100% within rows.
7. This includes anyone with a long-term health condition or disability which causes difficulties with communicating, socializing or everyday activities (20).
8. Long-term was defined as in contact with public mental health services for two years or more.
9. n/a=data not available.
The availability of evidence-based practices in vocational rehabilitation for people with psychiatric disabilities

An unpublished survey in 2014 (Te Pou, 2015) of employment support services in the New Zealand mental health sector, found that there was only one non-government employment provider aligning practices to the IPS method. This provider was delivering IPS in six of 20 District Health Board regions. Subsequent contact with these mental health and employment service providers revealed that in July 2017 IPS programs were being provided in seven of 20 District Health Board regions in New Zealand and were about to commence in two more regions.

From May 2015, an IPS technical assistance program supported the implementation of high fidelity IPS in two of these seven regions (Te Pou, 2017). However, even in regions where IPS is available, it is not yet accessible by every eligible adult under the care of public mental health services. For example, in one region IPS is available via four adult community mental health teams but is not yet linked to three specialist teams, namely the mental health team for indigenous Māori people, for Pacific people, or the early intervention in psychosis service. This is because only four employment specialist positions are funded (Te Pou, 2017).

In another region, although a full-time employment specialist is integrated into each of the adult community mental health teams, three of these clinical teams have more than 70 clinicians per team. When employment specialists are assigned to very large community mental health teams, they spend additional time building relationships with clinicians to develop efficient referral pathways, and to coordinate employment services with clinical services. Integration depends on taking referrals from each clinician, and this is unlikely in very large teams, simply because not enough of the clinical team can have a shared caseload with the employment specialist. Further, only a small proportion of the clinicians’ caseload, will have access to best vocational assistance (Te Pou, 2017).

Regions of New Zealand where IPS is not yet available have access to other forms of supported employment that accept people with psychiatric disabilities. Although these employment services have some features in common with IPS supported employment (Te Pou, 2015), they do not offer vocational rehabilitation integrated with mental health treatment and care services. In other regions government
and non-government organisations also provide alternative forms of vocational rehabilitation such as those based in sheltered workshops and in business enterprises (Te Pou, 2015).

**Relevant features of the New Zealand economic and social context**

New Zealand has a developed and growing market economy, and a small, diverse and largely urbanised population (see Table 15). However, there is an ongoing economic disparity and social inequity between the indigenous Māori and non-Māori people. This is represented by lower proportions employed (MBIE, 2016, p. 1) and more Māori people on welfare benefits than expected compared to the general population (MSD, 2014, p. 48). There is also evidence that greater proportions of Māori people experience common mental disorders than non-Māori people (Horwood & Fergusson, 1998; Oakley Browne et al., 2006). The dual disadvantages of indigenous status and poorer mental health may exacerbate overall labour force disadvantage, suggesting that this population subgroup may also need a tailored form of vocational rehabilitation.

As of 1 April 2017, the adult minimum wage was New Zealand $15.75 per hour. The purchasing power parity of the New Zealand Dollar to the USA Dollar (USD) was 1.47 in December 2015 (OECD, 2016), meaning that an equivalent minimum wage in the USA would be $10.71 USD per hour. The New Zealand minimum wage applies to all adults aged 16 or more, unless specifically exempted. Exemptions for a person with a disability in certain circumstances were introduced in 1983. A person with a disability can be exempted if it is deemed that the person’s disability reduces their ability to work and to earn the minimum wage (Minimum Wage Act 1983, s. 6). Consequently, some alternative vocational rehabilitation programs appear to have utilised this exemption and do not pay the minimum wage. Hence, it is possible that this exemption has inadvertently supported non-evidence-based practices to continue to be purchased and provided.
Table 15. Relevant features of the New Zealand (NZ) social and economic context, 2015-2016

<table>
<thead>
<tr>
<th><strong>Social context</strong></th>
<th><strong>Key facts</strong></th>
<th><strong>Source(^1)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>Total resident population estimated at 4.6 million</td>
<td>(a)</td>
</tr>
<tr>
<td></td>
<td>73% live in the 16 main urban areas</td>
<td>(a)</td>
</tr>
<tr>
<td></td>
<td>53% live in the four largest cites</td>
<td>(a)</td>
</tr>
<tr>
<td></td>
<td>55% live in the northern half of the North Island</td>
<td>(a)</td>
</tr>
<tr>
<td>Ethnic diversity of resident population(^2)</td>
<td>75% European</td>
<td>(a)</td>
</tr>
<tr>
<td></td>
<td>16% indigenous Māori people</td>
<td>(a)</td>
</tr>
<tr>
<td></td>
<td>8% Pacific Islanders</td>
<td>(a)</td>
</tr>
<tr>
<td></td>
<td>12% Asian</td>
<td>(a)</td>
</tr>
<tr>
<td></td>
<td>1% classed as other</td>
<td>(a)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Economic context</strong></th>
<th><strong>Key facts</strong></th>
<th><strong>Source(^1)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Economic growth</td>
<td>3.2% per annum</td>
<td>(a)</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>5.7%</td>
<td>(b)</td>
</tr>
<tr>
<td></td>
<td>Unemployment profiles are subject to regional and seasonal variation</td>
<td>(c)</td>
</tr>
<tr>
<td>Unemployment rate by ethnicity</td>
<td>4.4% Europeans</td>
<td>(d)</td>
</tr>
<tr>
<td></td>
<td>12.2% Māori people</td>
<td>(d)</td>
</tr>
<tr>
<td></td>
<td>11.6% Pacific Islanders</td>
<td>(d)</td>
</tr>
<tr>
<td></td>
<td>6.9% Asian</td>
<td>(d)</td>
</tr>
<tr>
<td>Adult minimum wage(^3)</td>
<td>$15.75 per hour (NZD)</td>
<td></td>
</tr>
<tr>
<td>Income support payments available for people with psychiatric disabilities(^4)</td>
<td>People can claim either Jobseekers Support (health and disability category) or the Supported Living Payment(^5)</td>
<td></td>
</tr>
</tbody>
</table>

**Notes.**
2. Since NZ census respondents are permitted multiple responses to ethnicity questions, these classifications are not mutually exclusive.
3. The NZ minimum wage applies to all adults aged 16 years or more, unless specifically exempted. Exemptions for a person with a disability in certain circumstances were introduced in 1983. A person with a disability can be exempted if it is deemed that the person’s disability reduces their ability to work and to earn the minimum wage. (Minimum Wage Act 1983, s.6).
4. To claim an income support payment due to a health condition or disability, the person must be temporarily or permanently unable to work due to a health condition or disability.
5. Supported Living Payment is paid when health conditions or disabilities severely limit the person’s capacity to work for more than two years and they cannot regularly work 15 hours or more a week in open employment. The Supported Living Payment is paid at a higher weekly amount than Jobseekers Support (Social Security Act 1964, Part 1E s. 40B).
New Zealand disability, welfare and health policy

New Zealand has the advantage of one single national parliamentary government uncomplicated by state or provincial governments. However, the disability, employment, and health services sectors operate under different government departments. Each department, or governing agency, can have different policy objectives, contracting methods, and different reporting requirements to different government Ministers. This creates a barrier to the implementation of evidence-based practices in vocational rehabilitation because people with psychiatric disabilities require an integrated policy and funding framework to ensure both their mental health and vocational needs are met in a coordinated way (Arends et al., 2014; Drake et al., 2016). Furthermore, service providers find it more challenging, time consuming, and therefore more costly to deal with multiple government agencies when delivering a specialised employment service that to be effective, needs close coordination with mental health services (King et al., 2006).

Nevertheless, the relatively small number of interested government departments and agencies permitted an exhaustive search to be completed. This search identified seven key policy documents whose contents were analysed for references to evidence-based practices in vocational rehabilitation for people with psychiatric disabilities. These documents were: the New Zealand Disability Strategy (Office for Disability Issues, 2016); the report of the Welfare Working Group (2011); the State Services Commission’s Better Public Services targets (State Services Commission, 2012); the MSD’s annual performance report 2013 (MSD, 2014); the New Zealand Mental Health Strategy (MoH, 2012); and the New Zealand Health Strategy (MoH, 2016b,c).

The New Zealand disability strategy

The Office for Disability Issues is currently responsible for developing disability policies and social inclusion strategies. The latest New Zealand disability strategy was published in November 2016. The strategy anticipates a future where New Zealand citizens with disabilities have a secure economic situation and can achieve their full potential (Office for Disability, 2016, p. 7). This strategy states that the future needs to ensure those people who need specialised supports and services have “ready access to them to secure and sustain employment” (Office for Disability Issues, 2016, p. 26). An
outcomes framework which sets targets and measures for a disability action plan were being developed at the time of writing. Although this disability strategy offers a supportive policy direction, there is not yet any mention of evidence-based practices or the need to make these available, or any reference to the specific employment assistance needs of people with psychiatric disabilities.

**Reducing long-term welfare dependency**

Providing income support payments and employment assistance are the responsibility of the MSD through its agency Work and Income New Zealand. Since 2012 the MSD has introduced policy reforms to reduce the numbers of people of working-age who receive income support payments. These reforms were largely in response to the recommendations of the Welfare Working Group (2011). These independent experts recognised the need for evidence-based practices in supported employment, highlighting that “for people with severe mental illness, individual placement support programs are more effective at helping people to find jobs than are pre-employment training schemes” (Welfare Working Group, 2011, p. 154). However, the MSD policy reforms have not yet adopted this observation by recommending increasing access to evidence-based vocational rehabilitation for people with psychiatric disabilities (MSD, 2014).

In 2012 the New Zealand Prime Minister launched a whole-of-government strategy to improve the effectiveness of public services. Known as the Better Public Services program, it identified 10 results-driven targets for all public service departments and agencies that require cross-agency collaboration. There is a lead agency for each target but all government agencies are required to show how they are contributing to meeting these targets (State Services Commission, 2012, p. 2). The first target is a reduction in long-term welfare dependence, led by the MSD. Since 2012 this strategy has reduced the overall number of people claiming income support payments (MSD, 2014, p. 1).

Despite an overall reduction in those receiving welfare payments, the number of people receiving income support payments due to a health condition or disability has not reduced (MSD, 2014, p. 4). In addition, the proportion claiming income support for psychiatric conditions has increased (MSD, 2016). For example, the proportion claiming Supported Living Payment for psychiatric
conditions in March 2012 was 30.4% \{25,452 of 83,657\} and by June 2016 this had increased to 33.5% \{28,307 of 84,609\} (MSD, 2016).

**Health policy**

The Ministry of Health delegates its overall responsibility for health policy and services to 20 District Health Boards. These Boards are responsible for providing directly, or by purchasing, health services in their respective regions. In 2012, the Ministry of Health published a five-year service development plan for mental health and addictions services, known as Rising to the Challenge (MoH, 2012). This plan identifies employment and education as priority areas for increasing social inclusion among people who experience mental health and addiction problems. District Health Boards are expected to “increase access to employment specialists delivering evidence-informed individual placement and support services” for people with low prevalence mental health conditions or high needs, which includes people with psychiatric disabilities (MoH, 2012, p 28). In addition, District Health Boards are accountable for employment and education outcomes (MoH, 2012, p. 33).

The need for District Health Boards to deliver evidence-based vocational rehabilitation as part of their core business was further reinforced in the 2016 New Zealand Health Strategy. This specifies an action point to “collaborate with other government agencies to implement an evidence-based program of vocational rehabilitation to keep people with long-term conditions in employment.” (MoH, 2016c, p. 11). Mental health conditions are included in the definition of long-term conditions (MoH, 2016b, p. 22).

In response to this national policy shift, some District Health Boards have specified actions in their annual plans to increase the labour force participation of people with psychiatric disabilities (Waitemata District Health Board, 2016, p. 40). These regional health plans provide an opportunity to develop a more detailed implementation framework to foster evidence-based practices, a framework that if developed could strengthen national health policy.

Whilst this health policy direction is supportive, there is no mention of return to employment as an important primary outcome and performance indicator for health care services (OECD, 2015b;
This means there is no incentive from health policy for health care funders and providers to make vocational rehabilitation a part of the core business of health services.

**Purchasing vocational rehabilitation in New Zealand**

Vocational rehabilitation is currently purchased through several government agencies or departments. The Ministry of Social Development purchases and provides employment assistance and vocational rehabilitation services directly through its delivery arm, Work and Income New Zealand, as well as through activity-based and results-based contracts with non-government vocational rehabilitation providers. The Ministry of Health purchases health services via the 20 District Health Boards using activity-based contracts. District Health Boards provide vocational rehabilitation services in addition to the mandated hospital and community based clinical services. They may also contract non-government organisations (including specialist providers targeting Māori people) to deliver mental health services, including vocational rehabilitation programs. The Accident Compensation Corporation is another relevant government owned agency. It both directly provides and purchases vocational rehabilitation for people who are currently employed but not attending their workplace because of injuries, a definition which includes psychological injury.

This means that vocational rehabilitation programs for people with psychiatric disabilities can be purchased by three government departments or agencies, and by up to 20 District Health Boards, all of whom can have different primary objectives, different contracting methods, and may report to different government Ministers. Although this can benefit providers through not being reliant on a single contracting source, a more standardised contract would be more efficient to administer, provided it did not hinder any particular evidence-based practices and did not create gaps in services.

**Strengths and limitations in contracting vocational rehabilitation**

Of the evidence-based vocational rehabilitation programs that are established in seven of the 20 District Health Boards in New Zealand, six are purchased through activity-based contracts between non-government employment providers and the regional District Health Board and one is purchased and delivered by the District Health Board. An advantage of enabling District Health Boards to undertake direct contracting is that this keeps the focus on adults with psychiatric disabilities currently
in contact with public mental health services and can also encourage coordination between mental health care and vocational services (Drake et al., 2003; OECD, 2015b) (see Table 16). The importance of close co-ordination has emerged over the past 20 years as a leading principle of evidence-based vocational rehabilitation, that when implemented well can alone improve vocational outcomes for people with psychiatric disabilities (Bond et al., 2012c).

Although the Ministry of Social Development has previously purchased evidence-based vocational rehabilitation for people with psychiatric disabilities, this is not a well-utilised funding mechanism for these types of intensive employment assistance programs. One possible reason is because Ministry of Social Development contracts often have contractual conditions that limit access to, as well as hinder the development of evidence-based practices and may inadvertently encourage a shift to helping only those with the least severe psychiatric disabilities (see Table 16). The strengths and limitations of the contracts offered by these two main government purchasers of vocational rehabilitation for people with psychiatric disabilities are outlined in Table 16, along with suggested changes to the contracts which would incentivise the provision of evidence-based vocational rehabilitation.
Table 16. Reviewing contracts offered by two government purchasers of vocational rehabilitation for people with psychiatric disabilities

<table>
<thead>
<tr>
<th>Type of contract</th>
<th>Strengths</th>
<th>Limitations</th>
<th>Possible solutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. District Health Boards: activity-based contracts.</td>
<td>S1. Contracts target all working-age adults who are in contact with public mental health services. This encourages providers to target services to the people most in need of intensive employment assistance.</td>
<td>L1. Outcome targets are usually set well below what is expected from evidence-based vocational rehabilitation so there is no performance incentive to apply evidence-based practices.</td>
<td>PS1. Contracts could specify that evidence-based practices should be applied. Providers could be required to report on the quality of the implementation of these practices.</td>
</tr>
<tr>
<td></td>
<td>S2. A provider can offer on-going in-work support for as long as the person is under treatment from the public mental health</td>
<td>L2. Payment is not usually linked to outcomes.</td>
<td>PS2. The existing contract between the Ministry of Health and District Health Boards could specify that public mental health services should where possible be integrated with, or closely coordinated with employment assistance services.</td>
</tr>
<tr>
<td>2. Ministry of Social Development: results-based contracts.</td>
<td>S1. Results-based contracts can incentivise providers to implement specific principles of evidence-based vocational rehabilitation.</td>
<td>L3. There are no incentives for the public mental health services to integrate with employment assistance services to deliver joint services which research shows are the most effective.</td>
<td>PS3. The Ministries of Health and Social Development and the Accident Compensation Corporation could agree on an integrated purchasing method and amend contract terms to include people with psychiatric disabilities under the care of general practitioners.</td>
</tr>
<tr>
<td></td>
<td>S2. Some contracts offer an independent assessment to determine each person’s need for support after one year of stable employment.</td>
<td>L4. People of working-age with psychiatric disabilities and who are receiving mental health services through their general practitioner are not usually covered yet this is a primary target group.</td>
<td>PS1. Other access pathways could be added to include referrals from independent health professionals and self-referrals.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L2. Some eligibility criteria can be inappropriate such as limiting the program to people with less severe mental health conditions, or to those seeking employment of at least 15 hours per week.</td>
<td>PS2. All existing and new contracts could be assessed to determine the extent to which they facilitate or inhibit purchasing of evidence-based vocational rehabilitation. Amend the contracts accordingly.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L3. Services to working-age adults with psychiatric disabilities who receive mental health services through their general practitioner are not usually covered yet this is a primary target group.</td>
<td>PS3. The Ministries of Health and Social Development and the Accident Compensation Corporation could agree on an integrated purchasing method, and on the principles of purchasing to encourage the implementation of evidence-based practices.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L4. Results-based contracts may incentivize a focus on the less important outcomes of vocational rehabilitation for people with psychiatric disabilities. This can cause providers to shift to assisting those with the most incentives attached.</td>
<td>PS4. Amend the fee structure so that only the most important employment outcomes are rewarded. Duration of employment needs to be a primary outcome while the number of hours worked should be a secondary outcome.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PS5. Create incentives within all contracts for providers to continue in-work support for more than one year if needed.</td>
</tr>
</tbody>
</table>
### Table 17. The challenges, implications, and possible solutions for expanding the implementation of evidence-based vocational rehabilitation

<table>
<thead>
<tr>
<th>Challenge</th>
<th>Implications</th>
<th>Possible solutions</th>
<th>Responsible</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. A lack of national data on labour force activity by disability type and health condition classification.</td>
<td>11. It is unclear how to identify candidates who most need intensive forms of vocational rehabilitation, namely those who face the highest levels of labour force disadvantage and disabilities in order to match to the most intensive vocational rehabilitation programs.</td>
<td>PS1. Add questions to existing surveys. PS2. Pool existing data using the Integrated Data Infrastructure. PS3. Design a new survey of general health, mental health, and labour force activity. PS4. Review the minimum wage exemptions for people with psychiatric disabilities.</td>
<td>DHBs, MoH, SNZ</td>
</tr>
<tr>
<td>2. The minimum wage exemptions for people with disabilities.</td>
<td>11. Traditional vocational rehabilitation, in the form of sheltered workshops, can continue. 12. Participation in the open labour market is discouraged.</td>
<td>PS1. Ensure the new disability outcomes framework and action plan recognise the employment assistance needs of people with psychiatric disabilities, specifies the need for EBPs, and the preference for more effective programs over less effective or non-effective ones. PS2. Revise welfare policy to include the EBPs for effective vocational rehabilitation and include these programs in its strategies to reduce the long-term welfare dependency of people with psychiatric disabilities.</td>
<td>MBIE</td>
</tr>
<tr>
<td>3. Current national disability and welfare strategies have no reference to the existence of EBPs in vocational rehabilitation for people with psychiatric disabilities.</td>
<td>11. Decisions on what vocational rehabilitation programs to purchase are left to local funders or providers who may not have up to date information on EBPs. 12. People with psychiatric disabilities are treated as a homogeneous group with all people with disabilities. 13. The numbers of people with psychiatric disabilities who are in receipt of welfare benefits is rising.</td>
<td>PS1. Expand the Auckland region technical assistance program to assist with the further implementation of evidence-based vocational rehabilitation particularly in those regions which have identified this as a priority area in their annual plans.</td>
<td>MSD, ODI, MoH, MSD</td>
</tr>
<tr>
<td>4. Current health policy does not provide a sufficiently detailed implementation framework for EBPs.</td>
<td>11. The availability of evidence-based vocational rehabilitation relies on the knowledge of local funders and/or providers.</td>
<td>PS1. Revise health service policy so that rehabilitation activities are part of the core business of health treatment and care services. PS2. Link health funding to meeting vocational rehabilitation outcomes.</td>
<td>MoH</td>
</tr>
<tr>
<td>5. Rehabilitation is not currently an outcome measure of health treatment and care services.</td>
<td>11. There is no incentive for health funders to purchase evidence-based vocational rehabilitation. 12. There is no incentive for mental health services to integrate with employment services.</td>
<td>PS1. Identify the lead agency for coordinating health and welfare policy in relation to purchasing vocational rehabilitation for people with psychiatric disabilities. PS2. Trial a coordinated investment approach to purchasing evidence-based vocational rehabilitation.</td>
<td>MoH, MSD, ACC</td>
</tr>
<tr>
<td>6. There is an uncoordinated approach to contracting and purchasing evidence-based vocational rehabilitation for people with psychiatric disabilities.</td>
<td>11. Evidence-based vocational rehabilitation programs are available in some parts of the country and not others. 12. There are also gaps in service provision and inequities of access to available provision within regions.</td>
<td>PS1. Revise health service policy so that rehabilitation activities are part of the core business of health treatment and care services. PS2. Link health funding to meeting vocational rehabilitation outcomes.</td>
<td>MoH, MSD, ACC</td>
</tr>
</tbody>
</table>

**Notes.** EBPs= Evidence-Based Practices. DHBs=District Health Boards. MoH=Ministry of Health. SNZ=Statistics New Zealand. MSD=Ministry of Social Development. ACC=Accident Compensation Corporation. ODI=Offices for Disability Issues. MBIE=Ministry of Business Innovation and Employment. 1. Responsible government department or agency. 2. The Integrated Data Infrastructure is a national research database which pools data about people and households in New Zealand. 3. Similar to the Australian national survey of high impact psychosis, Waghorn et al., 2012.
The challenges for the expansion of evidence-based vocational rehabilitation

This analysis identifies six main policy challenges to the expansion of evidence-based vocational rehabilitation in New Zealand (see Table 1). Possible policy solutions are suggested which could form the basis of a planned cross-government approach to increasing the labour force participation of people with psychiatric disabilities.

Discussion

The policy and regulatory context of a country is an important external factor in the delivery of nation-wide evidence-based practices (King et al., 2006; Lockett et al., 2016). This systematic examination of social and economic settings has identified that in New Zealand there are policy barriers, policy conflicts, and both opportunities and challenges. While national welfare and health policy reform has commenced this has yet to translate into a sufficiently detailed purchasing framework which enables and does not inhibit, the adoption of more evidence-based forms of vocational rehabilitation.

Other challenges in New Zealand concern the different purchasing and contracting approaches coexisting in both the health and welfare systems. Where health funders choose to purchase vocational rehabilitation, there is no requirement that it be evidence-based. Where the Ministry of Social Development purchases employment assistance, the focus is often generic and about moving the most people off income support payments. Neither contracting approach encourages the close coordination of vocational rehabilitation with clinical treatment and continuing care.

Uncoordinated and suboptimal approaches to purchasing are likely to restrict the availability of evidence-based practices in vocational rehabilitation for people with psychiatric disabilities in some parts of New Zealand. This is because service availability is dependent on a provider deciding to deliver evidence-based practices, or on a local funding agency specifically requesting evidence-based practices via terms of a contract. Beginning July 2016, District Health Boards were required to report on the labour force activity of adults under the care of public mental health services. This requirement
could help generate information to assist national and regional planners to target the most intensive forms of vocational rehabilitation to those most in need.

Taken together, these social and economic policy settings do not currently support the expansion of evidence-based vocational rehabilitation programs. This helps explain why the number of people with psychiatric disabilities who are not employed is increasing and not decreasing as intended.

Possible solutions

The new national disability strategy, with its accompanying outcomes framework and action plan provides an opportunity for policymakers to specify the particular needs of people with psychiatric disabilities and to prescribe the provision of evidence-based practices in vocational rehabilitation (see Table 17).

The new health strategy (MoH, 2016b,c) focus on increasing access to evidence-based vocational rehabilitation is an important part of proposed reforms to improve health outcomes and reduce welfare dependence, but as yet return to work is not a core outcome and defining feature of effective health service delivery (OECD, 2015b, Waddell & Aylward, 2010, p. 31). This issue has also emerged in most developed countries. OECD countries agree to encourage social and economic inclusion and may even measure it as a part of a suite of health outcome measures but may not regard rehabilitation as a core responsibility of a health service (Arends et al., 2014; King et al., 2006).

There is also the complicating issue of overlapping responsibilities among government departments, where it can be unclear which agency has the primary responsibility (Arends et al., 2014). To promote better policy coordination and to increase translation of the national disability strategy’s aspirations into reality, health policy could broaden the scope of the definition of health services to encompass rehabilitation, defined as returning to productive roles in society and specify what proportion of the health budget needs allocating to rehabilitation and to vocational rehabilitation in particular (see Table 17).
Comparison with other developed countries

These policy and funding issues are not unique to New Zealand and are found in other developed countries when evidence-based vocational rehabilitation is first introduced. In the USA, evidence-based vocational rehabilitation programs are funded through a complex blending of several state and federal government sources, Medicaid and vocational rehabilitation payments (Drake et al., 2016). Drake et al. (2016) argue that to increase access to evidence-based vocational rehabilitation, a simple, nationally integrated funding stream is needed (p. 1103).

Similar problems exist in England where IPS programs are purchased mainly by regional health and social care commissioning groups and local government (Boyce et al., 2008). A recent task force on the future of mental health in England recommended that in order to improve labour force participation of people with psychiatric disabilities, a joint unit for work and health should be established (The Mental Health Taskforce, 2016). Further, that the current investment in generic employment assistance programs by the UK’s Department of Work and Pensions should be re-directed to employment programs targeted to people with psychiatric disabilities. The basis for this recommendation is that these health-led employment interventions were found to deliver better employment outcomes as well as improved health outcomes (The Mental Health Taskforce, 2016).

To support implementations in the USA and parts of Europe, a national learning collaborative has been established to facilitate the expansion of evidence-based practices (Becker, Drake, & Bond, 2014). In the USA, public mental health services and employment agencies participating in the collaborative sign business associate agreements that stipulate they will adopt the practices needed to support the integration of employment services (Swanson, Courtney, Meyer, & Reeder, 2014). In addition, some state funders offer payment to mental health services for time spent on program integration as well as time spent on formal reviews of the quality of the collaboration (Swanson et al., 2014).

In Australia, there is a single national purchaser of disability employment services but only recently has the administering department moved to officially encourage the implementation of
evidence-based approaches for psychiatric disability. In the meantime, disability employment services aware of the need for evidence-based practices are independently approaching mental health services to establish jointly administered and coordinated services. In some cases, the impetus for change was initiated by the mental health service (Waghorn & Hielscher, 2015).

**Limitations**

This analysis used multiple methods to search for relevant data and policy documents (Dunn, 2015, p. 5). While it is likely that search saturation was eventually reached this may be more attributable to the relatively small size of the country, and the small number of agencies involved, than due to a comprehensive a-priori strategy. In countries with multiple levels of government as well as multiple agencies, a more formal and structured survey of key stakeholders may be required. Snowball sampling was useful in this instance and remains a useful method in more complex settings because it enables a wider search that can capture information which might otherwise be missed by narrowly defined search criteria.

**Conclusion**

The disparity between the vocational goals of individuals with psychiatric disabilities and their actual employment status represents unmet needs for evidence-based forms of vocational rehabilitation. This unmet need could be reduced by adopting the policy adjustments suggested in a planned cross-government approach to the expansion of evidence-based supported employment for people with psychiatric disabilities. A coordinated approach to purchasing evidence-based employment assistance for working-age adults with psychiatric disabilities could be trialled. Addressing this issue has the potential to contribute to a more inclusive society, reduce long-term welfare dependence, reduce mental health service utilisation, and increase national productivity.
Chapter Conclusions

This chapter presented a peer-reviewed report that investigated the broader New Zealand economic and social policy context. The paper provides evidence of how these contextual factors identified in the IFVR can help or hinder the availability and the performance of current evidence-based practices. The analysis found that the wider economic and social policy context is currently impeding the availability of IPS programs in New Zealand. In addition, even where evidence-based IPS programs are available, the context is limiting the ease and quality of implementation.

The analysis proposed a set of policy adjustments that if taken up could improve the availability and implementation of evidence-based practices. The analysis used the operationalisation of implementation fidelity in New Zealand to identify conflicts with existing policy. This information is then used to determine any policy adjustments that could further improve program implementation quality. The other eight dimensions of implementation quality would also benefit from similar operationalising within a specific social and economic policy context.

Chapters 6, 7, and 8 seek to partially operationalise the eight dimensions of implementation quality in the IFVR, to support the construction of a new tool for measuring the quality of implementation and program delivery more broadly. It is anticipated that any new tool would complement existing measures of program fidelity.
Chapter 6: Operationalising the Implementation Framework for Vocational Rehabilitation (IFVR)

Introduction

Study 4: The construction of a general measure of implementation quality and program delivery

As discussed in Chapter 2, to improve program performance the science and practice of vocational rehabilitation for people with severe mental illnesses has largely focused on: (a) improving adherence to evidence-based practices, implementation fidelity; and (b) enhancements to known evidence-based practices. Chapter 3 reported a systematic review and meta-analysis which examined the predictive validity of the two measures of implementation fidelity for current evidence-based practices in vocational rehabilitation, namely the 15-item and the 25-item IPS fidelity scales. Fidelity, as measured by the existing IPS fidelity scales, accounts for a small amount (16.8%) of the variation in program performance (Lockett et al., 2016). These findings align with previous investigations of the predictive validity of the scales (Becker et al., 2001b; Kim et al., 2015). Findings from the study reported in Chapter 3 imply that there are other factors, in addition to implementation fidelity, that explain the performance of vocational rehabilitation programs. If these factors can be identified and measured then it is highly likely that they will eventually make a positive contribution to increasing vocational rehabilitation program performance.

Chapter 4 reported a review and synthesis of the theoretical and empirical literature on psychosocial program implementation to identify and prioritise candidate influences for improving program performance. This review resulted in the development of a new Implementation Framework for Vocational Rehabilitation (IFVR) which broadens the conceptualisation of implementation beyond implementation fidelity (Lockett et al., 2018a). The IFVR identifies eight candidate dimensions of implementation quality, in addition to program fidelity, considered likely to directly influence program performance (see Figure 9). Thirty-two contextual factors are also identified. These contextual factors are considered to have an indirect influence on program performance.
Chapter 5 conducted a systematic examination of the wider social and economic policy context to identify how this wider context is helping or hindering the implementation of current evidence-based practices in vocational rehabilitation for people with severe mental illnesses (Lockett, Waghorn, & Kydd, 2018b). This examination used the operationalisation of implementation fidelity in New Zealand to identify conflicts with existing policy, and to determine any policy adjustments that could further improve program implementation quality. The other eight dimensions of implementation quality would also benefit from similar operationalising within a specific social and economic policy context, once valid and reliable measures for these dimensions have been constructed.

This chapter and Chapters 7 and 8 seek to address the fourth and final research question of this investigation: Can the new IFVR be operationalised to construct a general tool for measuring the quality of program implementation and program delivery to complement measures of program fidelity?

This chapter (Chapter 6) describes the methods and findings from the conceptual analysis for six of the dimensions of implementation quality identified in the IFVR. The findings in respect of the conceptual analysis of the seventh and eighth dimensions, employment specialist expertise and participant responsiveness, are reported and discussed separately in Chapters 7 and 8 because of the greater volume of previous research related to these two dimensions.

Each conceptual analysis targets the measurable attributes of each of the dimension of implementation quality considered most relevant to improving the performance of vocational rehabilitation programs. The result is a partial operationalisation of each construct which can inform the writing of scale items needed before progressing to empirical testing.

Three research questions guided the conceptual analysis of the dimensions of implementation quality in the IFVR.

1. Can discrete dimensions of program implementation quality, other than fidelity to existing evidence-based practices, be identified and described sufficiently to enable valid and reliable measures of each dimension to be developed?
2. Can these dimensions be operationalised into a general measure that can be successfully applied to vocational rehabilitation programs?

3. Can these dimensions of implementation quality be generalised to other forms of psychosocial interventions?

Figure 9. The nine dimensions of implementation quality in the IFVR

Notes. IFVR=Implementation Framework for Vocational Rehabilitation (Lockett, Waghorn, & Kydd, 2018a). There are eight dimensions of implementation quality in the ellipse which along with fidelity are hypothesised to have the strongest, direct influence on achieving high program performance. The interrelationships between the dimensions of implementation quality are not depicted in Figure 9. The IFVR also hypothesises that there are 23 contextual factors considered to be largely having an indirect influence on program performance. These contextual factors are also not depicted in Figure 9. For the contextual factors see Figure 7 on page 93.

Methods

The principles of classical psychometric scale development (Churchill, 1979; Clark & Watson, 1995; Loevinger, 1957) guided the construction of a new, broader measure of program implementation quality. The IFVR developed in Chapter 4 provides the underpinning theoretical model where the candidate constructs and their interrelationships are articulated (Cronbach & Meehl, 1955; Lockett et al., 2018a). Good articulation of the underpinning theoretical framework coupled with a thorough conceptual analysis of each of the relevant constructs are important first steps in establishing construct validity. These steps help to link the conceptual realm to the observable realm (Clark & Watson, 1995).
To construct a broader more general measure of program implementation quality and program delivery, a seven-step process was followed (see Figure 10). This steps supported a logical and systematic unpacking of each dimension of implementation quality and the measurable attributes considered likely to make up the concept. Table 18 provides an explanation of the main terms used in the conceptual analysis.

Table 18. The terms used in the partial operationalisation of the IFVR

<table>
<thead>
<tr>
<th>Term</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimension of implementation quality</td>
<td>The concept of implementation quality in the IFVR is hypothesised to consist of at least nine dimensions. The nine dimensions include implementation fidelity.</td>
</tr>
<tr>
<td>Measurable attributes</td>
<td>Each of the eight dimensions of implementation quality is hypothesised to consist of a set of measurable attributes. Those selected are those attributes, or indicators, that are thought to best represent the conceptual space for that dimension.</td>
</tr>
<tr>
<td>Targeted statements</td>
<td>Based on the review of available evidence, the attributes most likely to improve program performance were prioritised in the targeted statements. These targeted statements help to develop an item pool for each measurable attribute.</td>
</tr>
</tbody>
</table>

Notes. IFVR=Implementation Framework for Vocational Rehabilitation (Lockett, Waghorn, & Kydd, 2018a).

**Step 1: Developing the construct specification**

This first step in the conceptual analysis involved conducting a literature review for each of the eight dimensions of implementation quality to map the broad conceptual space for each. This helped to gather ideas and examples of how others have described and measured the construct, as well as closely related constructs. The review of the vocational rehabilitation and broader psychosocial literature helped to clearly conceptualise each of the dimensions of implementation quality, an important precursor to the development of an operational construct specification (Clark & Watson, 1995; Proctor, Powell, & McMillen, 2013). A construct specification defines the domain of the construct and specifies what is in the construct, as well as what is not (Churchill, 1979). This means that in addition to reviewing the literature on similar constructs it is also important to review the literature around entities
from which the target measure should be distinguished. At the same time, any potential overlaps with other dimensions of implementation quality and contextual factors in the IFVR can be documented.

To identify the relevant literature, search terms used the construct name alone, i.e. ‘program evaluation’ or ‘technical assistance’, and then a search was conducted combining the construct name with one or more of the following: ‘vocational rehabilitation’, ‘supported employment’, ‘improving program performance’, ‘program effectiveness’, ‘program outcomes’, or ‘best practices’. Google Scholar was the primary database used supplemented with searches of Medline, PubMed, and PsycINFO. As relevant articles were identified and reviewed their reference lists were searched to identify further potentially relevant literature. Searches were then extended to find articles citing the relevant articles that had already been identified. Priority was given to articles from the vocational rehabilitation literature.

**Step 2: Identifying the measurable attributes of each construct**

The process of developing the construct specification assisted with a logical process to identify the key attributes which capture the conceptual boundaries of each contributing construct. For example, the dimension of implementation quality, technical assistance was conceptualised as a collaborative process of providing regular and targeted information, and expert support to the program. This was theorised as consisting of a set of attributes covering what was delivered, to whom, how frequently, for how long, and in what way.

A second literature review was conducted focusing on these initial ideas for the attributes of each construct. Search terms in this step reflected these attributes. This enabled a more in-depth mapping of the conceptual space for each dimension of quality than the first review, and led to the identification and prioritisation of potential measurable attributes of each construct. The evidence identified in the literature reviews was classified into three categories. Category A: The evidence supports the inclusion of the attribute and its effect on program performance. Category B: The evidence does not support the inclusion of the attribute and its effect on program performance. Category C: No
empirical evidence was identified, but authors have written about the potential for the attribute to influence program performance.

**Step 3: Identifying existing relevant measures**

To avoid duplicating the work of previous researchers it is important to investigate how others have operationalised the construct, or closely related constructs (Clark & Watson, 1995). This third step involved identifying and reviewing existing measures to see if they were relevant to a new, more general measure of implementation quality. Priority was given to measures used in the vocational rehabilitation literature. This review of published measures also identified other potential measurable attributes, and so the process used in step two was repeated (see Figure 10).

**Step 4: Refining the construct specifications and measurable attributes**

The fourth step involved refining each construct specification, and the measurable attributes of each construct, based on the attributes most likely to be linked to improving program performance. For example, going through steps one to three initially identified seven attributes of technical assistance. However, based on the attributes that had the strongest evidence base, the list was refined to a logical set of five candidate measurable attributes of technical assistance that evolved from the literature as most likely linked to program performance. 1. Nature of the technical assistance. 2. Type of technical assistance. 3. Frequency of technical assistance. 4. Duration of technical assistance. 5. Style of technical assistance.
Step 1: The development of an operational specification for each construct.

Step 2: Identification of the measurable attributes making up the conceptual space for each construct.

Step 3: Identification of existing measures relevant to each measurable attribute.

Step 4: Further refinement of the operational specification and measurable attributes of each dimension.

Step 5: The generation of targeted statements for each of the measurable attributes.

Step 6: A process of scale rationalisation through a series of iterative steps:
- A review of the weight of evidence for each dimension.
- A review of all descriptors and removal of ambiguous statements.
- The clarification of the meaning of each measurable attribute.

Step 7: The development of a measurement item pool.

Figure 10. The steps employed in the conceptual analysis for each dimension of implementation quality in the IFVR
Step 5: Generating targeted statements for each attribute

Once this refinement process was completed the literature for each remaining attribute was re-examined. This process was conducted to improve the logical structure and face validity of each construct and its attributes and to identify any further evidence supporting the inclusion of specific attributes. This step also assisted with the generation of the targeted statements for each attribute. For example, one of the measurable attributes of technical assistance was hypothesised to be the nature of the technical assistance provided. The review identified that to be effective, technical assistance should be driven by the identified needs of the program. “The technical assistance is driven by the identified needs of the program” thus became the targeted statement for the measurable attribute, “the nature of the technical assistance”.

Step 6: Scale rationalisation

The primary objective of this step was to map the measurable attributes against the evidence to improve the face validity of the construct and the targeted measures, and to improve the utility of the scale. This involved a review of the weight of evidence for each dimension, a review and consolidation of all descriptors, removal of ambiguous statements, and clarification of meaning as needed. Any dual subject statements about an attribute were also split into single subject statements which could be observed without ambiguity. A table, summarising each measurable attribute and target statements with the corresponding evidence for the attribute, was also developed as part of this process of scale rationalisation.

Step 7: The development of a measurement item pool

The process of rationalising each target statement was an important pre-cursor to the development of a set of measures (the measurement item pool) for each targeted statement. This was possible because the preceding steps enabled a clear scope and range of content to be established for each construct. The content of the initial item pool was deliberately overinclusive. The wording of each candidate item was given careful attention to ensure it was non-ambiguous, objective, simple, and
This step was also an iterative process involving several reviews and re-writing of candidate items.

At the same time consideration was given to the response format. For example, whether to use poly categories to classify responses e.g. 3,2,1; to use dichotomous categories, i.e. yes-no; to use a Likert approach to ratings with three or more ranked response options; or to use a combination of these approaches depending on the attribute being targeted. Consideration was also given to sample questions that program leaders could use to seek evidence that the target attribute is present or absent. Steps one to seven were repeated for each of the eight dimensions of implementation quality in the IFVR.

**Results**

The partial construct operationalisation resulting from the conceptual analysis of six dimensions of implementation quality in the IFVR are reported in sequence. 1. Technical assistance. 2. Program evaluation. 3. Removal of non-evidence-based practices. 4. Complementary programs. 5. Program intensity. 6. Program delivery.

Each of the six results section begins by summarising the available evidence on the association between the construct and program performance. Next, the operational construct specification is presented along with the measurable attributes considered the best candidates for capturing the conceptual boundaries of the construct. The subsections that follow review the literature for each candidate measurable attribute and give examples of how others have attempted to conceptualise and measure the attribute. Suggestions are made of how program leaders could apply the findings to improve program performance.

As explained in Step 6, for each dimension of implementation quality a table synthesising the findings was developed. The table summarises the results from the iterative process of refining and rationalising the candidate measurable attributes into a set of targeted statements and suggested measurement items (an item pool). The table also provides suggested questions and a scoring method for program leaders to use to gather information to better describe, define, and assess the construct in
the context of vocational rehabilitation for people with severe mental illnesses. The final column in the table summarises the supporting evidence for each measurable attribute. The evidence is classified by the three categories of strength of evidence. Category A: The evidence supports the inclusion of the attribute and its effect on program performance. Category B: The evidence does not support the inclusion of the attribute and its effect on program performance. Category C: No empirical evidence was identified, but authors have written about the potential for the attribute to influence program performance.
Technical Assistance

In vocational rehabilitation, as with the implementation of other psychosocial programs, the provision of dedicated expert advice and guidance, in addition to the personnel who are managing and delivering the program, has been found to improve overall program performance (Fixsen et al., 2013; Torrey et al., 2012). However, the provision of technical assistance is an emerging activity in the field of implementation science, and there was limited evidence and standards for guiding how to define technical assistance and what constitutes effective technical assistance (Chinman et al., 2005; Hunter et al., 2009; Salyers et al., 2007; West, Clapp, Averill, & Cates, 2012).

Technical assistance is usually composed of a mix of instruction, skills training, and transmission of working knowledge. Technical assistance can also involve the analysis and transfer of data (United Nations Educational, Scientific and Cultural Organization (UNESCO), 2017). For the purposes of improving the performance of vocational rehabilitation programs, technical assistance can consist of consultation and advice from an external member of staff, or team to the program who has excellent knowledge and expertise in relevant evidence-based practices. The consultation and advice can be provided on-site, by phone, or by email (Bond, Drake, Becker, & Noel, 2016a).

Technical assistance and program performance

Technical assistance has been linked to program performance. West et al. (2012) conducted a systematic review to assess the effectiveness of technical assistance. The authors reviewed the published literature from 2000 to 2010 and the unpublished grey literature from 2010-2011. Twenty-three published articles met their inclusion criteria, but no reports from the grey literature met their inclusion criteria. Twenty-one of the 23 included studies came from the USA. Key findings were that: (a) needs assessment helps to focus the technical assistance; (b) technical assistance staff who have practical experience, are accessible, culturally competent, and have good communication skills are preferred by program staff; (c) technical assistance is better if it involves all stakeholders in assessing needs and planning activities; and (d) collaboration and continued two-way interaction between technical assistance staff and program staff can increase the evaluation capability of program staff.
In a review of the literature on techniques to improve community capacity to implement prevention programs, technical assistance was found to assist with tasks and activities which involved planning, implementation, and organisational maintenance. However, technical assistance did not improve the organisation’s capacity to conduct evaluation (Chinman et al., 2005). In this example, it may be that improving evaluation capacity was not specifically targeted in the technical assistance that was provided, or that it was in the assistance provided, but it was not effective.

Some evidence for the effectiveness of the provision of sustained, focused, and responsive technical assistance on program performance is provided by analysis of the IPS programs in the USA. In the USA the provision of technical assistance to IPS programs has become a routine and ongoing part of IPS establishment and sustainment to improve program adherence and program performance (Becker et al., 2014; Becker, Lynde, & Swanson 2008; Bond, Drake, Becker, & Noel, 2016b; Johnson-Kwochka, Bond, Becker, Drake, & Green, 2017; Margolies et al., 2015).

In 2001, the Johnson & Johnson-Dartmouth Community Mental Health Program was established. This learning collaborative began in three states to support the implementation of evidence-based practices in vocational rehabilitation (Becker et al., 2014). The learning collaborative now involves 24 states and five countries (The IPS Employment Center, 2018). The focus of the collaborative is on the shared collection and monitoring of program data, exchanging ideas and information, and providing training and technical assistance (Becker et al., 2014). Technical assistance is provided to sites through IPS fidelity reviews and full-time IPS trainers. IPS trainers are skilled in providing assistance based on best practice and emerging knowledge. Technical assistance is provided through on-site visits, mentoring, facilitating training, and through regular telephone communication. Specific technical tools have also been developed for sites, such as a web-based fidelity software to assist sites to carry out fidelity reviews, and online training courses for program staff, particularly targeted at employment specialists and their supervisors (Becker et al., 2008, 2014). In addition, technical assistance is offered to help states and countries address policy and funding barriers, in order to promote and better implement evidence-based practices. A similar approach has been taken to
support fidelity of implementation and improve performance across a learning collaborative of vocational rehabilitation programs in New York State. Trainers provide on- and off-site technical assistance, provide program specific consultation, and lead informational webinars (Margolies et al., 2015).

A 2015-2016 survey of all USA mental health and vocational rehabilitation state agencies found that of the 38 states that had at least one IPS program, 23 had either a technical assistance centre or employed a full-time IPS trainer (Johnson-Kwochka et al., 2017). Whilst states participating in the learning collaborative were more likely to have implemented higher levels of technical assistance, states outside the learning collaborative also developed technical assistance centres and employed state trainers. The study did not examine the impact of technical assistance on the number, quality, or outcomes from IPS programs.

Evidence of the link between technical assistance and program fidelity and program reach can be found through analysis of IPS program implementation in New Zealand. A two-year pilot aimed to understand the influence of adding dedicated expertise in the implementation of evidence-based practices, through one full-time IPS trainer, to two existing IPS programs. The IPS trainer provided on-site mentoring to employment specialists, training to clinicians on evidence-based practices in vocational rehabilitation, and was available on a weekly basis providing advice and guidance to all clinical and vocational rehabilitation program staff (Te Pou, 2017). Adherence to IPS practices were assessed at the outset of the pilot, using the 25-item fidelity scale, and after one year. Significant increases in total fidelity scores were found in both programs. Program A increased from a baseline score of 83, to 95 out of 125 ($p<.001$), and program B increased from 93, to 103 out of 125 ($p<.001$). There was also a steady increase in the total annual referrals made to and accepted by the vocational rehabilitation program. Whilst the study collected data on the proportion of employment commencements within each program, comparisons with program outcomes prior to the technical assistance being added were not made (Te Pou, 2017).
Technical assistance has also been linked to IPS program maintenance. A study examining the sustainment of IPS practices across the USA learning community surveyed sites use of technical assistance. In 2012, 89 of 122 IPS programs (73.0%) had on-site technical assistance in the past year, in 2014 this decreased to 83 programs (68.0%). Ninety-six percent of the sites in the study were sustaining IPS implementation two years on and the average quarterly employment outcomes for the sites also improved over the period. The authors considered technical assistance a key aspect of successful program sustainment and performance (Bond et al., 2016b). The nature, quality, and duration of technical assistance was not examined in sufficient detail to explain how the technical assistance was contributing to program performance and program maintenance.

A more precise description of the construct technical assistance is needed so that future research can better investigate the link between technical assistance and vocational rehabilitation program performance.

**Conceptualising technical assistance**

There is some consensus that in order to achieve improvements in program performance, technical assistance should have: (a) clear objectives based on issues identified through program evaluation; (b) be tailored to the specific needs of the program; (c) assist program staff and other stakeholders to apply new knowledge and technology; (d) be provided at an adequate frequency and defined period of time to achieve the desired results; and (e) be delivered in an accessible way (Gibbs, Hawkins, Clinton-Sherrod, & Noonan, 2009; Hunter et al., 2009; Priest, Lockett, & Grove, 2012; Wandersman, Chien, & Katz, 2012; West et al., 2012).

For the purposes of improving the performance of vocational rehabilitation programs for people with severe mental illnesses, technical assistance can therefore be conceptualised as a collaborative process of providing regular and targeted information, and expert support to the program. The focus of the technical assistance is on the implementation of evidence-based practices known to work in other programs. Technical assistance can be theorised as consisting of five measurable attributes, covering what was delivered, to whom, how frequently, for how long, and in what way.
1. Nature of the technical assistance.

2. Type of technical assistance.

3. Frequency of technical assistance.

4. Duration of technical assistance.

5. Style of technical assistance.

Technical assistance should be distinguished from financial assistance (UNESCO, 2017) and from tools, training, and other forms of quality improvement (Wandersman et al., 2012).

**Nature of the technical assistance**

A technical assistance system should continually assess the needs of a program or service and use this information to prioritise and tailor the technical assistance provided (Mitchell, Florin, & Stevenson, 2002). Technical assistance supports program staff to adapt and apply the evolving base of experience, knowledge, technology, and innovative practices (West et al., 2012). The ultimate aim of technical assistance is to improve program quality, effectiveness, and efficiency, to maximise the project impact (West et al., 2012). Program effectiveness is about real-world outcomes in a complex environment (Glasgow et al., 2003). Program efficiency is about the ratio of outputs to inputs. Program efficiency can be expressed in terms of timeliness, costs benefits and other indicators showing how minimum resources were used to attain maximum outcomes (Cimera, 2000).

To ensure that the technical assistance is driven by and targeted to the needs of the program, it is important that program leaders have clearly defined the problem and the success criteria for addressing it. There should also be a specific action plan for technical assistance and program leaders should check that the action plan is being implemented as planned. Ongoing evaluation against the pre-agreed success criteria will assist program leaders to identify if the original problem has been resolved, or if further technical assistance is needed.

**Type of technical assistance**

Technical assistance can be provided at different levels, for example, on-site to employment specialists, managers and clinicians, or at a systems-level such as advice to realign funding, policy, and
performance systems (Priest et al., 2012; Salyers et al., 2007; West et al., 2012). Technical assistance can be offered at any stage of program development, at program start-up, operation, and maintenance. It can be specific to the innovation being implemented but can also offer more general capacity building, such as leadership development, management support, access to resources, or improving administrative support (Wandersman et al., 2012).

Technical assistance can be provided to individuals or groups of individuals, for example, via a technical assistance centre (Torrey et al., 2012), and to a range of people (a) program staff, (b) service users, (c) families, (d) program administrators, and (e) program leaders. The type of technical assistance should be informed by the identified problem, and then targeted at the appropriate staff, at the appropriate level (Chinman et al., 2005).

Based on the identified problem, program leaders should check that all the relevant program staff and stakeholders were involved in the process. This is particularly important for building the capacity of the program to address similar problems that arise in the future. It is also important that the technical assistance personnel have sought feedback from relevant staff and stakeholders on what support they felt was needed, and whether the technical assistance activities or guidance are meeting their needs.

**Frequency and duration of technical assistance**

Tailoring the nature and type of technical assistance should be accompanied by an assessment of the appropriate frequency and duration of assistance to address the identified problem (West et al., 2012). However, the evidence of just how much technical assistance is enough, is mixed (Wandersman et al., 2012). A study investigating the role of technical assistance on implementing prevention programs found that technical assistance hours were correlated with improvements in implementation. The areas which received the most technical assistance hours showed measurable improvements, conversely the areas which received the lowest hours of technical assistance had no measurable improvements (Chinman et al., 2008). Other studies have not found a relationship between the amount of technical assistance provided and program performance (Wandersman et al., 2012; West et al., 2012).
2012). It has also been argued that sustained technical assistance is better than time-limited, or temporary technical assistance (Spoth, Clair, Greenberg, Redmond, & Shin, 2007).

In prevention programs, Hunter et al. (2009) found that each program received between one and three hours of technical assistance a week, over one to two years. The study also found that it was important that the time staff spent preparing for and meeting with technical assistance staff was recognised and resourced. It was also important that an exit strategy was developed. The exit strategy needed to consider how to transfer the skills, and the responsibilities, for the additional activities that had been provided by technical assistance personnel, to program staff.

In health care, evidence is emerging of the positive influence on patient outcomes when psychiatrists provide technical assistance to primary care teams to discuss and assess patients with depression, particularly patients not recovering as expected (Kates & Mach, 2007). When this technical assistance is frequent, and meetings are face to face, the psychiatrists can also develop the capacity of primary care clinicians to more effectively treat and manage patients with mental illnesses (Kates & Mach, 2007).

In the USA IPS learning collaborative, technical assistance consists of one full-time technical assistance staff providing dedicated support to no more than two vocational programs for four years (Becker et al., 2011b). The technical assistance for programs focuses on improving fidelity to known evidence-based practices and the assistance tailored accordingly, with site visits reducing once sufficient levels of program fidelity are obtained by the programs (Becker et al., 2014).

Program leaders can find out about the frequency and intensity of technical assistance by asking technical assistance personnel and program staff when they last had contact with each other. Program leaders can also ask about what happened during the contact. Technical assistance staff can also be asked to keep a log of their weekly activity from the outset of providing the assistance. The log could provide details on the types of contact with program staff, how long the contacts lasted, and would also provide data on the duration of the assistance. Program leaders can use the information gathered to assess whether this is sufficient to address the identified needs of the program.
**Style of technical assistance**

Technical assistance should be focused, include follow up, and be sensitive to the context in which it is being provided. It can involve both pull and push activities (West et al., 2012). Pull activities are initiated by the program implementers because they believe the program needs them, whereas push activities are initiated by the technical assistance team to help the implementers adopt the innovation. Technical assistance personnel act as facilitators, affecting the environment in which the innovation is being adopted, as well as working with staff to help them understand and implement the research, and change their practices (Rycroft-Malone, 2004).

In the prevention literature a framework for technical assistance, the Getting to Outcomes® (GTO) Framework, has been developed to guide the process of using evidence-based practices to improve the planning, implementation, and evaluation of programs (Chinman, Imm, & Wandersman, 2004). The main technique used is empowerment evaluation, with program staff leading the work and independent technical assistance taking the role of facilitator.

Hunter et al. (2009) used this GTO framework to examine the impact of technical assistance on two community coalitions representing six substance disorder prevention programs in the USA. The study measured technical assistance activity, by mode (in-person, phone, email), time spent, what assistance was provided, and coded written descriptions in the technical assistance staff log book designed for the study. Interviews were also conducted with program staff, and in addition staff were asked to complete a brief questionnaire about their impressions of the technical assistance provided. The study found that the majority of the technical assistance staff time was spent on communication with program staff, who particularly valued structured, brief, and direct communication (Hunter et al., 2009). The style of technical assistance was also important to staff, with program staff reporting they preferred a motivating and guiding approach. Staff also reported that technical assistance personnel should provide some structure and direction, but also listen and learn from program staff too.

Program leaders can check the style of technical assistance, that it has been delivered in the field rather than office-based, that it includes activities that respond to the program’s needs (pull
activities), as well as practices that the technical assistance personnel introduce because they are known to be effective in other similar programs (push activities). Program leaders can seek feedback from program staff on how they experience the external support. Program leaders can also observe technical assistance personnel to understand whether they are adopting a facilitative and collaborative style which suits the setting and context of the program.

**Construct operationalisation and measurement**

Table 19 outlines the five candidates for measurable attributes of technical assistance. These five candidates are considered sufficient to adequately represent the construct of technical assistance. For each measurable attribute one or more targeted statements are suggested, along with a possible item pool that program leaders could use to identify, measure, assess and improve the quality of technical assistance.
Table 19. Operationalising the construct of technical assistance

<table>
<thead>
<tr>
<th>Measurable attributes</th>
<th>Targeted statements</th>
<th>Measurement item pool</th>
<th>Method of measurement</th>
<th>Examples of questions and information to gather</th>
<th>Evidence for attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nature of the TA</strong></td>
<td>The TA is driven by the identified needs of the program</td>
<td>1. The program has clearly defined the problem. E.g. to do with fidelity, attrition, or other areas of program performance. 2. There are clear success criteria for addressing the problem. 3. An action plan for TA has been developed. 4. The action plan is implemented. 5. Evaluation demonstrates that the success criteria have been achieved. (Score 1 for each anchor that is present).</td>
<td>Interview with program staff and TA staff Review program records</td>
<td>Ask staff to describe how they identified what the TA should consist of. Look for whether the TA was driven by theory, evidence-based practices or evaluation results. Ask staff to provide an example of TA provided in the last four weeks, compare the answer with the action plan. Seek documentary evidence of clear success criteria, action plans, and evaluation.</td>
<td>Becker et al., 2008; Bond et al., 2016a,b; Chinman et al., 2005; Gibbs et al., 2009; Mitchell et al., 2002; Wandersman et al., 2012; West et al., 2012.</td>
</tr>
<tr>
<td><strong>Type of TA</strong></td>
<td>The TA is targeted to the appropriate staff</td>
<td>Based on the identified problem the TA assistance has (a) met with and involved the relevant staff and stakeholders (b) sought feedback from staff and stakeholders on the impact of the TA activities. (Score 1 for each anchor that is present).</td>
<td>Interview with program staff and TA staff Review program records</td>
<td>Ask TA personnel to provide examples of how they involve program staff. Look for who was involved i.e. delivery and strategic staff. Find out whether staff were asked what support they needed and whether regular meetings were held to check how the implementation plan was going? Find out where the TA is usually provided. Do TA personnel meet vocational staff with participants and model behaviour? Do TA staff meet vocational staff with the clinical team and model behaviour?</td>
<td>Becker et al., 2008; Chinman et al., 2005; Gibbs et al., 2009; Hunter et al., 2009; Margolies et al., 2015; Rycroft-Malone, 2004; Salyers et al., 2007; Wandersman et al., 2012; West et al., 2012; UNESCO, 2017.</td>
</tr>
<tr>
<td>Measurable attributes</td>
<td>Targeted statements</td>
<td>Measurement item pool</td>
<td>Method of measurement</td>
<td>Examples of questions and information to gather</td>
<td>Evidence for attribute&lt;sup&gt;1&lt;/sup&gt;</td>
</tr>
<tr>
<td>-----------------------</td>
<td>---------------------</td>
<td>-----------------------</td>
<td>-----------------------</td>
<td>-----------------------------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>Frequency of TA</td>
<td>The TA is of sufficient frequency</td>
<td>Based on the identified problem the TA assistance is (a) of sufficient intensity (b) of sufficient frequency. (Score 1 for each anchor that is present).</td>
<td>Interview with program staff and TA staff</td>
<td>How many hours of TA are provided each week and given the problem that is being addressed, is this sufficient? Do TA personnel work across multiple programs? When was the last contact between the program staff and TA personnel? When was it and what happened? Was this contact adequate to address the problem?</td>
<td>Becker et al., 2011b, 2014; Chinman et al., 2005; Hunter et al., 2009; Johnson-Kwochka et al., 2017; Mitchell et al., 2002; Wandersman et al., 2012; West et al., 2012.</td>
</tr>
<tr>
<td>Duration of TA</td>
<td>The TA is of sufficient duration</td>
<td>1. TA has been provided within the last year. 2. The TA has been provided for sufficient duration in relation to the problem identified. (Score 1 for each anchor that is present).</td>
<td>Interview with program staff and TA staff Review program documentation</td>
<td>Ask TA personnel and program staff when the assistance started, how long it is likely to continue for. Also check any program funding contracts or staff job descriptions which confirm length of assistance. Does the duration of TA match the size of the identified problem?</td>
<td>Becker et al., 2011b,2014; Bond et al., 2016a,b; Hunter et al., 2009; Spath et al., 2007.</td>
</tr>
<tr>
<td>Style of TA</td>
<td>The style of TA</td>
<td>1. TA is predominantly been provided in the field, rather than office based. 2. TA includes a mix of pull and push activities. 3. The TA is delivered in a facilitative style. (Score 1 for each anchor that is present).</td>
<td>Ask program staff how they experience the TA personnel. Is the style motivating and two-way in style? Does there appear to be trust and cooperation between program staff and technical assistance personnel? Do TA personnel understand the cultural context of the program and how this is relevant to the employment assistance journey?</td>
<td>Chinman et al., 2005; Hunter et al., 2009; Rycroft-Malone, 2004; West et al., 2012.</td>
<td></td>
</tr>
</tbody>
</table>

Notes. TA=Technical Assistance. 1. Green=Category A: The evidence supports the inclusion of the attribute and its effect on program performance. Red=Category B: The evidence does not support the inclusion of the attribute and its effect on program performance. Blue=Category C: No empirical evidence was identified, but authors have written about the potential for the attribute to influence program performance. 2. Start the questioning by asking if the program had any external technical support in the past 12 months (if No, do not continue and score zero for this item). However, ensure the initial questions check for other types of assistance, which may not be referred to explicitly as technical assistance, e.g. consultancy or training providers.
Program Evaluation

Program evaluation is a systematic way of collecting, analysing, and disseminating accurate and reliable information on program design, implementation, and performance (Centers for Disease Control & Prevention, 1999, 2011, p. 102). It is a method for understanding program effectiveness and efficiency where inputs, process, outputs, and outcomes are measured (Mercer, 1994). Program evaluation is not meant to be a one-off activity. Program evaluation is usually intended as an ongoing process which involves program staff as well as evaluators such that the process becomes a routine part of service delivery and quality improvement (Gibbs et al., 2009). The use and sharing of program process and outcome data are also integral parts of program evaluation (Centers for Disease Control & Prevention, 1999).

Program evaluation and program performance

Program evaluation is likely to be an important dimension of implementation quality as it has been linked to program performance (Australian Government, 2001; Gowdy et al., 2004). However there appears to be limited empirical research on the association between program evaluation and vocational rehabilitation program performance. Gowdy et al. (2004) identified differences in the attention to, and use of, outcome data in a study examining the organisational practices that differentiated higher performing vocational rehabilitation programs from lower performers. Five of nine vocational rehabilitation programs were classified as high performing programs and the four lowest performing as low performing. Performance was measured by the average proportion of participants in employment each quarter over two years. Data on program attributes were collected through on-site visits and in-depth interviews with vocational rehabilitation staff, administrators, and participants. Program leaders at all the higher performing sites discussed systematically monitoring program and participant outcomes, and in one interview the use of data for setting team goals was mentioned. In contrast, program leaders in the lower performing sites did not mention outcome data, or if it were mentioned, the tracking system was not analysed and linked to any program action planning.
Previous research in vocational rehabilitation has identified the importance of program evaluation as part of implementation quality, and attempted to measure the attributes of program evaluation, but has not yet linked program evaluation to program performance. For example, as part of the national evidence-based practices project, Finnerty et al. (2009a) developed the State Health Authority Yardstick (SHAY) to measure the facilitating conditions for implementation fidelity. The SHAY is a 15-item measure with two items relating to program evaluation: Item 13, an assessment of adherence to evidence-based practices; and Item 14, participant outcomes. The total score for each item is developed from eight behavioural anchors. A score of five for the item is given if all eight anchors are met, a score of 1 if none or just one are met. Examples of behavioural anchors include, for Item 14: “Client outcome data are used routinely to develop reports on agency performance”; and “Agency performance data are given to programs and used for purposes of quality improvement” (Finnerty et al., 2009b, p. 15). The SHAY was tested in seven states in the USA. Whilst the scale was found to have promising face validity, as judged by the state leaders in the study, further empirical testing was considered needed. In addition, the items relating to program evaluation, fidelity assessment, and individual outcomes were not highly rated by state leaders as important activities at a state level (Finnerty et al., 2009a). It is possible therefore that the measurement of program evaluation is more important at the program delivery level, than at a state level.

**Conceptualising program evaluation**

For the purpose of operationalising program evaluation in vocational rehabilitation programs for people with severe mental illnesses, program evaluation can be conceptualised as the regular and systematic collection, analysis, and feedback of information about program processes and program outcomes which can improve program effectiveness and program efficiency. Because program evaluation requires having people with appropriate knowledge and skills, measurement of program evaluation can be divided into two stages: Program evaluation capability and program evaluation practices.
Program evaluation capability

Program evaluation capability assumes that there is a sufficiently defined program that can, in principle, be evaluated. Programs that are not sufficiently defined can often not be evaluated because they do not have a program purpose, a statement of scope and intention (Mercer, 1994), a defined target group, program logic (Owen & Rogers, 1999, p. 13), or defined and measurable outcomes. The core measurable attributes of program evaluation capability are hypothesised to be

(a) the presence of a sufficiently described and defined program,

(b) the presence of a feasible evaluation plan,

(c) the means and resources to implement a feasible evaluation plan.

Program description and definition

The description and definition of a program should involve a clear description of the aims and objectives, including the underpinning program logic, and the presence of an evaluation plan (Australian Government, 2001; Mercer, 1994). The development of a clear program description sets the frame of reference for all subsequent evaluation work, is important in assisting programs to connect to their intended program steps, and a clear description allows comparison of evaluation findings with other similar programs. Developing a program description can help program leaders to start to improve program performance, even if data is not available yet to collect and analyse (Centers for Disease Control & Prevention, 1999; Mercer, 1994).

The evaluation plan

Program evaluation requires both process evaluation and outcome evaluation (Hermann, 2002). Programs achieving higher performance were found to identify their own key performance indicators and monitor these regularly and systematically (Australian Government, 2001). Whilst this Australian study did not report on the detail of the performance indicators used by the higher performing programs, or how performance measurement was fed back to improve the operating system, the programs did report having targets around job outcomes as well as other aspects of business performance. More detailed information on the evaluation plans in operation at the higher performing
sites were not available. Such information would have been useful to understand how higher and lower performing programs differed in their approaches to evaluation.

In terms of process evaluation, it is important that the evaluation plan includes an assessment of participants’ satisfaction with the employment assistance provided (Del Valle et al., 2014; Wehman et al., 2003). Monitoring participant satisfaction with the services delivered was one of the important aspects of program delivery quality identified by Evans and Bond (2008) in a survey of the views of nine vocational rehabilitation experts and 55 vocational rehabilitation staff.

In the context of vocational rehabilitation, a sufficient plan to evaluate program effectiveness would include at a minimum an assessment of:

(a) program intakes or referrals, including the proportion accepted into the program from the total seeking to enter the program;
(b) program attrition, including the proportion of participants who drop out before receiving a sufficient dose of the assistance available, as well as participants who gain no vocational benefits;
(c) program fidelity, the extent of adherence to recognised evidence-based principles or practices;
(d) participants’ satisfaction with the employment assistance; and
(e) relevant program outcomes.

More advanced evaluation plans may attempt to capture the perspectives of other stakeholders, such as employers, family members, health practitioners, or funding contractors. Some might even focus on investigating the experiences of those who received the least benefit from the program, or who indicated any degree of dissatisfaction. Other advanced plans may include measures of program efficiency such as timeliness, return on investment, and costs-benefits analysis.

**Means and resources**

Programs also need to have the means and resources to implement an evaluation plan. This could include financial resources and staff availability and capability. Staff need time allocated to
collect, check, and analyse data, and time for implementing changes that result from the evaluation. As part of planning for program evaluation program leaders should discuss and agree a budget and resource allocation for evaluation (Centers for Disease Control & Prevention 2011, pp. 18, 21).

**Program evaluation practice**

Program evaluation practice is the extent to which the evaluation plan is being executed and that principles of good evaluation are being practised. Program evaluation practice is therefore hypothesised to involve six measurable attributes. 1. Conducting unbiased and reliable evaluations. 2. Analysing and synthesising results. 3. Using the findings to develop an action plan. 4. Communicating results to stakeholders. 5. Implementing the action plan. 6. Building staff evaluation capability.

**Conducting unbiased and reliable evaluations**

Conducting unbiased and reliable evaluation usually consists of the collection of individual-level data and the quality control of that data. Whilst many programs design evaluations to assess the quality and performance of the program, the data collected are not always reliable and unbiased (Wehman et al., 2003). Program management information systems often collect data based on point-in-time monitoring and caseload profiles that can change each quarter. In the USA, learning collaboratives have been set up for vocational programs to share good practice, collect shared measures, and collectively identify and address barriers to implementation (Becker et al., 2014). These IPS learning collaboratives include aspects of program evaluation in their operation, particularly the collection and sharing of program process and outcome data (Becker et al., 2014; Margolies et al., 2015). Sites report quarterly data on outcomes, new entrants to the program, and program attrition (Becker et al., 2014).

One of the limitations of this caseload summary approach to the collection of outcome data is that the subsequent point-in-time monitoring does not involve the same cohort of participants over time. Nor does the caseload summary control for the number of participants who leave, or who commence assistance each quarter (Lockett et al., 2016). Individual-level data is more rigorous but also more time consuming to collect. However, individual-level data is considered an essential part of effective program evaluation as it allows a defined cohort to be monitored and tracked over time.
Monitoring defined cohorts over time provides an unbiased estimate of program outcomes as it takes into account new entrants into the program (each new person has a start date), the duration of employment assistance for each individual, and monitoring defined cohorts over time enables measurement of program attrition (Lockett et al., 2016). Program leaders need to ensure that program management information systems can collect data at an individual-level to track participants over time, measure a range of employment outcomes, and record details of program entry and exit (Australian Government, 2014; Bond et al., 2012a; Lockett et al., 2016). The data need to be retained in an accurate form and kept accessible for routine analysis (Attkisson, McIntyre, Hargreaves, Harris, & Ochberg, 1974; Wehman et al., 2003). Such data sets should contain a range of measures for a comprehensive and accurate picture of program performance (Australian Government, 2014; Bond et al., 2012a; Drebing et al., 2012). Date sets should be overseen by a nominated individual who is responsible for checking the quality of the data (Centers for Disease Control & Prevention, 2011 p. 34).

**Data analysis and synthesis**

To improve performance, data need appropriate analysis and synthesis to support an understanding of the strengths and weaknesses of current program performance, and to enable the development of an action plan for continuous program improvement (Australian Government, 2001; Gowdy et al., 2004). This means that program leaders need to have in the program team at least one staff member with training in program evaluation, data collection, and data analysis, or this role needs contracting out to evaluation professionals. Ideally data analysis and synthesis ought to be done by program staff as a part of routine service delivery (Gibbs et al., 2009). With evaluation expertise within the program and dedicated resources to program evaluation, an evaluation can be conducted annually (Centers for Disease Control & Prevention, 2011 p. 21).

**Using the findings to develop an action plan**

Action plans are needed to turn data collection and analysis into action (Centers for Disease Control & Prevention, 2011, p 31). A qualitative analysis of the organisational structures and practices from four higher performing vocational rehabilitation agencies for people with disabilities in the USA,
identified being data-driven as one of the promising organisational practices (Del Valle et al., 2014). This involved using information management systems to measure return on investment, and program outcomes, and using the data collected to set program goals and monitor program performance. Effective action plans should involve a strategy to communicate a summary of the findings to key stakeholders such as families, potential participants, current participants, program leaders, clinicians, program funders, and employment program staff, as part of a quality improvement cycle (Australian Government, 2001; Del Valle et al., 2014; McHugo et al., 2009).

**Communicating results to stakeholders**

Some authors have suggested that communication of evaluation results should include well-structured information on program strengths and weaknesses, recommendations for change, and priorities for future evaluations (McHugo et al., 2007; Rapp et al., 2005). Possible measures of the extent of communication with stakeholders could include (a) frequency of outcome reporting to key stakeholders, (b) the number of stakeholders who are aware of the evaluation strategy, or (c) the number of stakeholders who have current knowledge of the most recent program outcomes.

With regard to communicating results to staff in the program, the influence of audit and feedback on professional practice and healthcare outcomes were examined in a Cochrane Collaboration review and meta-analysis by Ivers et al. (2012). The authors reviewed 70 studies comparing audit and feedback to usual care to examine the effect on clinical professional practice. A multivariable meta-analysis found that feedback was more effective when “baseline performance is low, the source is a supervisor or colleague, it is provided more than once, it is delivered in both verbal and written formats, it includes explicit targets, and an action plan” (p. 2). The effect size of the associations varied depending on the clinical behaviour being targeted (Ivers et al., 2012).

Once a summary of evaluation findings is available, program leaders could set up a planned series of meetings with key stakeholders to share and discuss the evaluation results. This could help lever additional targeted resources for the program, for example technical assistance personnel and
strengthen support and advocacy for the continuation and expansion of the program (Centers for Disease Control & Prevention, 2011 p. 35).

**Implementing the action plan**

According to Gibbs et al. (2009) the evaluation needs to drive program improvement. Communicating and dissemination of results is therefore not sufficient, evaluation teams and program staff should be proactively implementing actions from the evaluation findings that will improve program performance. Staff involvement in the evaluation process is considered important so that planned actions can be brought into routine service delivery (Centers for Disease Control & Prevention, 2011, p. 9). Program leaders should therefore be regularly assessing program activities against the actions outlined in the plan, tracking progress being made against the identified actions.

**Building staff capability in evaluation**

Building evaluation into routine practice is considered an important attribute of effective program evaluation to improve program performance (Gibbs et al., 2009). Evaluators can help program staff build evaluation into routine program delivery and plan program improvement by putting program staff in control of evaluation design and delivery. To increase the evaluation capacity of program staff, effective evaluation should include: (a) good collaboration between the evaluators and program staff; and (b) high levels of participation by program staff in evaluation activities, so that skills are transferred. Observable outcomes could be more staff conducting their own evaluations, and more staff actively using program evaluation findings to improve program performance (Gibbs et al., 2009).

**Construct operationalisation and measurement**

Table 20 outlines the three measurable attributes of program evaluation capacity and the six measurable attributes of program evaluation practice. These attributes are considered sufficient to adequately represent the construct of program evaluation. For each measurable attribute one or more measures are suggested, along with a possible item pool that program leaders could use to begin assessing the quality of program evaluation.
Table 20. Operationalising the construct of program evaluation

<table>
<thead>
<tr>
<th>Measurable attributes</th>
<th>Targeted statements</th>
<th>Measurement item pool</th>
<th>Method of measurement</th>
<th>Examples of questions and information to gather</th>
<th>Evidence for attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construct specification: The regular and systematic collection, analysis, and feedback of information about program process and program outcomes which can improve program effectiveness and program efficiency. Measurement of program evaluation consists of two parts. 1. Program evaluation capacity, the feasibility that the program can currently be evaluated. 2. Program evaluation practice, the extent that principles of good evaluation are being practised when implementing program evaluation.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Part 1. Program evaluation capability</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Program definition and description</td>
<td>A program is sufficiently well-described and defined to be capable of being evaluated.</td>
<td>1. The program must be sufficiently well-defined to be able to understand how it can cause or influence change for participants in accordance with its written aims and objectives. (Score 1 if anchor is present).</td>
<td>Program documentation</td>
<td>Review a description of the program, if available. Is the program sufficiently well-defined to be evaluated?</td>
<td>Centers for Disease Control &amp; Prevention, 2011; Australian Government, 2001; Mercer, 1994.</td>
</tr>
<tr>
<td>Evaluation plan</td>
<td>The presence of a feasible and appropriate evaluation plan.</td>
<td>An evaluation design has been proposed that covers: (a) intake or referrals (b) attrition (c) program fidelity (d) relevant program outcomes (e) participant satisfaction (f) performance targets. (Score 1 if (d) and (e) are present; score 2 if (d) and (e) and one other is present; score 3 if all are present).</td>
<td>Program documentation</td>
<td>Has a feasible evaluation plan been designed? Review the evaluation plan, including proposed measures.</td>
<td>Australian Government, 2001; Evans &amp; Bond, 2008; Finnerty et al., 2009a,b; Gowdy et al., 2004; Hermann, 2002; Wehman et al., 2003.</td>
</tr>
<tr>
<td>Resources and means</td>
<td>The program has the resources and means to implement an evaluation plan.</td>
<td>1. The program has sufficient financial resources allocated to program evaluation. 2. Relevant program staff have time allocated in their roles to spend on program evaluation. (Score 1 for each anchor present).</td>
<td>Program documentation, Interview</td>
<td>Review the program budget, if available. Ask program leaders whether there are funds and staff time dedicated to program evaluation activities.</td>
<td>Centers for Disease Control &amp; Prevention, 2011.</td>
</tr>
<tr>
<td>Measurable attributes</td>
<td>Targeted statements</td>
<td>Measurement item pool</td>
<td>Method of measurement</td>
<td>Examples of questions and information to gather</td>
<td>Evidence for attribute¹</td>
</tr>
<tr>
<td>------------------------</td>
<td>---------------------</td>
<td>-----------------------</td>
<td>-----------------------</td>
<td>-----------------------------------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>Part 2. Program evaluation practice</td>
<td>Conducting unbiased and reliable evaluations</td>
<td>The collection of individual-level data.</td>
<td>Observation of the system</td>
<td>Review the organisation’s information management system. Are individual-level data collected on participants? Individual-level vocational outcome data could include: employment status, tenure, wages, hours and types of jobs.</td>
<td>Australian Government, 2014; Becker et al., 2014; Bond et al., 2012a; Del Valle et al., 2014; Lockett et al., 2016; Margolies et al., 2015.</td>
</tr>
<tr>
<td></td>
<td>Quality control of data</td>
<td>The program collects individual-level data that is (a) accurate and defined (b) accessible (c) tracks individuals over time (d) measures individual vocational outcomes (e) measures program entry and exit. (Score 1 if (d) and (e) are present; score 2 if (d) and (e) and one other is present; score 3 if all are present).</td>
<td>Interview Program documentation</td>
<td>Ask who can access all the data in the information system, and how quality is managed? Is there an accurate data dictionary to accurately define each data item, and how it is collected? What steps are taken to manage data quality?</td>
<td>Atkinson et al., 1974; Centers for Disease Control &amp; Prevention, 2011; Drebing et al., 2012; Wehman et al., 2003.</td>
</tr>
<tr>
<td></td>
<td>Data analysis and synthesis</td>
<td>A suitable written report is available.</td>
<td>Program documentation</td>
<td>Are written evaluation reports prepared? How frequently? Review the evaluation report.</td>
<td>Australian Government, 2001; Finnerty et al., 2009a,b; Gibba et al., 2009; Gowdy et al., 2004; Wehman et al., 2003.</td>
</tr>
<tr>
<td></td>
<td>Using the findings to develop an action plan</td>
<td>A suitable action plan is available.</td>
<td>Program documentation</td>
<td>How do evaluation findings inform program delivery? Review the action plan.</td>
<td>Centers for Disease Control &amp; Prevention, 2011; Del Valle et al., 2014; Gibba et al., 2009.</td>
</tr>
<tr>
<td>Measureable attributes</td>
<td>Targeted statements</td>
<td>Measurement item pool</td>
<td>Method of measurement</td>
<td>Examples of questions and information to gather</td>
<td>Evidence for attribute</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------------</td>
<td>-----------------------</td>
<td>-----------------------</td>
<td>-----------------------------------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Communicating results to stakeholders</td>
<td>The knowledge gained is systematically communicated to relevant stakeholders.</td>
<td>The following stakeholders understand the findings from the evaluation (a) clinicians (b) participants and families (c) program funders (d) vocational program staff (e) program managers. (Score 1 if (d) and (e) are present; score 2 if (d) and (e) and one other is present; score 3 if all are present).</td>
<td>Interview</td>
<td>Is any knowledge gained from evaluation communicated to relevant stakeholders? Ask who was the evaluation report sent to and ask for examples of how and where the findings were communicated. Talk to stakeholders where possible, assess knowledge of evaluation findings, current program outcomes, strengths and weaknesses.</td>
<td>Centers for Disease Control &amp; Prevention, 2011; Del Valle et al., 2014; Finnerty et al., 2009a,b, Ivers et al., 2012; McHugo et al., 2009; Rapp et al., 2005.</td>
</tr>
<tr>
<td>Implementing the action plan</td>
<td>The findings from the evaluation have informed program delivery in specific ways.</td>
<td>The program is actively implementing the actions in the action plan. (Score 1 if 25% or less of the action plan is being implemented; score 2 if between more than 25% but 50% or less is being implemented; score 3 if more than 50% of the action plan is being implemented).</td>
<td>Interview Program documentation</td>
<td>Interview the program leader and discuss areas of the action plan, asking for examples of where action has been taken.</td>
<td>Centers for Disease Control &amp; Prevention, 2011; Gibbs et al., 2009.</td>
</tr>
<tr>
<td>Building staff evaluation capability</td>
<td>Program evaluation involves staff in order to develop and enhance their evaluation skills.</td>
<td>1. There is a high level of participation of program staff in evaluation activities. 2. Staff report improved evaluation skills. 3. Staff report using evaluation findings to improve program performance. (Score 1 for each anchor present).</td>
<td>Interview and observation</td>
<td>Can staff give examples of their involvement in evaluation activities and any changes in their knowledge and ability to conduct evaluation? E.g. establishing a new data collection system, training staff in evaluation, their ability to interpret data.</td>
<td>Gibbs et al., 2009; Australian Government, 2001</td>
</tr>
</tbody>
</table>

Notes. 1. **Green**=Category A: The evidence supports the inclusion of the attribute and its effect on program performance. **Red**=Category B: The evidence does not support the inclusion of the attribute and its effect on program performance. **Blue**=Category C: No empirical evidence was identified but authors have written about the potential for the attribute to influence program performance.
Removal of Non-Evidence-Based Practices

In the context of vocational rehabilitation, research has led to clear differentiation between evidence-based programs and non-evidence-based ones, both in terms of the underlying theory as well as the practices (Bond, 2004). Traditional, non-evidence-based programs in vocational rehabilitation have either not shared the same goals for participants, or take a ‘train then place’ approach. This involves assuming that almost all participants require a period of preparation prior to entering employment. Typical non-evidence-based practices can include programs that lack a focus on competitive employment in the open labour market, programs with high caseloads of between 30-150 participants with high assistance needs, the segregation of the vocational service from the health service, or the extensive use of work readiness assessments at program intake and prior to commencing a job (Crowther et al., 2001; Waghorn et al., 2009c). When an accumulation of these non-evidence-based practices exist, or when program goals diverge from individual competitive employment in the open labour market, vocational rehabilitation programs tend to take alternative forms. These may include sheltered workshops, business enterprises, work crews, group skills training, and other vocational preparation activities (Waghorn et al., 2009c).

The link with program performance

The identification, removal, and replacement of non-evidence-based practices is considered important to positively influence program implementation quality and improve the performance of vocational rehabilitation programs (Becker et al., 2001a; Drake et al., 1994; Lockett et al., 2018a). In the 1990s and early 2000s, conversion studies conducted in the USA evaluated the process and outcomes from transitioning psychiatric day programs and traditional vocational rehabilitation programs to supported employment programs. Drake et al. (1994) found that converting the practices in a psychiatric day treatment program increased the proportion of participants commencing competitive employment from 33.3% to 56.6%. Similarly, Becker et al. (2001b) found an increase in the proportion of participants employed, to 44.2% and 56.7%, in two converted programs compared to 19.5% in the program that remained a day treatment program. These studies provided the basis for identifying,
defining, and describing the agreed set of evidence-based practices in vocational rehabilitation, and identified the benefits of replacing non-evidence-based practices for improving program performance. However, even when programs have good overall adherence to evidence-based practices, it is possible that the identification and replacement of non-evidence-based practices and programs could further improve program performance. This is because the presence of non-evidence-based practices in the implementation of evidence-based vocational rehabilitation programs has been found to negatively impact on program performance (Bejerholm et al., 2015; Latimer et al., 2006; Waghorn et al., 2014). A Swedish RCT comparing an IPS program to traditional vocational rehabilitation illustrates this point. In Sweden, the welfare system stipulates that people on benefits must complete an internship prior to commencing competitive employment. The presence of this non-evidence-based practice significantly delayed time to commencing first competitive job in the IPS intervention group. The average days from program intake to commencing first employment were 462.5 (Bejerholm et al., 2015). Whilst this was still quicker than the average time from intake to commencing first employment in the control group (541.2 days), 462.5 days is high for an IPS program when compared with the average of 167.7 days in nine IPS programs in the RCTs in the Bond et al. (2012b) systematic review. This delay in commencing first job in the IPS group might also explain why there was no significant difference between the IPS group and the control group at six months, in terms of the number of participants commencing competitive employment (Bejerholm et al., 2015).

**Conceptualising the removal of non-evidence-based practices**

For the purposes of improving the effectiveness of vocational rehabilitation programs the removal of non-evidence-based practices can be conceptualised as the removal of practices or programs which do not align to the evidence base or that conflict with the primary program goals. Removal of non-evidence-based practices can be conceptualised as consisting of three measurable attributes. 1. The identification of non-evidence-based practices and conflicting programs. 2. The separation of any conflicting programs. 3. The conversion of non-evidence-based practices to evidence-based practices, or the removal of non-evidence-based practices.
Identification of non-evidence-based practices

Programs could begin by identifying practices which do not align with the current evidence base, and by identifying programs that might be conflicting with the primary goal of competitive employment. The IPS fidelity scales (IPS-15 and IPS-25) both contain measures to help programs identify non-evidence-based practices. For example, in the IPS-25 scale: the level of integration with mental health treatment teams (Item 5); the presence of exclusion criteria on who can participate in the program (Item 9); the way in which vocational assessments are used (Item 14); the average length of time from acceptance onto the program to commencing job search (Item 15); and the diversity of jobs found and employers used (Items 19 and 20) (Becker et al., 2015, p. 4). The scales have good discriminative validity, i.e. the scales can distinguish between programs practising supported employment and those that are not (Bond et al., 1997a, 2011).

However, even within a program largely focused on evidence-based practices, it is possible that there are non-evidence-based practices present, which if identified, removed, or replaced could improve program performance (Bejerholm et al., 2015; Latimer et al., 2006; Waghorn et al., 2014). Such practices could include non-time-limited group programs, paternalistic attitudes and engagement strategies, or a requirement to attend training or a set number of appointments prior to acceptance onto the vocational program (Loveland et al., 2007).

It is therefore likely to be important, along with conducting fidelity reviews, that program leaders conduct regular reviews of program practices. A review can identify any practices which do not align with the current evidence base and potential obstacles to changing these practices (Handler, Doel, Henry, & Lucca, 2003). Program leaders could also measure the average time from acceptance onto the program to commencing the first job. Traditional approaches take on average 34 weeks (236 days), compared to evidence-based approaches that on average support a person into employment within 24 weeks (167.7 days) (Bond et al., 2012b). These measures routinely collected, could assist program leaders to identify, monitor, and modify non-evidence-based practices which could delay job commencements.
Identification of non-evidence-based programs

It is also possible that seemingly complementary vocational programs, such as those supporting people into education and voluntary work, or programs which support people to retain employment, can adversely affect program performance. This could happen in circumstances where the programs are delivered by the same staff, managed under the same organisation, or provided in the same premises. This is possible because they can compete with the primary goal of supporting people into competitive employment. These complementary programs could also be competing for referrals, program management resources, and for funding.

This potential for seemingly complementary programs to be competing with the primary program is illustrated in a study examining the effectiveness of a vocational program. Staff were required to provide employment assistance to people referred to the program because they were at risk of losing their jobs, while at the same time as delivering employment assistance to unemployed participants (Pittam, Boyce, Secker, Lockett, & Samele, 2010). The study found that although the program was successful with the former group, the outcomes for the unemployed group were negatively impacted. Staff prioritised working with people whose jobs were at risk over those who did not yet have a job.

Similarly, programs can have funding contracts which provide payments for educational and voluntary outcomes. This has been found to reduce the effectiveness of the vocational program by diverting the attention of staff away from the primary goal of competitive employment (van Veggel, Waghorn, & Dias, 2015). The identification, separation, and replacement of these practices and programs is therefore likely to be an important mechanism to improve program performance.

The IPS fidelity scales do not attempt to assess potential conflicting programs which may be impacting on the effectiveness of the vocational rehabilitation program. To identify possible conflicting programs, program leaders could review the programs that share the same premises, those that come under the same manager, funder, or parent organisation, and those that use the same referral pathway. This would not have to be a lengthy review process. The review could just be a regular process in
which consideration is given to whether local arrangements around sharing premises, management, or referral processes are having a negative or a positive influence on the vocational rehabilitation’s primary objective of supporting participants into competitive employment. Also included in this review, could be a check on contractual requirements to determine if there are performance measures for outcomes which compete with the primary goal of competitive employment (Lockett et al., 2018b).

**Separation**

When there are programs which have been found to conflict with the primary program goal, a plan is needed to support their separation (Necker et al., 2001; Handler et al., 2003). The plan could involve one or more of the following activities. 1. Changing premises, and staff roles and responsibilities (Drake et al., 1994). 2. Consulting with participants, families, and other stakeholders (Becker et al., 2001a). 3. Talking to funders to change contracts or address shared referral pathways (Lockett et al., 2018b). 4. An organisational re-structure to remove the conflict of shared managers (Loveland et al., 2007), or multiple management responsibilities (van Erp et al., 2007).

The separation of staff roles and responsibilities may have improved outcomes in the implementation of four sites implementing IPS programs in the Netherlands. Program leaders reported minimal time to manage the IPS program. One leader reported having just 12 hours per week formally allocated to the IPS program, the other three had no formal time allocated. All leaders were managing other programs and attending to a range of organisational duties, such as arranging finance for the program (van Erp et al., 2007).

**Conversion**

It has been argued that the removal and replacement of non-evidence-based practices is a multi-faceted issue and therefore needs an approach to change that spans the whole organisation (Kitson et al., 1998; McKenna, Ashton, & Keeney, 2004). Identifying specific barriers within the working environment in which they occur is considered important to helping the conversion of practices (McKenna et al., 2004). To move away from routines and rituals there also needs to be a stronger focus across the organisation, or across the community, on encouraging the use of research in practice.
(Becker et al., 2001a; Newman et al., 1998). The removal and replacement of non-evidence-based practices will also take time and is likely to go through a process of gradual transition (Handler et al., 2003).

In a two-year study examining strategies which improved the implementation fidelity of evidence-based practices in vocational rehabilitation, leaders in high-fidelity sites had actively discontinued services, revised paperwork requirements, and aligned organisational policies which conflicted with the evidence-based practice that was being implemented (Bond et al., 2009). In one agency, the funder removed the financial incentives associated with supporting program participants into non-competitive jobs (Bond, McHugo, Becker, Rapp, & Whitley et al., 2008b). Other important influences in this transformation process were director-level support, sites volunteering to participate in the change, and on-site technical assistance (Bond et al., 2008b).

Novak, Rogan, Mank, and DiLeo (2003) conducted a national survey of state vocational rehabilitation agencies across 50 states in the USA to identify the impact of systems change on the conversion of segregated traditional vocational rehabilitation programs to supported employment programs. Training to vocational rehabilitation and to mental health staff was prioritised by 46 of the states and the training was rated as having a high impact by 43 states (Novak et al., 2003). This was an area of top priority alongside the provision of technical assistance. State agencies who responded also highlighted the importance of policy directives and financial incentives which target and reward the downsizing, or closure of traditional vocational rehabilitation programs (Novak et al., 2003).

The replacement of any non-vocational support services previously being provided by the ineffective practices is important alongside the removal of these practices (Becker et al., 2001a). For example, in the conversion of a day treatment program to a supported employment program, Drake et al. (1994) reported that a replacement drop-in centre was established at the same time for participants who wished to maintain a service with a focus on socialisation with peers.

A study which examined the impact on vocational rehabilitation practices of a change in funding from fee-to-service to outcomes-based payment, found that practices shifted away from pre-
employment supports to job development and job placement. However, there were limitations related to the quality of jobs that were secured and the target population served by the program (Gewurtz, Cott, Rush, & Kirsh, 2015). This example highlights the role of financial incentives in modifying the practices of vocational rehabilitation programs. Careful consideration needs to be given to ensure financial rewards are incentivising the adoption of evidence-based practices and there are no unintended consequences which move practices away from the evidence base (Gewurtz et al., 2015).

Attitudes of staff towards research and knowledge of evidence-based practices are also important influences on whether staff implement the latest evidence-based practices (Norcross, Hogan & Koocher, 2008, p. 211). In drug education, researchers found that it was important to provide teachers with training to help distinguish evidence-based practices from other practices and approaches that seemed very similar (Hansen & McNeal, 1999). One hundred and forty-six classes were observed to understand the performance of teachers delivering a drug education program. The authors concluded that training staff in the theory and concepts underlying evidence-based practices should be conducted, as well as training which helps teachers to unlearn as much as to learn (Hansen & McNeal, 1999).

Similarly, Fairhurst and Huby (1998) in a study of the attitudes of general practitioners towards research found that the doctors did not feel the findings of RCTs translated to the patients in front of them where people presented with complex social and psychological contexts. In this study the general practitioners preferred to use the opinions of other doctors, i.e. opinions of others, over the research findings. Other barriers to conversion include a lack of authority to make change, lack of access to research, and limited time to implement new ideas (Handler et al., 2003; McKenna et al., 2004).

**Construct operationalisation and measurement**

Table 21 outlines the three candidates for measurable attributes of the removal of non-evidence-based practices. These attributes are considered sufficient to adequately represent the construct. For each measurable attribute one or more targeted statements are suggested, along with a possible item pool that program leaders could use to identify, measure and assess this dimension of implementation quality in the IFVR.
Table 21. Operationalising the construct of removal of non-evidence-based practices

<table>
<thead>
<tr>
<th>Measurable attributes</th>
<th>Targeted statements</th>
<th>Measurement Item pool</th>
<th>Method of measurement</th>
<th>Examples of questions and information to gather</th>
<th>Evidence for attribute¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identification</td>
<td>Practices in the vocational service: Within the last year, a specific review of program practices has been conducted with the purpose of identifying any practices that do not align with the current evidence base. (Yes=1, No=0). The average time from referral to commencing first job is six months or less. (Yes=1, No or unknown=0).</td>
<td>Interview Review program documentation and client management system.</td>
<td>Ask staff directly about their knowledge of EBPs and their attitudes towards research. Ask program staff how they would know if they deliver non-EBPs in the program.</td>
<td>Bejerholm et al., 2015; Bond 2004; Handler et al., 2003; Latimer et al., 2006; Loveland et al., 2007.</td>
<td></td>
</tr>
<tr>
<td>Conflicting programs</td>
<td>Within the last year a review has been conducted to identify any programs that are conflicting with the primary goal of competitive employment. (Yes=1, No=0).</td>
<td>Interview Review program documentation</td>
<td>The review should include whether there are programs nearby that use the same referral pathway, operate from the same premises, have the same manager, or the same parent organisation.</td>
<td>Loveland et al., 2007; Novak et al., 2003; Pittam et al., 2010.</td>
<td></td>
</tr>
<tr>
<td>Contract conflicts</td>
<td>Within the last year a review of the program contract(s) has been conducted to identify any performance outcomes that may conflict with the primary goal of competitive employment. (Yes=1, No=0).</td>
<td>Interview Review program documentation</td>
<td>Ask program managers if there are contractual terms which reward other outcomes such as education, voluntary work, pre-vocational training, supporting people at risk of job loss.</td>
<td>Gewurtz et al., 2015; Lockett et al., 2018b; van Veggel et al., 2015.</td>
<td></td>
</tr>
</tbody>
</table>

Construct specification: The replacement of practices or programs which do not align to the evidence base or that conflict with the primary program goals.
<table>
<thead>
<tr>
<th>Measurable attributes</th>
<th>Targeted statements</th>
<th>Measurement Item pool</th>
<th>Method of measurement</th>
<th>Examples of questions and information to gather</th>
<th>Evidence for attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td>Separation</td>
<td>Plan to separate conflicting programs</td>
<td>There is an action plan in operation to separate the conflicting program(s). (Yes or there are not any conflicting programs=1, No action plan=0). The action plan has been signed off by program directors. (Yes or there are not any conflicting programs=1, No action plan=0).</td>
<td>Review program documentation</td>
<td>The action plan might include a meeting with the funder to review contract terms and conditions, a planned organisational re-structure, or office re-location.</td>
<td>Becker et al., 2001a; Drake et al., 1994; Handler et al., 2003; Loveland et al., 2007.</td>
</tr>
<tr>
<td>Conversion</td>
<td>Replacing practices</td>
<td>There is an action plan in operation to remove practices that are known to be ineffective. (Yes or there are not any conflicting practices=1, No action plan=0). The plan includes alternative ways to meet any non-vocational needs of program participants that were previously being met by the non-effective practices. (Yes or there are not any conflicting practices=1, No action plan=0). The action plan has been signed off by program directors. (Yes or there are not any conflicting practices=1, No action plan=0).</td>
<td>Review program documentation</td>
<td>The action plan might include training staff in the value of evidence-based practices, a commitment to measuring time to first job, talking to funders to resolve contractual conditions incentivising the delivery of non-EBPs, re-aligning staff roles and responsibilities.</td>
<td>Becker et al., 2001a; Bond et al., 2008a, 2009; Drake et al., 1994; Handler et al., 2003; McKenna et al., 2004; Ncrcross et al., 2008.</td>
</tr>
</tbody>
</table>

Notes. EBPs=evidence-based practices. 1. Green=Category A: The evidence supports the inclusion of the attribute and its effect on program performance. Red=Category B: The evidence does not support the inclusion of the attribute and its effect on program performance. Blue=Category C: No empirical evidence was identified, but authors have written about the potential for the attribute to influence program performance.
Complementary Programs

The success of some types of rehabilitation programs can depend on the availability, quality, and effectiveness of complementary programs. A program can be considered complementary when it can potentially positively influence intended rehabilitation outcomes, yet represents interventions not normally provided as part of the rehabilitation program i.e. non-vocational programs such as mental health, general health, housing, legal assistance, or financial support programs. For instance, in medical rehabilitation, programs do not include medical treatment and care, even though rehabilitation outcomes may depend on the prior provision of optimal medical treatment and health care to reduce impairments due to untreated health conditions. In some types of vocational rehabilitation the close involvement of a complementary program may be considered necessary for success. In such cases the complementary program may be specified as an essential practice principle within the scope of the rehabilitation program. This is the case in evidence-based vocational rehabilitation designed for people with severe mental illnesses (Becker et al., 2015, p. 2). Here the complementary program is continuing mental health treatment and care. This is considered essential to the success of this form of vocational rehabilitation (Drake et al., 2003). The integration of medical care with the vocational rehabilitation program is therefore measured as part of measuring implementation fidelity (Becker et al., 2015, p. 49). However, the availability, quality, and effectiveness of the medical treatment and care program is not measured as part of IPS implementation fidelity (Becker et al., 2015, p. 4; Lockett et al., 2018a).

Complementary programs and program performance

The importance of complementary programs on program performance has been previously recognised (Durlak & DuPre, 2008; Fixsen et al., 2005). To evaluate a primary rehabilitation program it may be necessary to monitor, measure, or control for the effectiveness of any complementary programs. This is because differences in the amount, type, and quality of complementary interventions associated with the primary program intervention and control groups in controlled trials, may account for differences in primary program outcomes (Durlak & DuPre, 2008). For example, in a USA-based IPS RCT participants in the intervention group as well as receiving the IPS program, had more mental
health case management and mental health medication management services, and more contacts with mental health services, than those in the control group (140 contacts versus 83 contacts). Significantly more participants in the intervention group commenced competitive employment than participants in the control group (63.6% participants versus 26.0%; $p<.001$) (Gold et al., 2006). The effects of any differential access to complementary programs need to be taken account of in the analysis and therefore should be measured and reported in RCTs (Durlak, 2015).

In vocational rehabilitation there are several candidate complementary programs. Mental health interventions, mental health support services, substance use and addiction services, general health services, family support and education services, peer support, housing, income support and welfare benefits, financial counselling, other forms of disability support, and legal assistance (Drake et al., 2003; Lockett et al., 2018a; Nepe, Pini, & Waghorn, 2011; Whiteford et al., 2014). The key questions about the nature of these complementary programs concern: (a) to what extent the complementary program is facilitating the goals of the primary program; (b) to what extent the complementary program is available, and meeting the needs of individuals; (c) how well the complementary program aligns with any recognised standards for that program; and (d) how well is the complementary program coordinated with the primary program? (Durlak & DuPre, 2008; Dane & Schneider, 1998). Answering these questions requires formal measurement otherwise the influence of complementary programs may remain unknown confounders in trials of the primary program outcomes. Measurement is also important so that vocational rehabilitation programs can maximise the support provided by complementary programs to improve program performance.

**Conceptualising complementary programs**

For the purposes of improving the performance of vocational rehabilitation programs and for improving implementation quality, complementary programs can be conceptualised as the availability, quality, and effectiveness of complementary programs with respect to active participants in the vocational rehabilitation program. Complementary programs are considered complementary if they can potentially influence rehabilitation outcomes. Positively when present, and negatively when absent.
Complementary programs can be broadly conceptualised and operationalised as consisting of six measurable attributes.

1. Detection and identification of complementary programs.
2. Relevance of the complementary program to the goals of the primary program.
3. Availability of the complementary program to participants of the primary program.
4. Quality of the complementary program with respect to standards set for that type of complementary program.
5. Effectiveness of the complementary program with respect to the same target population as the primary program.
6. The extent the complementary program can be coordinated with the primary program.

The quality dimension concerned with complementary programs differs from the dimension concerning the removal of non-evidence-based practices. This is because complementary programs can either positively or negatively affect the primary program. The influence of complementary programs can be positive if the complementary program is available to target participants, coordinated with the primary program, and is of sufficient quality to be effective. By implication the absence of a potentially effective program can be negative, for instance if a complementary program has been identified as potentially able to enhance the goals of the primary program, but is not available to program participants, or is unable to be coordinated, or if it is not of sufficient quality in its current form.

**The detection and identification of complementary programs**

Vulnerable and disadvantaged people in the community can have multiple, and often complex unmet needs which may require services from a range of health and social services (Areberg et al., 2013; Morgan et al., 2016; New Zealand Productivity Commission, 2015; Whiteford et al., 2014). These range of services span several service sectors, including housing, welfare, social care, financial, and legal issues. These services also span the government and non-government sectors and involve clinical and non-clinical services (Whiteford et al., 2014). Most of these providers specialise in a specific service delivery area and so will be restricted in the type and range of services they can
provide. To support multiple unmet needs simultaneously and improve individual outcomes, leaders of the primary program often need to identify the most likely candidates for complementary programs within the service delivery context. Once potential candidates for complementary programs are identified, the attributes of the most relevant ones can be measured.

Vocational rehabilitation programs for people with severe mental illnesses usually begin by identifying the barriers to employment hindering each participant in each phase of (a) job searching, (b) starting employment, and (c) retaining employment. Program leaders can also identify the other services available that complement the support provided by the vocational rehabilitation program.

People with severe mental illnesses seeking vocational rehabilitation assistance are also likely to be educationally disadvantaged, live in unstable accommodation, have restricted social networks, have co-morbid drug or alcohol problems, and have poor physical health (De Hert et al., 2011; Hancock, Smith-Merry, Gillespie, & Yen, 2017; Morgan et al., 2016). Analysis of the 2010 Australian national survey of psychotic disorders, the Survey of High Impact Psychosis which had a sample of 1,825, illustrates the multiple unmet needs of participants. Six top-ranked challenges were identified by people diagnosed with psychotic disorders.

1. Financial matters.
2. Loneliness, and social isolation.
3. Lack of employment.
4. Poor physical health.
5. Symptoms of mental illness.
6. Lack of stable suitable housing (Morgan et al., 2016).

Complementary programs that are therefore potentially likely to facilitate the success of the vocational rehabilitation program include those that simultaneously offer assistance with housing and tenancy arrangements, financial planning and debt management, managing mental health and general health, drug and alcohol problems, childcare, or other family commitments (Poremski, Whitley, & Latimer, 2014).
To best utilise complementary programs available in the local service delivery context, vocational rehabilitation program leaders can assess the current non-vocational needs of program participants, alongside the assessment of their vocational support needs. Some programs have found ways to explore broader needs. For example, in the Australian Partners in Recovery program a broad range of participants' needs are assessed along with the range of existing services available that can help meet those needs. The program employs facilitators in non-clinical roles to develop and maintain partnerships with complementary services in the local area. These facilitators meet participants face-to-face to understand their needs and follow up regularly to understand whether these needs had been met or have changed (Smith-Merry, Gillespie, Hancock, & Yen, 2015).

The logical starting point for the measurement of this construct is that program leaders regularly assess the potential candidates for complementary programs. This assessment could be based on the identified needs of the target program participants. For evidence-based vocational rehabilitation programs one of the complementary programs considered essential is mental health treatment and care (Drake et al., 2003; Swanson et al., 2014). However, a wider assessment of the non-vocational needs of program participants will likely identify other community services that can benefit participants in vocational rehabilitation and their families. These should therefore be detection and local candidate complementary programs identified from the outset (Nepe et al., 2011).

**Relevance of the complementary program**

Once primary program leaders have identified the types of complementary programs that can also support participants needs, the next step is to assess the relevance of the complementary program to the goal of the primary program. Primary program leaders can look for evidence that the complementary program influences the goal attainment of the primary program. They could ask the question: Is there any evidence that the complementary program also influences employment outcomes? Taking the example of the complementary program substance use treatment there is evidence that completion of a substance use program can enhance employment outcomes (Burnett-Zeigler et al., 2012; Platt, 1995). In contrast, the evidence for the relevance of housing stability as a
predictor of employment outcomes is mixed (Nelson, Aubry, & Lafrance, 2007). Program leaders should therefore carry out a local assessment of their complementary program’s potential importance to program participants employment outcomes.

**Availability of the complementary program**

Primary program leaders could ascertain the extent that a complementary program is available. Availability should be on an unrestricted basis for all primary program participants throughout their time with the primary vocational rehabilitation program. This includes continuing access while participants are in all phases of vocational rehabilitation, from job searching through to continuing employment. This is important because some vocational rehabilitation programs have found that mental health treatment teams may discharge people from mental health treatment and care following referral to a vocational rehabilitation service (Killackey & Waghorn, 2008). Continuing mental health treatment in parallel with and throughout the vocational rehabilitation program can be important. Continuing mental health treatment can assist participants with issues such as anxiety management when starting a new job, dealing with problems at work that can challenge mental health, job loss, relapse prevention planning, and social skills training (Areberg et al., 2013; Sherring, Robson, Morris, Frost, & Tirupati, 2010).

Primary program leaders can ascertain the availability of a complementary program by monitoring access to and attendance by participants of the primary program. Whilst services might be available they may not be accessed readily by program participants. This may be due to costs, to transport barriers, or for fear of discrimination. An example is access to general practice health care, which whilst available is often under-used by people diagnosed with severe mental illnesses (Royal Australian and New Zealand College of Psychiatrists [RANZCP], 2015). To measure the availability of important complementary programs primary program leaders could ask participants how often they need the complementary program and ask about the last time they accessed the complementary program. With this knowledge, primary program leaders can take action to improve access to relevant complementary programs.
The quality of complementary program

The complementary programs should align with any clinical guidelines or quality standards for that program. In medicine it is an almost universal requirement now to use only evidence-based treatments to avoid doing harm. In developed countries medical procedures, devices, and pharmaceutical products are strictly controlled by governments. There are standards for efficacy when new products are developed and adverse events and side-effects are monitored throughout their sanctioned use. The medico-legal and insurance system also helps to deliver quality control. For example, in Australia the Australian Commission on Safety and Quality in Health Care routinely publish clinical case standards which set out the clinical pathway, based on current best evidence, for specific conditions and procedures. For example, stroke, acute coronary syndromes, hip fracture care (Australian Commission on Safety and Quality in Health Care, 2018). Similarly, in the UK the National Institute for Health and Care Excellence (NICE) sets guidance, clinical pathways, and quality standards for a range of conditions and diseases, including for example quality standards for the assessment, treatment, and management of drug and alcohol use disorders (NICE, 2012, 2014).

However, psychosocial interventions are not scrutinised as intensively either during development or many years later and may not be removed even if ineffective. Whilst standards do exist for housing programs, for example the Pathways Housing First model (Stefancic, Tsemberis, Messeri, Drake, & Goering, 2013), they invariably have not undergone the same intensity of research investigation and efficacy testing as those for medical treatments and procedures. Vocational rehabilitation program leaders can identify the relevant standards appropriate to the complementary program and review the extent the local complementary program follows any specified standards. Some complementary programs may follow service standards which explicitly encourage coordination with vocational rehabilitation. For example, the NICE guidance on drug use disorders (NICE, 2012) and the Pathways Housing First fidelity scale (Stefancic et al., 2013) specifically mention coordination with vocational services.
Complementary programs may have recent clinical audit reports that are publicly available, or a fidelity review or a service audit report they are willing to share with leaders of vocational rehabilitation programs. If these are not available primary program leaders could ask to meet with the staff from the complementary program to spend time in the program, if appropriate. During these meetings program leaders could ask the program staff directly about their alignment with service standards as well as their quality assurance processes. Alternatively, staff in the primary program could ask program participants about their experiences and views on the complementary programs.

**Effectiveness with respect to the shared target population**

It is also important to assess whether the complementary program is effective with respect to the same target population as the primary program (Galletly et al., 2016). For example, leaders of a vocational rehabilitation program working with people with severe mental illnesses could look for evidence of the effectiveness of the complementary program with respect to the shared target group, namely people with severe mental illnesses. Leaders of vocational rehabilitation programs can also examine the published literature applicable to the complementary program to ascertain effectiveness with respect to the target population. An example of a complementary program targeting the same population is the Pathways Housing First model. Housing First was designed to support homeless people with severe mental illnesses into stable accommodation (Stefancic et al., 2013). On the other hand, general health care services may not have sufficient experience at providing physical health care to people with severe mental illnesses (RANZCP, 2015). If this is the case, primary program leaders may need to actively seek out general practices with mental health specialists, or general practitioners with specialist expertise in mental health, to encourage greater use of these complementary programs.

**The importance of coordination**

Coordination of support between the primary program and any identified complementary programs is important for supporting participants with multiple needs (Hancock et al., 2017; Hermann et al., 2002). In a review of the literature examining the effectiveness of intersectoral collaboration, Whiteford et al. (2014) identified eight different strategies services use to achieve better coordination.
1. Shared care plans or a common medical record.
2. Shared information systems across providers.
3. Formal interagency collaborative arrangements, such as a memorandum of understanding.
4. Staff from the different providers attending the same training.
5. Interagency coordinating and planning committees.
6. The co-location of services.
7. Blended funding or management by a single lead agency.
8. Joint service provision (Whiteford et al., 2014).

The authors (Whiteford et al., 2014) found that organisations differed in the number of strategies they implemented. Although the quality of studies reviewed limited the conclusions that could be drawn, the authors found that coordination activities improved communication between services, increased access to the resources available to programs and to program participants, and strengthened empathy for other service providers. Enablers of better coordination included strong senior leadership, co-location of services, monitoring of joint service effectiveness, and optimising the intensity and frequency of contacts between services.

The development of formal communication structures between services has improved service coordination, particularly where co-location or direct employment of employment specialists in public mental health services has not been possible. Formal communication structures can include scheduled weekly meetings between services, exchange of assessment information, monthly case reviews, and regular joint appointments (Sherring et al., 2010).

Another strategy for improving service co-ordination has been the development of facilitated referral processes between programs and the tracking of referrals between the co-ordinated services (Cooley, McAllister, Sherrieb, & Clark, 2003; Stange et al., 2010). Facilitated referrals processes, also referred to as warm handoffs, are used in health settings to improve transitions between services (Richter et al., 2016). Warm handoffs involve face-to-face or phone transfers directly between providers. These handoffs have been found to improve treatment engagement (Richter et al., 2016).
The standards for some complementary programs highlight the importance of coordination with vocational rehabilitation. For example, the NICE guidelines on the management of drug use includes a standard stating that people in drug treatment should be offered support to access services that promote recovery and reintegration, including employment (NICE, 2012). In contrast, the NICE standards for alcohol treatment and management do not explicitly reference employment (NICE, 2014). The recently updated guide for psychiatrists in Australia and New Zealand on the treatment and management of psychosis and schizophrenia highlights the importance of access to evidence-based vocational rehabilitation. Although this guide does not outline the importance of coordination with employment services (Galletly et al., 2016).

To improve the performance of the vocational rehabilitation program, program leaders can regularly assess the level of coordination with the complementary program and address any lack of coordination issues that arise. Program leaders can ask staff in the primary program to give examples of how often, and in what way, they communicate with staff in the complementary programs. Program leaders can instigate formal co-ordination arrangements with complementary programs and establish joint committees. Program staff may also have ideas and suggestions for how to improve local collaborative arrangements.

**Construct operationalisation and measurement**

Table 22 outlines the six candidates for measurable attributes of complementary programs. These six attributes are considered sufficient to adequately represent the construct. For each attribute one or more targeted statements are suggested, along with a possible measurement item pool and sample questions that program leaders could use to identify, measure, and assess the quality of complementary programs. Once each complementary program is identified program leaders could assess each of the other five measurable attributes of the complementary program, and then if needed, calculate an average score across all complementary programs (see Table 22).
<table>
<thead>
<tr>
<th>Table 22. Operationalising the construct of complementary programs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Measurable attributes</strong></td>
</tr>
<tr>
<td>Detection and identification</td>
</tr>
<tr>
<td>Relevance</td>
</tr>
<tr>
<td>Availability</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Measurable attributes</td>
</tr>
<tr>
<td>------------------------</td>
</tr>
</tbody>
</table>
| Quality                | The CP aligns with recognised standards for that program | 1. The program can identify clinical guidelines or quality standards for the CPs.  
2. The CPs are delivered according to recognised EBPs or clinical practice guidelines.  
3. The program has knowledge of how program participants view the CP. (Score * for each anchor that is present). | Interview | Ask program staff if they are aware of the clinical guidelines or quality standards that are available for the CP. Ask if staff are aware of whether the CP adheres to those clinical guidelines or service standards? Ask program staff if they are aware of program participants views on the CP. | NICE, 2012; Stefancic et al., 2013. |
| Effectiveness          | The CP is effective with the same target population | 1. The program has looked for evidence of the CPs effectiveness with the same target population as the VR program.  
2. The primary program actively seeks out CPs which have been effective with the same target participants. (Score * for each anchor that is present). | Interview | Ask program staff to give examples of how CPs have supported the same target population. Ask staff if they know the primary populations served by the CP. | Galletly et al., 2016; Stefancic et al., 2013. |
| Coordination           | The CP are coordinated with the primary program | The primary program has the following arrangements in place with the CP  
(a) co-located services  
(b) formal collaboration arrangements  
(c) a coordinating committee.  
(d) shared care plan or individual record  
(e) shared Information Systems  
(f) shared staff training  
(g) facilitated referrals  
(h) scheduled weekly meetings.  
(i) joint monitoring of program outcomes. (Score 3 if three or more are present, 2 if two are present, 1 if one is present). | Interview | Ask staff to give examples of ways in which they formally and informally communicate with the CP. For example, what type of referrals processes are there between the primary program and the CP? Ask staff to give an example of the last time they communicated with staff in the CP and the next time they think they will talk to them. Ask staff to talk about the staff in the CP to see how well they know them. Review any documentation which shows joint working arrangements e.g. MOU, minutes of committee meetings. | Galletly et al., 2016; Hancock et al., 2017; Hermann et al., 2003; NICE, 2012; Richter et al., 2016; Sherrington et al., 2010; Stefancic et al., 2013; Whiteford et al., 2014. |

Notes. CP=Complementary Programs. EBPs=Evidence-Based Practices. MOU=Memorandum of Understanding. 1. Green=Category A: The evidence supports the inclusion of the attribute and its effect on program performance. Red=Category B: The evidence does not support the inclusion of the attribute and its effect on program performance. Blue=Category C: No empirical evidence was identified, but authors have written about the potential for the attribute to influence program performance. 2. CPs are considered complementary if they potentially influence rehabilitation outcomes, e.g. housing services, tenancy support, legal assistance, medication management, psychological therapies, peer support, family support, general healthcare. 3. Once the CPs are identified, program leaders should measure the attributes of each complementary program, and then calculate an average score across all CPs.
Program Intensity

In vocational rehabilitation a program is a planned and coordinated set of activities tailored to an individual with the specific goal of restoring an individual's vocational functioning and their economic participation in the community (Waddell et al., 2008, p. 11). The concentration of these planned activities and inputs relevant to each participant’s vocational goal is what is meant by program intensity. Intensity can be estimated by summing the frequency and extent of relevant activities of support given to a cohort of participants. This can then be expressed as an average value per program participant. The units of measurement could involve the time taken by those activities, or the number of contacts of a standardised type, or capture the extent of program activity per person and per time unit (McGuire et al., 2011; Rogers, MacDonald-Wilson, Danley, Martin, & Anthony, 1997). For instance, where a vocational program has one full-time vocational rehabilitation practitioner with an active caseload of 50 participants. If each participant were seen face to face or spoken to by telephone once a week for 15 minutes, this program could be considered half as intensive as a similar program with the 25 active caseload participants contacted once a week for 15 minutes, provided other conditions remained constant. If the weekly contact per participant went up to twice a week, this would also increase program intensity.

Program intensity and program performance

Program intensity is a potentially important dimension of program implementation quality because it has been linked to program performance. For example, low program intensity can explain low program performance even when adherence to a standardised program (fidelity) is good (Lockett et al., 2018a). In a UK RCT of IPS, the IPS intervention programs achieved more than 80% fidelity (scores of 67/75 and 69/75 on the IPS-15 fidelity scale) yet only 22.1% of participants obtained competitive employment (Heslin et al., 2011). A mean of 15 contacts were recorded between program participants and a single employment specialist over a two-year period (Heslin et al., 2011). This is less than a third of the intensity (mean 48 contacts over a two-year study period) reported by McGuire et al. (2011) in the USA. This USA intervention at higher intensity and similar high fidelity (72/75),
achieved 75.0% of participants commencing competitive employment (Bond et al., 2007). There is other evidence that program intensity is a factor in program performance. Taylor and Bond (2014) found the frequency of contacts between the employment specialist and participants was positively associated with program performance. Program performance was measured by the monthly proportion of participants employed, and by the number of participants attaining 90-day job tenure.

**Conceptualising program intensity**

For improving the performance of vocational rehabilitation programs a measure of program intensity aims to discover whether or not each participant received sufficient assistance to achieve the goals of the program. In this conceptualisation therefore, program intensity is considered a function of both the number of participants and the number of relevant program activities provided in a defined period. Employment assistance is usually provided for a participant by a single employment specialist or vocational professional but can involve multiple vocational professionals. In the latter case, program intensity can be estimated from the sum of the extent and frequency of combined activities provided by all vocational professionals involved, expressed as a ratio per participant within a defined period.

The intensity of assistance provided must be relevant to the participant's current need for assistance otherwise it could drive program disengagement and separation from the program. Programs with low intensity are likely to be ineffective for those with higher intensity assistance needs. For these participants, less help than sought or needed is likely to encourage disengagement and inactivity. Whereas more optimal levels of program intensity, matching the assistance sought by participants, are likely to support and encourage active participation in the employment assistance being offered (Fleming, Del Valle, Kim, & Leahy, 2012).

Program intensity must therefore be flexible and free to vary with demand in a timely way so that individual participants can access more help when needed without waiting so long that employment opportunities are missed, or motivation is lost. This is important because assistance needs can also be a function of the stage of the employment journey. For example, intensity of assistance needs is likely to differ when participants are in job search compared to when they are employed (Bond
& Kukla, 2009). These aspects of program intensity imply attributes of availability at short notice, and that the intensity can be adjusted to demand based on the needs of an individual participant (Waddell et al., 2008, p. 49). In addition, because intensity is counted via the sum of all program activity (frequency and amount) in a defined period, both direct contacts with the person and indirect activities (undertaken on the participant's behalf) need to be taken into account.

This implies that the proportion of persons actively participating in a program at any time may also reflect the program's intensity. Active participation has been measured previously and was defined as at least one contact with the vocational professional within the last month (Becker et al., 2015, p. 38). Program leaders may observe changes in active participation if program intensity is not matched to participant’s assistance needs.

For the purpose of improving the performance of vocational rehabilitation programs, program intensity can be described as the provision to each participant of the optimum amount of relevant employment assistance while they remain active in the program. It is hypothesised to consist of four measurable attributes.

1. Tailored, that the intensity of assistance is tailored in relation to the participant’s needs and circumstances.
2. Available, that the participant can access the program for as long as they need employment assistance to get a job.
3. Direct contact, the frequency and amount of direct contact.
4. Non-direct contact, the frequency and amount of non-direct contact.

Although intensity is measured with respect to a defined period, it is not the same as duration of participation in a program. Duration is best considered an exposure variable which complements intensity but is not a dimension of intensity. The amount of exposure to a program could be determined as intensity multiplied by duration of each exposure, expressed as the sum of all exposures.

Previous studies on vocational rehabilitation programs have rarely attempted to define and measure program intensity. Where studies have done so the conceptualisation appears incomplete. The
measures reported were often restricted to duration of participation in the program, the total number of face-to-face contacts between participants and vocational staff, or to the frequency of contacts (Bond & Kukla, 2011; MacDonald-Wilson, Revell, Nguyen, & Peterson, 1991; McGuire et al., 2011; Roger et al., 1997; Taylor & Bond, 2014).

The IPS fidelity scales (IPS-15 and IPS-25) include two items which acknowledge the potential importance of program intensity. Item 1 specifies a maximum caseload size of 20 or less active participants for the maximum item-level fidelity score of 5 (Becker et al., 2015, p. 37). Item 3 assesses the extent that the employment specialist being assessed provides only employment support services (p. 43). This latter item is related to fidelity because providing other services such as mental health case management would necessarily dilute vocational program intensity. Other than these two items in the frequently used fidelity scales, and measures of contact frequency, few studies have acknowledged the potential importance of measuring program intensity directly and independently of program fidelity.

**Tailoring employment assistance**

Tailoring the intensity of employment assistance according to the needs and circumstances of participants is an important aspect of effective vocational rehabilitation programs for people with a range of health conditions as well as people with disabilities (Waddell et al., p. 49; Wehman et al., 2003). Low-intensity supports are likely to be appropriate and effective for participants with a recent independent employment history who may have recently lost a job, or who only need minor forms of support to stay at work. More intensive and structured support is more likely to be needed by participants with more substantial impairments who request additional help to secure and sustain employment (Waddell et al., 2008, p. 46). Meeting each person’s unique needs when help is requested, and matching the pace and intensity of support, has been claimed as critical for success (Grove, 1999).

Tailoring support at the right level and pace emerged as an important requirement for participants in a UK study. The authors examined the experiences of 182 people with severe mental illnesses participating in vocational rehabilitation programs (Johnson et al., 2009). Offering employment assistance at a pace the participant was comfortable with, to prevent the participants
feeling rushed or pushed, was rated as important. Participants also noted the importance of active and intensive support, criticising support when it was not intensive enough. Participants valued employment assistance that allowed them to maintain control over the process, whilst at the same time was proactive and facilitative enough not to leave participants dealing with things on their own (Johnson et al., 2009).

There is also a need to for program intensity to vary with the needs of the participant’s vocational situation (Gowdy et al., 2004). For example, the intensity of support may need to increase when participants start a new job, when health deteriorates while in employment, or when there are major changes in the workplace (Wehman et al., 2003). Studies of the frequency of contacts found that contact frequency was higher initially, then reduced rapidly, and was followed by periodic spikes in contact thereafter (Bond & Kukla, 2009, 2011; MacDonald-Wilson et al., 1991). Whilst the frequency of contacts in these studies appeared to be responding to participant needs, it was not clear how contact was actually tailored to the needs of each individual. This analysis suggests a potentially productive opportunity for program leaders. The opportunity is to establish a process for identifying and determining at regular intervals the intensity of support needed by each participant. This could include establishing agreement about the timing and duration of contacts, meetings, and other support activities. The availability of more or less intense support can be discussed so that participants are aware that this can change alongside their changing assistance needs and circumstances. Programs offering tailored support could be more explicit about how program intensity can vary for different participants, as well as at different stages of a participant’s vocational rehabilitation journey, and by their changing personal circumstances.

The availability of employment assistance

A widely recognised attribute of effective community mental health services is that they remain available for as long as needed by participants. This is because of the long-term and fluctuating nature of severe mental illnesses, and the exacerbation of symptoms and loss of functioning that can occur when services are suddenly withdrawn (Brekke, Long, Nesbitt, & Sobel, 1997). A similar principle in
vocational rehabilitation is described as “time unlimited follow-along supports” (Becker et al., 2015, p. 122). However, this is rarely measured directly in sufficient depth (Waghorn, Hielscher, & Shield, 2015). The two IPS fidelity scales consider only the continuing availability of in-employment support (Becker et al., 2015, pp. 118-123). The general availability of employment assistance regardless of how long it takes participants to secure employment, is likely to impact on program performance, but it is rarely measured. Sometimes the availability of employment assistance is stipulated by funding agreements. For example, in New Zealand some vocational rehabilitation contracts have required a program to exit a participant from the program if a job outcome has not been obtained within six months (Lockett et al., 2018b). This contracting practice has also been reported in the UK (Burns et al., 2015). Although the likelihood of commencing employment reduces over the time participants are in the program, participants still secure jobs for at least 18 months from program intake (Metcalfe, Drake, & Bond, 2018a).

In summary, program availability can be measured in terms of whether or not there is any formal or informal policy or practice which limits access to the program over time when the program goals have not been met. The nature of these restrictions can then be explored to understand how more continuous program availability can be established.

**Direct and non-direct contact time**

Two types of program activities contribute to program intensity. 1. Direct contact time spent with participants. 2. Non-contact time spent assisting specific individuals or time spent working on their behalf. Administration time or time in training courses which cannot be attributed to helping a particular participant, do not contribute to program intensity since these activities are not considered directly relevant. Relevance of program activities is more difficult to measure with precision. Yet, it can be taken into account at a basic level. Unsuitable or irrelevant program activities could be defined as those outside the scope of the participants’ goals or interests and that were unlikely to be perceived as helpful by the participant. Direct contact and non-contact time can be simply measured by summing the number of contacts and the duration of each contact over a defined period.
McGuire et al. (2011) measured program intensity through the number of contacts between the employment specialist and the participant over three-months. The authors collected data over two years and found that participants averaged six contacts per quarter. This measure counted direct contacts, it did not include non-contact time so is likely an under-reporting of program intensity (McGuire et al., 2011). Studies by Bond, Miller, and Dietzen (1992) and MacDonald-Wilson et al. (1991) defined program intensity as the amount of contact over a month, finding that participants were averaging eight to nine hours of monthly contact. MacDonald-Wilson et al. (1991) measured intensity through a billing database whereas Bond et al. (1992) measured this via service logs completed by the employment specialists.

Rogers et al. (1997) measured program intensity as both the number of contacts and the amount of contact over a defined period. The authors included direct and indirect contacts but excluded travel and preparation time. Across the nine study participants the average contact duration during the 12 month period was 37 minutes, ranging from one minute to nine hours. More than half the contacts were under 15 minutes. The average number of contacts per month per participant was 16.5 (Rogers et al., 1997).

In the Rogers et al. (1997) study the authors relied upon the employment specialists keeping logs of all direct and indirect contacts. The employment specialists coded the mode of contact (telephone, face-to-face), where the contact took place, with whom the contact was with, i.e. participant, family member, or clinician, and recorded the purpose of the contact, i.e. job search, in-work support. Collecting and analysing this data enabled program leaders to better understand program intensity. They found that participants were getting frequent, but often brief contact, that the intensity of support changed depending on whether the participant was job searching, starting employment, or in employment. The authors also learned that employment specialists were making contact at the weekends. This information identified the need for the organisation to support staff through more flexible working hours. This study developed and tested a thorough measure of program intensity but
was not designed to examine any possible relationship between program intensity and program performance (Rogers et al., 1997).

Measuring the frequency and amount of both contact and relevant non-contact could assist program leaders to understand patterns of program intensity and how this needs to vary by participants, and by stage of the vocational rehabilitation process. Since previous evidence suggests a positive relationship between program intensity and program performance there is a potential opportunity to gather and utilise new information to improve the effectiveness of vocational rehabilitation programs.

**Construct operationalisation and measurement**

Table 23 outlines the four candidates for measurable attributes of program intensity. These four attributes are considered sufficient to adequately represent the construct. For each measurable attribute one or more targeted statements are suggested, along with a possible item pool that program leaders could use to identify, measure, and assess program intensity.
Table 23. Operationalising the construct of program intensity

<table>
<thead>
<tr>
<th>Measurable attributes</th>
<th>Targeted statements</th>
<th>Measurement item pool</th>
<th>Method of measurement</th>
<th>Examples of questions</th>
<th>Evidence for attribute¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tailored</td>
<td>The suitability of the intensity of assistance in relation to the participant's assistance needs and circumstances.</td>
<td>1. On starting the program, the program has identified and agreed the intensity of support needs for each participant. I.e. timing, amount, content. 2. The program delivers the agreed support, making changes as needed. 3. The program provides different levels of intensity for different participants. 4. Program intensity for an individual participant can vary across the vocational rehabilitation process. (Score 1 for each anchor that is present).</td>
<td>Interview Program records</td>
<td>Ask program staff about specific caseload participants. Ask them to describe how often they meet each of them, for how long, and how the timing of these meetings were decided on. Review records that outline agreed support and see if these match the assistance given. From the examples, identify if the assistance varies by participants and if the assistance varies by what stage the participants is at, i.e. job searching, in employment.</td>
<td>Bond &amp; Kukla, 2009; Grove, 1999; Johnson et al., 2009; MacDonald-Wilson et al., 1991; McGuire et al., 2011; Taylor &amp; Bond, 2014; Waddell et al., 2008; Wohman et al., 2003.</td>
</tr>
<tr>
<td>Available</td>
<td>The participant can access the program for as long as they need employment assistance to get a job²</td>
<td>There are no formal or informal requirements that participants leave the program if they do not secure employment within a set time period. (Score 1 if the anchor is present).</td>
<td>Interview Program records</td>
<td>Ask program staff if there are any restrictions on how long participants can stay in the program if they have not got an outcome. Also review contracts and program information.</td>
<td>Brekke et al., 1997; Fleming et al., 2012; Metcalfe et al., 2018a.</td>
</tr>
<tr>
<td>Direct contact time</td>
<td>Frequency of contact</td>
<td>1. The program measures the total number of contacts of a standardised type or time unit a participant has over a defined period with the employment specialist or vocational professional. 2. Program staff regularly review this information to understand trends and patterns. (Score 1 for each anchor that is present).</td>
<td>Interview Review client management system and program records</td>
<td>Is there a system to measure contacts? If yes, how is this used by program staff? Does it measure total contacts over a defined time?</td>
<td>Heslin et al., 2011; McGuire et al., 2011; Roger et al., 1997; Taylor &amp; Bond, 2014.</td>
</tr>
</tbody>
</table>

Construct specification: The provision to each participant of the optimum amount of relevant employment assistance while they remain active in the program.
<table>
<thead>
<tr>
<th>Measurable attributes</th>
<th>Targeted statements</th>
<th>Measurement item pool</th>
<th>Method of measurement</th>
<th>Examples of questions</th>
<th>Evidence for attribute</th>
</tr>
</thead>
</table>
| Amount of contact     |                     | 1. The program has a process in place to measure the total hours of contact a participant has over a defined period with the employment specialist or vocational professional. 2. Program staff regularly review this information to understand trends and patterns. (Score 1 for each anchor that is present). | Interview  
Review client management system and program records | Is there a system to measure contacts?  
If yes, how is this used by program staff? Does it measure total hours of contact over a defined time? | Bond et al., 1992; MacDonald-Wilson et al., 1991; Roger et al., 1997. |
| Non-contact time      | Indirect assistance | 1. The program has a system in place to measure the total hours the employment specialist or vocational professional have spent in relation to the participants’ specific case when they haven’t been in contact with the participant. 2. Program staff regularly review this information to understand trends and patterns. (Score 1 for each anchor that is present). | Interview  
Review client management system and program records | Is there a system to measure non-contact time? If yes, how is this used by program staff? | Rogers et al., 1997. |

Notes. 1. **Green**=Category A: The evidence supports the inclusion of the attribute and its effect on program performance. **Red**=Category B: The evidence does not support the inclusion of the attribute and its effect on program performance. **Blue**=Category C: No empirical evidence was identified, but authors have written about the potential for the attribute to influence program performance. 2. The IPS-25 fidelity scale measures that the participant receives ongoing support once in employment. This suggested item in the attributes of program intensity is distinct from the measure in the IPS-25 fidelity scale, as the former investigates unconstrained continuous support pre-employment. 3. This includes contacting employers on the person’s behalf, looking for job vacancies. It does not include any time spent communicating with clinicians or benefits counsellors, attending training courses, or carrying out administration.
**Program Delivery**

Measuring program delivery is a quality dimension most informed by the broader concept of process evaluation (Centers for Disease Control, 2011, p. 1). Program delivery is the application, receipt, and setting of an intervention (Moore et al., 2015; Oakley, Strange, Bonell, Allen, & Stephenson, 2006). Taken quite literally program delivery can be considered to be the way in which the program is discharged to and received by participants. As in the delivery of a courier package or a pizza. The attributes of program delivery that might be important are:

(a) whether it is delivered on time;

(b) the condition of the item on arrival;

(c) how the delivered item matches the customers’ expectations or hopes;

(d) whether the item is delivered in a friendly, efficient, and helpful way;

(e) whether the item is delivered to a place which suits the customer; and

(f) the vehicle, or other mechanism which delivered the item is safe and fit for purpose (Lee, 2017; Parasuraman, Zeithaml, & Berry, 1988).

Viewed in this way the attributes of program delivery likely to be important will vary depending on the product or program being delivered. Perceptions of delivery quality will also vary among recipients of the product, or the participants in the intervention. Participant expectations and participant satisfaction are often measured as if they alone represent the quality of program delivery. However, program delivery is broader than participant satisfaction. It is possible for program delivery to be suboptimal yet received by participants with high satisfaction. The limitations of typical measures of program delivery have been identified previously. For instance, in the service industry measures of program delivery have been criticised for limiting the focus to participant satisfaction (Cronin & Taylor, 1994).

**Program delivery and program performance**

Improving the quality of program delivery could improve program performance (Berkel et al., 2011; Dane & Schneider, 1998; Durlak & DuPre, 2008; Lockett et al., 2018a). Mowbray et al. (2003)
argued that for program sustainability and effective program replication it is important to spend time developing and testing reliable and valid process measures, including measures of program delivery. In the course of testing these measures the critical components of effective programs can be revisited and program theory re-examined at a deeper level.

Identifying the attributes that distinguish the delivery of a vocational rehabilitation program which achieves higher performance from the delivery of a program with lower performance could assist program leaders to improve program performance, particularly once sufficient fidelity to basic program principles has been reached. Measurement of program delivery is important since this has the potential to maximise the effectiveness of an intervention implemented in real-world settings (Oakley et al., 2006). In vocational rehabilitation, whilst qualitative studies have identified specific attributes of IPS programs most important to participants (Areberg et al., 2013; Johnson et al., 2009), the concept of program delivery has not been systematically unpacked to enable clear conceptualisation and agreed measurement. This is likely to part explain why there is so little research investigating the attributes of program delivery which distinguish the operation of higher performing programs from lower performing ones, beyond program adherence to prescribed content (Drake et al., 2012, p. 44).

**Conceptualising program delivery**

Different views have emerged as to how program delivery should be defined. In prevention research, Dane and Schneider (1998) interpreted program delivery as meaning “the qualitative aspects of program delivery that are not directly related to the implementation of prescribed content” (p. 45). The authors identified staff enthusiasm and responsiveness and staff attitudes towards the program being implemented as important influences on delivery quality. In contrast, Hasson (2010) described program delivery in relation to program outcomes, that is “the appropriateness of the delivery process for achieving what was intended” (p. 2).

In education, Dusenbury et al. (2005) examined the way teachers delivered a standardised drug prevention program in schools. Program delivery was described as the “quality of process”, namely the ways in which teachers deliver the prescribed curriculum (p. 302). The authors developed a quality of
process scale based on observational ratings of teachers in the classroom. Their measure of program delivery quality encompassed teacher-student interaction, teacher enthusiasm, the way teachers communicated the intervention goals, the levels of student engagement and attentiveness, and how much students expressed their own opinions. Teachers’ understanding of the intervention had a positive correlation with scores on this quality of process scale. Program delivery quality as measured by this scale was found to be higher when teachers had a good understanding of the educational program (Dusenbury et al., 2005). In this example, the item or product to be delivered was a standardised drug education program, and the teacher was conceptualised as the primary mechanism for delivering the product, hence the measure of program delivery had a focus on teacher attributes.

In vocational rehabilitation, program delivery is how well a defined set of evidence-based practices are supplied to and received by participants in the vocational rehabilitation program. The primary, but not the sole mechanism for supplying evidence-based practices is the vocational staff. In evidence-based practices vocational staff are likely a strong, direct influence on program implementation and performance. The IFVR distinguishes between the dimensions of implementation quality, employment specialist expertise and program delivery (Lockett et al., 2008a). The operational specifications need to reflect this distinction. Therefore, the attributes of program delivery in vocational rehabilitation programs are broadly conceptualised as follows, and do not include attributes of the employment specialist. However, they do include attributes of non-vocational staff.

Program delivery includes, but is not limited to, the delivery of a sufficiently resourced program that is accessible and appealing to participants, that facilitates continuing engagement with participants, and provides planned and consistent employment assistance in line with recognised industry standards and participant expectations. Based on this broad conceptualisation of program delivery six candidates for measurable attributes were identified.

1. Performance standards. The program needs to follow minimum performance standards particularly those which protect participant rights, comply with health and disability standards, and ensure the safety and security of participants. The delivery of the program also needs to
observe and conform to mandatory standards, for example those required by relevant legal, government, or industrial standards (Hermann & Palmer, 2002; Hermann & Provost, 2003; Lee, 2017; Parasuraman et al., 1988).

2. Adequacy of program resources. These are the material features of program delivery considered essential for the successful delivery of a vocational rehabilitation program. Adequate resources need to be available and fit for purpose (Lee, 2017). For example, these can include the functional state of the physical premises, sufficient vocational and non-vocational staffing, and sufficient and reliable equipment, such as computers and photocopiers.

3. Program accessibility. This is the ease with which participants can access the program premises, i.e. the physical building itself, including parking areas, and the location of the building, as well as access to program equipment.

4. Participant perceptions of the delivery environment. This is how the operating environment looks and feels to participants through all stages of participation (Lee, 2017). It can include how the physical environment is perceived and perceptions of how non-vocational staff interact with participants. For example, how welcoming, friendly, and respectful staff are perceived to be, and how clean and comfortable the physical environment is perceived. The physical environment can include access areas, waiting areas, meeting rooms, availability of refreshments, and the perceived state of bathrooms.

5. The communication of the delivery content. This attribute is concerned with the information provided to participants on what the program will consist of. So that participants know in advance what to expect from the program, and how it will operate.

6. Quality assurance. This is the delivery of the program in a planned, consistent, and reliable way to each participant (Wehman et al., 2003).

**Performance standards**

One logical starting point for measuring and improving program delivery is the review of, or the identification of, any existing standards and requirements for program delivery quality. Programs
leaders could review relevant legal, industry, and government standards and performance expectations outlined in contracts. In New Zealand, vocational rehabilitation programs need to comply with the Health and Disability Standards (Standards New Zealand, 2018). The standards are intended to ensure the safe provision of services to participants through good service quality and continuous quality improvement. The standards cover participant rights, organisational governance and management arrangements, infection control, and minimising restraint. Organisations providing vocational rehabilitation programs need to comply with those standards that are relevant to vocational rehabilitation and follow the auditing processes. This is a usual requirement from program funders and is typically written into the terms and conditions of contracts.

In addition to government standards there may be relevant industry standards. For example, the New Zealand Disability Support Network (NZDSN) developed eight practice guidelines for vocational rehabilitation programs working with disabled people. These guidelines were developed following a literature review, and consultation with employment service providers and disabled people (NZDSN, 2018). Together the guidelines stipulate a focus on participant choice and self-determination, involving other people and agencies who are relevant to helping the participant in their working life, ensuring a personalised service, the way staff engage employers, the knowledge and competencies of the vocational rehabilitation staff, and conducting program evaluation. Whilst there is no legal or financial requirement to comply with these standards yet, vocational rehabilitation providers are encouraged to align practices to the standards as part of their commitment to improving employment outcomes for disabled people, and their commitment to quality improvement (NZDSN, 2018; Prinz, Lockett, Sternsek, & Arends, 2018).

Whilst it is necessary for vocational rehabilitation programs to comply with minimum service standards, an over-burdensome compliance culture can hinder improving program performance (Kitching, 2006). Program leaders should therefore monitor time spent on compliance activities, to reduce the burden of compliance, and optimise compliance efficiency. Compliance strategies can also
be combined with quality improvement processes that are intended to enhance program performance (Kitching, 2006).

It is important that program leaders do not just look inwards to monitor program performance, but seek information from national, and international standards, and best practices for comparable programs. Looking outwards is considered more likely to align the program standards with community expectations and contribute to performance improvement (Australian Government, 2001). It is also important for program leaders to identify those attributes that distinguish excellent program delivery from good or average delivery (Hermann & Provost, 2003). Whilst programs may need to have systems in place that ensure compliance with performance standards particularly those which protect participant rights, comply with health and disability standards, and ensure the safety and security of participants, it is also suggested that programs look beyond compliance with industry and government standards and strive for excellence (Hermann & Provost, 2003).

**Adequacy of program resources**

The adequacy of program resources is another measurable attribute within the construct, program delivery. Adequacy of resources could include the material features of program delivery essential for the delivery of a vocational rehabilitation program to be effective. This would include the physical premises, the equipment, and the number of non-vocational staff.

To improve program performance program leaders could regularly assess the adequacy of their premises, equipment, and staffing focusing on those aspects most important to the provision of planned and consistent employment assistance and the resources identified as important by program participants. This is likely to include, but not be limited to computers that are up to date, reliable, and with reasonable internet speed, and facilities to print information and to copy resumes. Equipment should be regularly serviced, safe, and in working order (Lee, 2017). Adequacy of resources would also include whether for example, there are sufficient rooms to hold one-to-one meetings with participants and cars for staff to use for job development, if public transport is not adequate.
**Program accessibility**

Program accessibility covers the premises and the program equipment. It includes the ease with which participants and staff can access the program premises, i.e. the physical building itself and the location of the building, as well as participants access to the program equipment. In a study examining the perceptions of people with disabilities on the important qualities of vocational rehabilitation programs, Timmons, Schuster, Hamner, & Bose (2002) interviewed 19 program participants. Nine of the 19 participants had a diagnosis of a mental illness. Access to program resources was identified as important. Resources included computers, job vacancy lists, information on training opportunities and information on transportation routes (Johnson et al., 2009). Accessibility also includes how program staff (vocational and non-vocational) help participants navigate the use of the program resources (Corbière & Lanctôt, 2011; Timmons et al., 2002).

Program leaders need to regularly review how accessible their physical premises and the program resources are to the target participants. This could include whether the premises are easily accessible by local public transport routes, whether there are spaces for parking cars which are nearby and affordable, and the availability of lifts. In terms of program resources this could include clear instructions located by equipment, or the provision of training in how to use the equipment when participants are accepted onto the program.

**User perceptions of the delivery environment**

The program delivery environment represents how the operating environment looks and feels to participants. For example, this includes how welcoming, friendly, and respectful it feels. Timmons et al. (2002) made a distinction between how participants perceive the social environment and how they perceive the physical environment. The former was about participants sense of belonging, how comfortable they were with all program staff, how warmly they were greeted on arrival, and the level of genuine interest staff took in them (Timmons et al., 2002). The physical environment was described as how welcoming and accommodating the physical premises are perceived. The look and feel of the environment might be influenced by positive messages in waiting areas about supporting people with
mental illnesses to commence employment, stories of success, and quiet spaces for people to wait if they are feeling anxious. Environments can reflect and support individual recovery journeys. To do so they should feel open, accessible, hopeful, and non-stigmatising. Staff only areas can exacerbate a sense of exclusion or difference, sometimes described as ‘otherness’ (Deegan, 1988).

The value of participants’ perceptions of the delivery environment is illustrated by a study of nine participants in a Swedish IPS program. The study found participants endorsed the importance of the first meeting and the space in which that takes place (Topor & Ljungberg, 2016). The study did not however examine the contribution of this attribute of program delivery on participants’ employment outcomes. In another study examining a new approach to strengthen inter-agency collaboration in New Zealand, particular attention was directed at creating a non-stigmatising neutral community environment. The aim was to create a feeling more like a public library, and less like a mental health service (Nepe et al., 2011).

As well as direct influencing program performance, it is likely participants perceptions of the delivery environment impact on other dimensions of implementation to influence program performance. For example, Waghorn, de Souza, Lloyd, and Rampton (2009b) found that the quality of the physical environment was associated with a positive perception of the working alliance between the participant and the employment specialist. The environment was rated by participants into one of three-categories (below average, average, above average). Aspects of the environment rated were the noise levels, the availability of private spaces for conversations between participants and staff, how comfortable the furniture was, and the general appearance of the waiting areas (Waghorn et al., 2009b).

To improve the perceptions of the physical and social environment program leaders can check that their signage conveys a positive, safe, and inclusive tone. Program leaders can ensure that the very first contact with program staff is welcoming and positive, and that this type of contacts continues throughout the individual’s participation in the program. Information can be gathered through their own observations and through talking to participants.
Communication of the delivery content

The fifth candidate for a measurable attribute of program delivery quality is the communication of the delivery content. It is important that participants as well as referrers, have information on how the vocational rehabilitation program will work and the processes that will be followed so they know what to expect (Areberg et al., 2013; King, Cleary, Harris, Lloyd, & Waghorn, 2011). For example, such background information could include the type of vocational assessment used, the purpose of the assessment, and how the vocational assessment will assist the person. This could help participants to anticipate what personal information is most relevant, and to understand whether and how their personal information will be shared with other parties. This is particularly important in situations where participants could be concerned about providing information that might lead to their exclusion from a program, when in fact the information is used to tailor the nature and intensity of employment assistance provided (King et al., 2011).

Quality assurance

Employment assistance should be delivered in a planned, consistent, and reliable way to each participant. Planned, consistent, and reliable employment assistance is enhanced by having clear processes in place to formally accept participants onto the program and to let participants know what to expect next, so there are few surprises. Programs leaders should have a system in place to check the quality of program delivery for each participant. Programs are expected to benefit from regularly communicating with participants about what aspects of program delivery are important to them, what is working well, and whether there are any gaps (Hermann et al., 2002).

In some cases however, a lack of value of an attribute by participants may be influenced by lack of experience, or exposure to that attribute. For example, the findings of the lack of value placed on peer support by participants (Johnson et al., 2009; Liu, 2011), may in part be explained by the fact that the use of peer workers in vocational rehabilitation programs is not yet commonplace. So peer support with job search or in-work support may not be something participants had experienced and therefore had not considered as an important attribute of program delivery. Some participants will not articulate
what is important to them until after they have experienced it. It is important to evaluate any changes made to the program delivery, particularly those not identified by participants, so that the extent that these changes are improving service delivery or not, can be understood.

**Construct operationalisation and measurement**

Table 24 outlines the six candidates for measurable attributes of program delivery. These six candidates are considered sufficient to adequately represent the construct of program delivery quality. For each measurable attribute one or more targeted statements are suggested, along with a possible item pool that program leaders could use to identify, define, assess, and improve the quality of program delivery.
Table 24. Operationalising the construct of program delivery

<table>
<thead>
<tr>
<th>Measurable attributes</th>
<th>Targeted statements</th>
<th>Measurement item pool</th>
<th>Method of measurement</th>
<th>Examples of questions</th>
<th>Evidence for attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance standards</td>
<td>The identification of minimum performance standards Compliance with standards</td>
<td>1. Within the last year a review of the attributes of program delivery required by legal, government, or industrial standards has been conducted. (Score 1 if the anchor is present). 1. Program delivery complies with relevant and required standards. 2. The program monitors the time spent on compliance activities. 3. Adjustments to the time spent on compliance are made to optimise compliance efficiency. (Score 1 for each anchor that is present).</td>
<td>Interview Review program documentation</td>
<td>Find out from program staff whether they have looked for the minimum standards of program delivery required for the type of program in operation. When was this undertaken? Interview Review program documentation</td>
<td>Corbière &amp; Lancôt, 2011; Hermann &amp; Provost, 2003; NZDSN, 2018; Timmons et al., 2002; Topor &amp; Ljungberg, 2016; Welman et al., 2003; Hermann et al., 2002; Hermann &amp; Palmer, 2002; Hermann &amp; Provost, 2003.</td>
</tr>
<tr>
<td>Strives for excellence</td>
<td>The program has (a) identified the attributes of program delivery that are associated with higher performance (b) a quality plan to improve the identified attributes. (Score 1 for each anchor that is present).</td>
<td></td>
<td></td>
<td>Ask program staff about national and international standards, or best practices for comparable programs. What do they know about these standards, how is program delivery aligned with the standards.</td>
<td>Australian Government, 2001; Hermann &amp; Provost, 2003.</td>
</tr>
<tr>
<td>Adequacy of program resources</td>
<td>The material features of program delivery are available, adequate, and fit for purpose (a) premises (b) equipment (c) staffing. (Score 1 for each anchor that is present).</td>
<td>Observation Interview Review program documentation</td>
<td>Ask program staff how they know the program resources are adequate and fit for purpose. Ask specifically about the size and availability of meeting rooms, access to cars, if needed, age of computers, if any equipment is awaiting repair and for how long. Ask about demand for the program, and the availability of staff. Also review any documentation e.g. service records.</td>
<td>Corbière &amp; Lancôt, 2011; Koletsis et al., 2009; Timmons et al., 2002; Lee, 2017.</td>
<td></td>
</tr>
<tr>
<td>Program accessibility</td>
<td>The following are accessible to the target program participants (a) physical premises (b) location of premises (c) equipment (Score 1 for each anchor that is present).</td>
<td>Observation Interview participants</td>
<td>Ask about local public transport routes, parking for participants. Observe signage, stair and lift access. Find out what equipment is available for participants to use, ask participants about their experience of using any equipment.</td>
<td>Corbière &amp; Lancôt, 2011; Johnson et al., 2009; Timmons et al., 2002.</td>
<td></td>
</tr>
<tr>
<td>Measurable attributes</td>
<td>Targeted statements</td>
<td>Measurement item pool</td>
<td>Method of measurement</td>
<td>Examples of questions</td>
<td>Evidence for attribute&lt;sup&gt;1&lt;/sup&gt;</td>
</tr>
<tr>
<td>------------------------</td>
<td>---------------------</td>
<td>-----------------------</td>
<td>-----------------------</td>
<td>-----------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>The perceptions of the delivery environment</td>
<td>The perceived nature of the delivery environment</td>
<td>Participants report that (a) the physical areas are clean, comfortable welcoming (b) all interactions are positive, welcoming, and respectful (c) basic refreshments are available (d) the signage is respectful, clear and inclusive. (Score 3 if three or more are present, 2 if two are present, 1 if one is present).</td>
<td>Observation interview participants</td>
<td>Ask participants how they perceived the welcome for their first and subsequent meetings. Ask about the feel of the waiting areas, meeting rooms, and bathrooms, and their ongoing interactions with non-vocational staff. Check that basic refreshments are available for participants i.e. tea, coffee, water.</td>
<td>Areberg et al., 2013; Deegan, 1988; Nepe et al., 2011; Parasuraman et al., 1988; Topor &amp; Ljungberg, 2016; Timmons et al., 2002; Waghorn et al., 2009.</td>
</tr>
<tr>
<td>The communication of delivery content</td>
<td>Participants know how the program will operate</td>
<td>Information is available in different formats that lets participants know what to expect (a) before and during referral (b) once accepted onto the program. (Score 1 for each anchor that is present).</td>
<td>Interview participants Review program documentation</td>
<td>Review all information provided to participants. Is it in a suitable format for the target population e.g. language, text-based, online, recovery focused. Find out if program participants report understanding the way in which the program will operate.</td>
<td>Areberg et al., 2013; King et al., 2011.</td>
</tr>
<tr>
<td>Quality assurance</td>
<td>The delivery of planned, consistent, and reliable employment assistance</td>
<td>Quality assurance processes are in place. They cover (a) formally accepting participants onto the program (b) agreeing a vocational plan which includes how the person will deal with disclosure, any work restrictions and any financial impact of commencing employment (c) agreeing a job search plan (d) agreeing an on-going support plan offering planned and unplanned responses (e) checking with all participants how they are experiencing the employment assistance. (Score 3 if three or more are present, 2 if two are present, 1 if one is present).</td>
<td>Observation Review program documentation</td>
<td>Ask program staff for a copy of their quality assurance policy and practices. Review and observe program practices, is there a clear set of employment assistance practices that staff follow which include the five anchors outlined.</td>
<td>Wehman et al., 2003.</td>
</tr>
</tbody>
</table>

<sup>1</sup> For explanation of color coding: 
- **Green** = Category A: The evidence supports the inclusion of the attribute and its effect on program performance. 
- **Red** = Category B: The evidence does not support the inclusion of the attribute and its effect on program performance. 
- **Blue** = Category C: No empirical evidence was identified, but authors have written about the potential for the attribute to influence program performance.

Notes: 1. **Green** = Category A: The evidence supports the inclusion of the attribute and its effect on program performance. **Red** = Category B: The evidence does not support the inclusion of the attribute and its effect on program performance. **Blue** = Category C: No empirical evidence was identified, but authors have written about the potential for the attribute to influence program performance.
Discussion and Limitations

This sixth chapter has explained the methods used to develop clear operational construct specifications for the eight dimensions of implementation quality identified in the IFVR. Next, the results from the conceptual analysis for six of the dimensions of implementation quality were presented including the implications of the findings for program leaders.

The first question guiding the conceptual analysis explored whether discrete dimensions of program implementation quality other than fidelity to existing evidence-based practices can be identified and described sufficiently to enable valid and reliable measurement? The results from the conceptual analysis indicate that this is feasible. To help generate discrete and independent dimensions, construct specifications were generated to ensure the development of each measurable attribute took into account related measurable attributes of other dimensions of implementation quality.

Whilst the results of this conceptual analysis were reported in sequence, the process of working on each was more iterative than linear. By conducting the conceptual analysis for each dimension in parallel with the other seven, possible construct overlaps were recognised and addressed throughout the process. For example, the conceptualisation of program evaluation included building staff capacity in program evaluation. This attribute was not duplicated in the conceptualisation of employment specialist expertise. Likewise, the construct defining non-evidence-based practices did not duplicate a measurement item to identify whether individual program staff had knowledge of evidence-based practices. This is because the construct of employment specialist expertise required an independent focus on the knowledge and skills of individual employment specialists.

Whilst each dimension of implementation quality in the IFVR is hypothesised to have a direct influence on program performance there are also likely to be multiple interrelationships among the nine dimensions of implementation quality. For example, the delivery of targeted technical assistance requires that some form of program evaluation has already identified the problem to be targeted. In addition, program intensity tailored to the employment assistance needs of individual participants is also likely to positively influence participant responsiveness. Although steps were taken to increase the
independence of each dimension and to remove redundancy, these interrelationships mean that some overlap among dimensions remains and cannot be removed altogether. Once the measures have been further developed and empirically tested, future studies of the internal structure can explore and quantify these interrelationships.

The second question guiding the analysis asked: Can these dimensions be operationalised into a measure that can be successfully applied to vocational rehabilitation programs? The results of the conceptual analysis for six of the dimensions of implementation partially operationalised in this chapter suggest that a new measure for vocational rehabilitation can be so constructed. The targeted statements and example measurement items are a good starting point to target each dimension and attributes as conceptualised, but each construct and set of measurable attributes will need further development and empirical testing, in particular, to explore the construct validity of the tool as a whole.

The last question guiding this conceptual analysis was: Can these dimensions of implementation quality be generalised to other forms of psychosocial interventions? The analysis has drawn on the wider literature from across the health and social sciences. Although there is little direct evidence that this is the case, the many parallel findings in the literature suggest that with further operationalisation, generalisability to other psychosocial programs is likely.

There was limited empirical evidence from the science of vocational rehabilitation relating to the six dimensions of implementation quality outlined in this chapter. The conceptual analysis therefore also drew on empirical and theoretical literature from the wider social sciences. As a result, it is possible that some of the measurable attributes identified are less important influences on the performance of vocational rehabilitation programs for people with severe mental illnesses than has been outlined. As the available empirical literature for vocational rehabilitation grows, these measurable attributes and targeted statements can be updated and refined. Whilst there are limitations to this analysis, particularly the need for further development of the measures through empirical testing, the targeted statements provide a useful starting point for program leaders to begin measuring dimensions of implementation quality, to complement measurement of program fidelity.
Chapter Conclusions

This chapter presented the conceptual analysis for six of the dimensions of implementation quality in the IFVR. 1. Technical assistance. 2. Program evaluation. 3. Removal of non-evidence-based practices. 4. Complementary programs. 5. Program intensity. 6. Program delivery. It is anticipated that this conceptualisation of these constructs could inform the further development and testing of a tool to identify, define, and measure discrete dimensions of implementation quality to complement existing measures of implementation fidelity. The resulting measure is expected to assist practitioners, researchers, policymakers, and program administrators to further improve program performance. The next chapter, Chapter 7, partially operationalises the seventh dimension of implementation quality in the IFVR, namely employment specialist expertise.
Chapter 7: Operationalising Employment Specialist Expertise

Introduction

Chapter 6 outlined the methods for the conceptual analysis for the eight new candidate dimensions of implementation quality outlined in the IFVR. The chapter next reported the results of the conceptual analysis for six of these dimensions. This chapter (Chapter 7) describes the conceptual analysis for the seventh dimension of implementation quality in the IFVR, namely employment specialist expertise. The chapter maps the conceptual space for each measurable attribute considered essential to adequately represent the construct. Identifying and measuring the attributes most likely to represent the construct could help program leaders to identify skills gaps, formulate human resource strategies, and improve the expertise of the workforce (Kumar, 2014). To support a logical and systematic conceptual analysis the following questions were investigated.

1. How should employment specialist expertise be described and defined in the context of vocational rehabilitation for people with severe mental illnesses?
2. Are there identifiable and measurable attributes that distinguish higher performing employment specialists from lower performers?
3. How can program leaders use the findings from this conceptual analysis to identify, describe, and improve the expertise of employment specialists in vocational rehabilitation programs?

The chapter starts by describing the analytical methods for operationalising the construct of employment specialist expertise. Next, the reasons employment specialist expertise is likely to be a candidate for explaining some of the unexplained variance in program performance are considered, along with an overview of the role of employment specialists in vocational rehabilitation programs for people with severe mental illnesses.

Based on the results of the first review of the literature a broad construct specification for employment specialist expertise is proposed. This specification identifies four potential candidates for measurable attributes of employment specialist expertise: knowledge, skills, attitudes, and other employment specialists’ characteristics. Other employment specialist characteristics could include age,
gender, qualifications attained, ethnicity, and level of education. The challenges involved in measuring these attributes are then discussed.

The next section of the chapter describes the findings from a targeted review of the empirical evidence. This identifies a set of behavioural indicators thought to be common to higher performing vocational staff. These behavioural indicators are used as a framework to identify and discuss the mostly likely knowledge, skills, attitudes, and other individual characteristics underpinning these behaviours. The discussion of the behavioural indicators is informed by the wider literature review and examples of how others have conceptualised and operationalised the construct. Suggestions are also made for how program leaders could use these measures or develop new measures, to identify, describe, and improve employment specialist expertise.

The chapter concludes with a discussion of the findings and limitations of this conceptual analysis. This includes a synthesis of the findings showing how the dimension of implementation quality, namely employment specialist expertise, can be partially operationalised, using the behavioural indicators identified from the empirical literature, that are common to high performers.

**Methods**

The methods followed the seven-step process of conceptual analysis outlined in Chapter 6. This involved an iterative process of reviewing the empirical and conceptual literature against the construct specification and candidates for measurable attributes, continually revising the specification, measurable attributes, and targeted statements pool.

Given the greater volume of previous research in this area compared to other dimensions, a targeted review of the empirical evidence was conducted. The aim of this additional step was to identify the likely behavioural indicators associated with higher performance. This led to a discussion of the findings from the wider conceptual and empirical literature, in order to target the most likely candidate attributes for improving employment specialist expertise beyond baseline training competencies.
In the targeted review of empirical evidence, studies were included if they combined at least one measure of a candidate attribute of employment specialist expertise with a relevant measure of program performance. The measure of program performance could be either participant employment commencements or participant job tenure. Studies were excluded from this step in the conceptual analysis if they did not report a measure of program outcomes or did not measure program provider competencies at the level of the individual employment specialist. Strength of empirical evidence was classified into three levels to prioritise the main findings.

1. At least one empirical study identified an association between the attribute of employment specialist expertise and higher program performance.

2. There was no empirical evidence from vocational rehabilitation identified which linked the attribute to program performance, yet the attribute remains a promising candidate since it has good conceptual clarity, and there is a logical and potentially causal link between the attribute and program performance.

3. The attribute has been written about by researchers in vocational rehabilitation. To date no empirical evidence links the attribute to program performance. In addition, the attribute lacks conceptual clarity, or the potential for a causal connection between the attribute and program performance seems unlikely.

**Identifying relevant literature**

To identify the relevant literature, the following initial search terms were used: ‘expertise, competency, experience, abilities, staff performance’. These were combined with the terms: ‘program outcomes, program performance, vocational rehabilitation’, and ‘supported employment’. As articles were sourced and the operational specification developed, the literature search continued in a second iteration using additional terms relevant to newly identified measurable attributes. The terms used included: ‘knowledge, skills, attitudes, behaviours, values, control, self-efficacy, characteristics’. These terms were combined with the previously identified search teams. Google Scholar was the primary database used supplemented with searches of Medline, PubMed, and PsycINFO. As relevant articles
were identified and reviewed their reference lists were searched to identify further potentially relevant literature. Citation searches were then conducted to find further articles citing the articles that had already been identified.

Following a similar approach to the conceptual analysis for the other six dimensions of implementation quality, the review of literature gave priority to research from vocational rehabilitation. However, because this dimension of implementation quality has a focus on the performance of individuals, it also drew on theories of human motivation and behaviour considered relevant to vocational rehabilitation, as well as theories from the broader literature on the science of psychosocial program implementation. In parallel to this step, the literature was examined for published measures relevant to employment specialist expertise and to each potentially measurable attribute. The purpose was to see how others had attempted to conceptualise and measure employment specialist expertise.

To prioritise the attributes of employment specialist expertise, all the evidence gathered was organised into the three categories used in the operationalisation of the other dimensions of implementation quality. Category A: The evidence supports the inclusion of the attribute and its effect on program performance. Category B: The evidence does not support the inclusion of the attribute and its effect on program performance. Category C: No empirical evidence was identified, but authors have written about the potential for the attribute to influence program performance.

The Employment Specialist’s Role in Program Performance

The expertise of staff implementing an intervention or program is already known to be an important dimension of implementation quality (Berkel et al., 2011; Breitenstein et al., 2010; Dusenbury, et al., 2005). In evidence-based practices in vocational rehabilitation the program staff who have the most contact with participants are those in employment specialist roles. These staff are likely to have the greatest direct influence on participant motivation, behaviour, and engagement with the program, and therefore on program outcomes (Knaeps, Neyens, Donceel, van Weeghel, & van
Audenhove, 2015; Lockett et al., 2018a). For many participants the employment specialist personifies the vocational rehabilitation intervention (Gammelgaard et al., 2017).

Employment specialists have a key one-to-one role in all aspects of supporting participants to return to work, and to stay at work (Corbière et al., 2014a). Whilst some participants obtain their own jobs with little direct assistance, most require direct assistance from the employment specialist to find, secure, and retain suitable employment. This is because vocational rehabilitation programs are designed specifically for those with the biggest challenges to obtaining employment (Wehman et al., 2003). The more intensive the vocational rehabilitation program, the greater influence the employment specialist is likely to have on participant outcomes (Knaeps et al., 2015). The expertise of the employment specialist for supporting people with severe mental illnesses is considered likely to account for some of the variance in the performance of evidence-based programs in vocational rehabilitation (Corbière et al., 2014a; Drake & Bond, 2011; Drebing et al., 2012; Lockett et al., 2018a).

The key components of the employment specialist role in vocational rehabilitation have been described as follows.

1. Working on a one-to-one basis with program participants to plan job commencement or return to employment, based on the participants’ preferences, skills, and interests (Corbière et al., 2014a);
2. Coordinating support with other stakeholders providing non-vocational services, around participants’ employment aspirations (Whiteford et al., 2014).
3. Providing ongoing advice, general problem-solving, and practical assistance to participants about commencing or returning to employment, and assistance to maintain employment, or to transfer to another job (Waghorn et al., 2015).
4. Working with families, local businesses, and other community members to help participants to get and keep jobs (Swarbrick et al., 2017).
5. Working directly with employers to identify suitable jobs, arrange job starts, prepare people for specific jobs, and provide on-site training to new workers. These activities involve
supporting employers as well as program participants (Glover & Frounfelker, 2013; King & Waghorn in press-a, b).

There is an emerging consensus on the central role of the employment specialist. However, the specific influence of different attributes of employment specialists on program performance remains unknown and unclear (Catty et al., 2008; Corbière et al., 2014a, 2017; Glover & Frounfelker, 2011, 2013; King & Waghorn, 2018; Taylor & Bond, 2014; Whitley, Kostick, & Bush, 2010). There are mixed findings as to which attributes of employment specialist expertise are common among higher performing staff and therefore most likely to have the strongest impact on program outcomes. For example, some employment specialists have low caseload attrition and regularly secure employment outcomes for almost everyone on their caseload, when others using the same recommended practices do not (Morris et al., 2014).

**Developing the Construct Specification**

Competency in the delivery of an intervention can be underpinned by personal attributes and characteristics such as experience in the role, education, communication skills, values and beliefs, self-efficacy, technical abilities, amount of training provided, and interpersonal skills (Breitenstein et al., 2010). Expertise however is more than baseline competencies. Expertise has been summarised as the competencies required for excellent performance in a role (Campion et al., 2011; Schippmann et al., 2000). Expertise in a role is usually multi-dimensional, being composed of all the actions that an individual carries out to effectively support the organisational objectives (Campbell & Wiernik, 2015). Expertise is therefore best conceptualised as a bundle of skills, knowledge, or attitudes, combined with other personal attributes and characteristics (Campion et al., 2011; Corbière et al., 2014a; Marrelli, Tondora, & Hoge, 2005; Taylor & Bond, 2014). To address the practical problem of how to improve the effectiveness of vocational rehabilitation programs, the focus narrows to an operational specification that is about the knowledge, skills, attitudes, and other individual attributes which distinguish higher performing employment specialists from lower performers.
Some attitudes, knowledge, skills, and personal qualities can be difficult to directly observe since they are internal states. Internal states are often inferred from behaviours or practices (Campion et al., 2011). In addition, it can be difficult to distinguish among attitudes, knowledge, and skills, and difficult to differentiate these from observable behaviours. This is because some attributes overlap in terms of both theory and observation. For example, by observing an employment specialist listening carefully to a participant’s preferences and interests, it could be inferred that the employment specialist places a high value on participant perspectives, or perhaps believes that understanding each participant is critical to finding a successful pathway into employment for that person. Deeper attitudes concerning the employability of people with mental illnesses could also be present. Any of these attitudes help the employment specialist take the time to fully understand each participant.

Another way to classify employment specialist attributes is in terms of the extent that they are fixed or trait-like, such as personality characteristics, age, gender, and ethnicity, versus the extent that the attributes represent more fluid psychological states, such as opinions, attitudes, beliefs, self-efficacy, motivation, and self-esteem. This is a potentially useful distinction because attributes that are more changeable may also be more amenable to training and development (Kumar, 2014).

Previous research on employment specialist competencies and expertise has often failed to make these conceptual distinctions. This lack of conceptual precision may have contaminated the utility and construct validity of previous studies attempting to form and measure a representative construct (Kanfer, Chen, & Pritchard, 2008, p. 21). This could partly explain why so many findings relating to employment specialist expertise have been mixed or inconclusive (Coursey et al., 2000; King & Waghorn, 2018; Williams, Fossey, Corbière, Paluch, & Harvey, 2016).

**The attributes of employment specialist expertise**

Identifying the relevant personal qualities, attitudes, knowledge, and skills, through observable behaviours is challenging. To overcome these difficulties the following definitional principles were applied to each candidate attribute.
**Attitudes**

Attitudes include beliefs, values, individual self-efficacy, expectations of success, and sense of personal agency and control (Campion et al., 2011; Taylor & Bond, 2014). These qualities can remain unexpressed or unobserved and can therefore be difficult to measure. Yet attitudes can be indirectly observed or inferred through observing patterns of behaviour. For instance, by directly observing interactions between employment specialists and participants, between employment specialists and clinicians, and by observing interactions between employment specialists and employers.

**Knowledge**

Knowledge is the accumulated theoretical and practical understanding about matters such as facts, rules, principles, and guidelines, that are needed to perform a job (Campion et al., 2011). Knowledge is acquired through learning and experience on the job. Some knowledge can be measured, but much can be complex, abstract, and not easily measurable (Marrelli et al., 2005). Two types of knowledge have been distinguished, described as tacit and explicit knowledge.

Tacit knowledge is considered to be unconscious and therefore difficult to verbalise and to store explicitly. For example, knowing how to ride a bicycle can be measured by a riding test, but the person with the knowledge can find it impossible to explain this knowledge to others. Hence, people are often unaware of their tacit knowledge until they are specifically asked about that. Even when asked, many people are unable to explain to others how they successfully perform everyday activities such as driving a car. Tacit knowledge is acquired through hands-on practical experience and practice. Tacit knowledge can be transferred through interactions, observing others with that knowledge, and through a build-up of shared understanding and trust (Haldin-Herrgard, 2000). For example, apprentices learn about the art of carpentry from spending time and working alongside an experienced carpenter, but this knowledge is not necessarily accessible and may not be able to be used directly in training others.

Conversely, explicit knowledge can be articulated, verbalised, accessed, and more easily transferred to others. Explicit knowledge can be stored, for example in books, teaching materials and manuals, and online (Haldin-Herrgard, 2000). Whilst a distinction can be made between these two
types of knowledge in their conceptualisation, it can be hard to distinguish between them in practice. Both can be indirectly observed or inferred through how well novel problems are solved. The presence of explicit knowledge can be inferred through questioning and examination (Dreher, Bond, & Becker, 2010). For example, in advanced driving instruction for police, instructors can require the student to maintain a running verbal commentary to describe what they see, their actions, and the decisions they make during a fast pursuit. The driving exhibited represents tacit knowledge and skills, while the commentary represents the explicit knowledge and meta-cognitive skills involved in observing and reporting complex behaviours, while at the same time performing those actions (Antoñanzas, Salavera, & Teruel, 2015).

**Skills**

Acquired skills are the demonstrated abilities and methods for completing tasks leading to goal attainment and high performance. They can be domain-general, or domain-specific. For example, a general skill for most employment specialist roles is time-management, another is verbal communication. Domain-specific skills could relate to the type of job being sought for the participant, and to the technology and equipment found in that workplace. The measurement of skills can be job-specific or task-specific (Campion et al., 2011). The development of general and specific skills in vocational staff is likely to be an important area for program leaders to focus on to improve program performance (Corbière et al., 2014a; Glover & Frounfelker, 2011). This is because skills can be taught through appropriate training, supervision, feedback, and through leaders who role-model the skills (Campbell & Wiernik, 2015; Coursey et al., 2000; Kukla & Bond, 2012).

**Other individual characteristics**

Other relevant individual characteristics include relatively fixed personal attributes namely, age, gender, ethnicity, language spoken, culture, educational attainment, vocational qualifications, years of experience in role, communication style, and other personality traits. Personal characteristics of this type are potentially observable and can often be directly measured or classified (Marrelli et al., 2005).
However, their relevance to the performance of employment specialists is less clear and the findings of previous studies are mixed (Corbière et al., 2017; Coursey et al., 2000; Kukla & Bond, 2009).

The measurable attributes relevant to employment specialist expertise differ in their observability. The least observable attributes are those related to internal states such as attitudes, beliefs, values, and tacit knowledge. The most observable are the more fixed demographic characteristics such as age, gender, qualifications, and experience (see Figure 11).

![Figure 11. The observability of the candidate attributes of employment specialist expertise](image)

**The Practices of Higher Performing Employment Specialists**

Eleven papers were identified in the targeted review of the empirical evidence (see Table 25). These eleven papers provide an important starting point from which to identify the most promising candidates for measurable attributes of employment specialist expertise. There are limitations in the design of almost all these studies. For example, a lack of control groups or the use of self-report measures where responses can be biased by social desirability (Breitenstein et al., 2010). Yet despite different approaches to data collection, observation, and measurement, a synthesis of findings converged towards a common set of practices distinguishing higher performing employment specialists.
from lower performers. From these practices a set of specific and targeted behavioural measures can be constructed (Campion et al., 2011). Eight behaviours distinguishing higher from lower performing employment specialists were identified as having the strongest level of evidence (Category 1).

1. Establishes a working alliance with caseload participants.
2. Secures jobs that align with the desires and preferences of individual participants.
3. Develops participants’ job search skills.
4. Provides employment assistance in the workplace and supports employers.
5. Develops a network of employers
6. Builds effective working relationships with a range of stakeholders.
7. Puts in extra effort to understand and address each participants’ employment assistance needs and related barriers.
8. Effectively manages a caseload of participants.

Two other potential candidates with less evidence but good conceptual clarity and potential for influencing program performance (Category 2) were also identified.

10. Monitors and evaluates own performance in the role.

The results from mapping the conceptual space for each of the behavioural indicators are discussed in the next sections, organised by the candidate behavioural measures (1) to (10) identified from the empirical evidence. Each section explores the likely knowledge, skills, attitudes, and other personal attributes that may underlie these behaviours. This analysis is important because it has the potential to assist program leaders to identify the knowledge, skills, attitudes, and characteristics associated with higher performance. Identifying and describing these attributes of higher performers may have important applications in the recruitment, selection, training, and development of all vocational staff.
<table>
<thead>
<tr>
<th>First author, date</th>
<th>Sample size</th>
<th>Country of study</th>
<th>Brief description of study</th>
<th>Measures used</th>
<th>ES attributes influencing program performance</th>
<th>Study limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>King, in press-a</td>
<td>76</td>
<td>Australia</td>
<td>A qualitative study to identify the job development practices differentiating higher performing ES to lower performers.</td>
<td>Behavioural interview</td>
<td>Compared to lower performers, higher performing ES • generated most new jobs by cold calling employers and targeting specific businesses • used advertised vacancies as job leads to approach employers personally • had larger employer networks and made significantly more employer contacts per week • treated employers as customers, spent time with them to learn about their business • created informal ‘meet and greet’ opportunities rather than formal job interviews.</td>
<td>Did not measure the influence of implementation fidelity or ES characteristics on outcomes.</td>
</tr>
<tr>
<td>King, in press-b</td>
<td>76</td>
<td>Australia</td>
<td>A qualitative study to identify the job retention strategies differentiating higher performing ES to lower performers.</td>
<td>Behavioural interview</td>
<td>Compared to lower performers, higher performing ES • maintained frequent contact with employers, supporting and building a relationship with the employer as well as with the participant • made frequent and on-going visits to the workplace • participated in training provided by the employer • provided tailored financial advice and counselling • supported the majority of participants to disclose health issues.</td>
<td>Did not measure the influence of implementation fidelity or ES characteristics on outcomes.</td>
</tr>
<tr>
<td>First author, date</td>
<td>Sample size</td>
<td>Country of study</td>
<td>Brief description of study</td>
<td>Measures used</td>
<td>ES attributes influencing program performance</td>
<td>Study limitations</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------</td>
<td>------------------</td>
<td>---------------------------</td>
<td>--------------</td>
<td>-----------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>King, 2018</td>
<td>76</td>
<td>Australia</td>
<td>ES were categorised into higher and lower performers based on the employment outcomes for participants on their caseload. Behavioural interviews were conducted asking each ES to recall what they remembered doing. ES supervisors were also interviewed. Deductive and inductive thematic analyses were used to synthesise the data.</td>
<td>Behavioural interview</td>
<td>Higher performing ES differed from lower performers in the areas of: • level of job seeker engagement and support • discussions on welfare benefits and disclosure • levels of cold calling to develop jobs. • knowledge of different workplaces and employers' needs • level of job retention activities.</td>
<td>Did not measure the influence of implementation fidelity or ES characteristics on outcomes.</td>
</tr>
<tr>
<td>Corbière, 2017</td>
<td>97</td>
<td>Canada</td>
<td>To identify the relative contribution of ES competencies to participants commencing employment using multi-level analysis and modelling. Controlled for participant characteristics and implementation fidelity.</td>
<td>BAKES scale^4 The Working Alliance Inventory^5</td>
<td>Higher performance was associated with higher scores on the subscales • relationships to employers and supervisors • the ES's perspective of the working alliance.</td>
<td>Self-report scale. Did not control for ES characteristics.</td>
</tr>
<tr>
<td>Corbière, 2014a</td>
<td>153</td>
<td>Canada, Netherlands</td>
<td>A review of the literature to construct an item pool and asked seven ES their opinion to prioritise the items for the scale. Regression analysis examined associations between scores on the BAKES scale and the number of participants commencing employment and the number maintaining employment for six months or more.</td>
<td>The BAKES scale to measure ES self-reported competencies.</td>
<td>Higher performance was associated with higher scores on the subscales • relationship to employers and supervisors • supportive and client-centred approach. No differences found in characteristics of higher and lower performers e.g. age, gender, work history, type of SE program, country, and caseload size.</td>
<td>Self-report scale. Did not control for implementation fidelity.</td>
</tr>
<tr>
<td>First author, date</td>
<td>Sample size</td>
<td>Country of study</td>
<td>Brief description of study</td>
<td>Measures used</td>
<td>ES attributes influencing program performance</td>
<td>Study limitations</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-------------</td>
<td>------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Taylor, 2014, 2010</td>
<td>57</td>
<td>USA</td>
<td>Regression analysis examined the influence of self-report and supervisor rated performance measures on competitive employment outcomes, 90-day job retention, and attrition. The covariates participant and ES demographics, and regional unemployment rate and community size were controlled for.</td>
<td>Clinician Optimism Scale$^6$; ES self-efficacy scale$^7$; IPIP Conscientiousness Scale$^8$; IPS-Q$^9$; Employment Process Measure$^{10}$</td>
<td>Higher performance was associated with • supervisor-rated ES self-efficacy • supervisor-rated ES performance • frequency of contact with participants • greater time in the community • ES with the lowest educational levels • ES with higher caseloads. Optimistic attitudes, knowledge of SE, self-efficacy and conscientiousness were not predictive of higher performance.</td>
<td>Self-report scale. Supervisor’s knowledge of ES outcomes could have influenced their ratings. Did not control for implementation fidelity.</td>
</tr>
<tr>
<td>Glover, 2013, 2011</td>
<td>12</td>
<td>USA</td>
<td>Ethnographic study observing ES working in vocational rehabilitation programs in Chicago. ES were categorised as lower or higher performing based on new job starts, and proportion of participants on their caseload in competitive employment.</td>
<td>Each ES was observed for an average of 12 hours in their vocational role.</td>
<td>Higher performers had • strong listening, interpersonal, and communication skills • good time management skills, and worked efficiently and flexibility • networking and advocacy skills • the ability to work in a team • recovery-orientated attitudes • strong, egalitarian working alliances • collaborative relationships with clinicians, vocational team, and employers.</td>
<td>Researcher who conducted the observations was not blind to ES categorisation.</td>
</tr>
<tr>
<td>First author, date</td>
<td>Sample size¹</td>
<td>Country of study</td>
<td>Brief description of study</td>
<td>Measures used</td>
<td>ES attributes influencing program performance</td>
<td>Study limitations</td>
</tr>
<tr>
<td>--------------------</td>
<td>--------------</td>
<td>------------------</td>
<td>---------------------------</td>
<td>---------------</td>
<td>-----------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Kukla, 2009</td>
<td>n/a</td>
<td>USA</td>
<td>Regression analysis of data from a RCT comparing two vocational rehabilitation programs. The two employment outcomes measured were based on time in a job over the two-year study period. Covariates known to be related to employment outcomes, like work history, were controlled for.</td>
<td>A six-item scale for the working alliance was developed for the study¹².</td>
<td>No relationship was found between the working alliance and employment outcomes. Feedback and instrumental support had a negative association with employment outcomes. The working alliance in both programs was high, but significantly higher in the IPS program.</td>
<td>Working alliance was assessed from participant’s perspective. Did not control for ES characteristics or ES perceptions of participants.</td>
</tr>
<tr>
<td>Catty, 2008</td>
<td>n/a</td>
<td>Six European countries</td>
<td>Data from a six-site RCT identified individual, ES, and program characteristics which would predict employment commencements.</td>
<td>Helping Alliance Scale¹¹</td>
<td>Employment outcomes were associated with a good working alliance between the ES and the participant, as rated by the participant and the ES.</td>
<td>Large number of variables in multiple regression model, some significant associations found could be spurious.</td>
</tr>
</tbody>
</table>

Notes. ES=Employment Specialists. SE=Supported Employment. USA=United States of America. RCT=Randomised Controlled Trial. 1. Sample size= the number of ES in the study. 2. High performance was based on (a) the number of job seekers commencing employment each month averaged over the ES’ length of time with the organisation, and (b) the average proportion of the ES’ current caseload in competitive employment. 3. Job seeker engagement and support was further categorised into (a) developing working alliances, (b) addressing negative experiences, (c) using psychological interventions, and (d) supporting employers. 4. BAKES scale=Behaviours, Attitudes, Knowledge of Employment Specialists scale (Corbière, Neduha, & Lancôt, 2007). A self-report measure consisting of 90 items and 12 subscales. 5. Horvath and Greenberg (1989). 6. Fiorentine and Grusky (1990). Additional items were added to the seven-item scale to increase the scale’s relevance to vocational rehabilitation. 7. A 10-item self-report scale which the authors developed for the study. 8. A 10-item scale created through the International Personality Item Pool (Goldberg et al., 2006). 9. IPS-Q=Individual Placement and Support Quiz, a 30-item multiple-choice quiz (Dreher et al., 2010). 10. A six-item measure of frequency of contacts with employers and consumers, time spent working in the community, numbers of ES caseload in employment for 90 consecutive days, and number on caseload that leave with no vocational outcome. 11. Priebe and Gruyters (1993). 12. The six items were: emotional support, instrumental/informational support, frequency of performance feedback, stressfulness of the relationship, how critical the ES was of the participant, and participants’ overall satisfaction with the relationship.
Establishes a Working Alliance with Caseload Participants

Both participants and vocational staff consider a strong participant-employment specialist relationship important for job success (Areberg et al., 2013; Henry & Lucca, 2004). Studies have examined the way in which employment specialists engage with participants as an influence on employment outcomes (Kukla & Bond, 2009; Whitley et al., 2010). This has been described as a working alliance, or rehabilitation alliance. The working alliance is thought to mirror the properties of the therapeutic alliance in psychiatry, psychology, and other forms of general health treatment (Bordin, 1979; Corbière & Amundson, 2007).

The cluster of skills, knowledge, and personal qualities likely underpinning the establishment of a working alliance are good interpersonal and communication skills. Also considered relevant are good engagement and support skills (Corbière et al., 2014a; King & Waghorn, 2018; Glover & Frounfelker, 2013), and the ability to establish trust (Poremski, Whitley, & Latimer, 2016). Collaborative working relationships are based on participants feeling in control and having an increased sense of power (Areberg et al., 2013). Therefore, employment specialists’ attributes of valuing participant self-determination, choice, and control are likely to be important (Tilson & Simonsen, 2013; Wehman et al., 2003). To establish a working alliance with all participants, employment specialists need the knowledge, skills, and values which enable them to work across cultural boundaries. This is because cultural differences and unintentional biases have been identified in the psychotherapy literature as interfering with the working alliance (Vasquez, 2007).

How employment specialists build relationships with participants was one of three factors identified as differentiating higher from lower performers (Glover & Frounfelker, 2013). Employment specialists were categorised as higher or lower performers based on two measures. 1. The proportion of people in their caseloads each month who were in competitive employment. 2. The proportion of people on their caseload who started a job each month. Higher performers communicated in a transparent and explicit way, expressing hope, enthusiasm, and optimism, and focused on facilitating a
partnership between the two of them. For example, one employment specialist was observed saying, “You know when you do a job search, it is you and me looking” (Glover & Frounfelker, 2013, p. 313).

In a study of 153 employment specialists in Canada and the Netherlands, supportive and individually-focused employment specialists achieved more job starts and longer job tenure for their clients (Corbière et al., 2014a) (see Table 24). The attributes of being supportive and client-centred were measured through eight items in the BAKES scale (Behaviours, Attitudes, Knowledge in Employment Specialists) (Corbière, Neduha, & Lanctôt, 2007). The scale consists of self-ratings from employment specialists on a scale of 7 (always do) to 1 (never do). Questions included asking employment specialists how often they reassured participants when participants were feeling the pressures of starting employment and how often they maintained a supportive relationship at the time when participants commenced employment.

In a more recent study, Corbière et al. (2017) also measured the working alliance between the employment specialist and the participant and found an association between ratings of this alliance and employment outcomes (Corbière et al., 2017). The working alliance was measured using the Working Alliance Inventory (Horvath & Greenberg, 1989) from the perspective of the employment specialist. The working alliance between the employment specialist and the participant has also been measured using the Helping Alliance Scale (HAS) (Priebe & Gruyters, 1993). The relationship between the participant and employment specialist as assessed by the participant and when assessed by the employment specialist were positively associated with participants commencing employment (Catty et al., 2008).

In a qualitative study interviewing 27 participants in an RCT comparing supported employment to usual services, analysis of interview transcripts identified trust as an important influence on the development of a collaborative working alliance. Trust was mentioned more frequently by participants who gained employment (Poremski et al., 2016). Distrust was also spoken about by participants. This included both the participants in the intervention group and the usual services group. Distrust was related to participants experiencing employment specialists failing to listen to their needs or breaking
promises. Trust appeared to develop over time and through repeated contacts with the same employment specialist (Poremski et al., 2016).

Whilst employment specialists’ level of cultural competency has received little investigation in the literature it is likely that cultural competency is an important aspect in the development of a good working alliance (Becker & Drake, 2003, p. 164; Tilson & Simonsen, 2013). In vocational rehabilitation the cultural competencies of vocational program staff have been found to be related to participant employment outcomes, as measured through state vocational rehabilitation case closures (Bellini, 2003). Behaviours demonstrating cultural competencies would include being able to identify the role one’s own race, ethnicity, and experiences have in shaping one’s world view, and being able to interact and communicate with people from different cultures, or ethnic groups (Brach & Frasier, 2000). Employment specialists should be able to give examples of respectful interactions with participants on their caseload. This could include the way in which they work to build an understanding of individual participant’s experiences, values, norms, and traditions, especially where this differs from their own (Tilson & Simonsen, 2013).

Whilst the working alliance that employment specialists establish with participants has been found to be important, there is less consensus in the literature on what constitutes a strong working alliance and a standardised way to measure it (Kukla & Bond, 2009). In the absence of evidence, program leaders could consider using some of the items from existing measures, such as the Working Alliance Inventory or the Helping Alliance Scale. Selected items could be built into participant satisfaction surveys and into discussions in supervision sessions with employment specialists. There is a short-form version of the Working Alliance Inventory which has been found to have similar predictive validity as the long version when tested in counselling relationships (Busseri & Tyler, 2003). Building an understanding of the extent that individual employment specialists are developing working alliances with participants on their caseload could inform the training and supervision of employment specialists.
**Secures Jobs Aligned with the Desires and Preferences of Participants**

Effective job matching is associated with greater job satisfaction (Becker, Drake, Farabaugh, & Bond, 1996; Mueser, Becker, & Wolfe, 2001a) and longer job tenure (Becker et al., 1996; Bond & Kukla, 2011). Employment specialists who have the skills to match individuals to appropriate jobs and help to secure those jobs are likely to get more participants to commence employment, and to help them retain those jobs for longer. It has been argued that job matching is more than just considering occupational preferences. Job matching should include responding to participants’ preferences in terms of location, hours of employment, and wage rate (Bond, Campbell, & Becker, 2013).

Job matching skills are likely to be more important than career counselling skills in the context of vocational rehabilitation for people with severe mental illnesses (Waghorn & Lloyd, 2005). This is because for many program participants the onset of severe mental illness interrupts education and early work experiences. In turn this reduces exposure to the typical range of life and work experiences and roles and responsibilities that help people to develop sufficient career maturity to benefit from career counselling (Waghorn & Lloyd, 2005). Not every participant will have low career maturity, so career counselling can be made available when required and may not need to be a core attribute of employment specialist expertise.

The BAKES scale has three items which ask employment specialists about their job matching practices: “Do you use your knowledge and experience to select jobs that correspond to your clients’ needs?” (p. 1); “Do you assess the work environment to determine whether it corresponds to your clients’ needs?” (p. 6); and “Do you assist your clients in choosing and keeping a job that is right for them?” (Corbière et al., 2007, p. 11). These items were included in the scale because job matching was identified as an important competency for employment specialists through the literature review and feedback from employment specialists (Corbière et al., 2014a). However, studies using the BAKES scale scores on these items have not correlated with employment outcomes (Corbière et al., 2014a, 2017).
It is also important for job retention that employment specialists continue to review the suitability of the job match when a participant is working. This is because if the job changes, or the participant finds the job is no longer aligned with their preferences and skills, employment specialists can support participants to secure a better job match, either with the same employer, or in a different business (Wehman, 2012).

Effective job matching, and job negotiation skills are likely underpinned by knowledge of evidence-based practices particularly

(a) the importance of helping participants to secure employment aligned to their interests and preferences, and

(b) the belief that the nature of the employment assistance provided should follow the preferences of individual participants.

To assess the presence of effective job matching and job securing practices that align with individual preferences, program leaders could regularly ask employment specialists to give examples of how they identify the job preferences of individual participants on their caseload.

The BAKES scale provides a useful starting framework for program leaders. However, the scale could be strengthened further with behavioural questions. Rather than asking for self-reported statements about hypothetical situations or asking for ratings of frequency, program leaders could ask employment specialists to give specific examples of what they have actually done. To check whether these skills are developing, program leaders could monitor the number of participants assisted by each employment specialist who start a suitable new job each month. Program leaders could also train employment specialists in the evidence for the importance of job matching and provide detailed and regular training in how to identify individual preferences and in how to find matching positions. Once the basic training curriculum is designed, program leaders can supplement this with more advanced training informed by the practices of higher performers contrasted with that of lower performers. Some aspects of this training require field testing and supervision. This is particularly important in terms of teaching employment specialists to build relationships with employers. This should be understood in
the classroom first and then practiced in the field with mentoring and supervision to further develop these skills (Swanson et al., 2011).

Develops Participants’ Job Search Skills

There is evidence that employment specialists also need to be able to support and help participants to engage in job search activities. Corbière et al. (2017) examined this aspect in a multi-level analysis of program, individual, and participant characteristics that predict employment outcomes. The authors found that the use of active job search strategies by participants increased the likelihood of them commencing employment by eight times. This Canadian study involved 489 participants with severe mental illnesses and 97 employment specialists. Active job search strategies included talking to friends and past employers about job openings, using social networks to find jobs, sending resumes to employers, completing application forms, and phoning employers (Corbière et al., 2017). This profile of results is consistent with an evaluation of employment programs in Australia. Higher performing sites worked more closely with participants to develop their job search skills, empowered participants to feel confident to find their own jobs and provided guidance to keep the job search focus on realistic jobs (Australian Government, 2001).

Lack of help to look for suitable jobs was the most widely reported resource issue among participants who did not get into employment (Koletsi et al., 2009). The study interviewed 48 participants who did not get a job in a randomised controlled trial of an IPS program. All participants had a diagnosis of a psychotic disorder. Lack of help was an issue reported by both participants in the IPS program and participants in the alternative vocational program (the control group). Participants who did not get a job also reported that the employment specialist did not have a good network of employers to draw upon, or they did not know much about the participant’s field of employment (Koletsi et al., 2009). Similarly, King and Waghorn (in press-a) found higher performers generally had larger employer networks than lower performers.
Behavioural questions in which employment specialists are asked to describe aspects of their practice such as how they actually conducted job search activities with specific participants could help program leaders to understand how well they are working with participants. For example, a general hypothetical question would be “How do you conduct job search with participants?” The behavioural version of this would be “Tell me about the last person you helped with job searching. Tell me what you actually did to help them”. The behavioural question is more specific to an actual participant and requires the employment specialist to give more detail of a specific event or series of events.

Behavioural interviews require that respondents recall specific events where their behaviour demonstrates that they have applied particular skills and knowledge. Respondents are asked to give specific examples of what they have actually done, rather than what they would do (McClelland, 1998; Taylor & Small, 2002). The downside is that behavioural interviews can be time-consuming and more costly than other survey methods. Conducting behavioural interviews by telephone or video could offset these costs (King & Waghorn, 2018).

The behavioural approach to gathering information is important because lower performers are often able to articulate evidence-based practices but have difficulty in performing them (King & Waghorn, 2018; Waghorn et al., 2015). Program leaders can probe specific examples of good practices by asking about how particular participants were assisted. E.g. When did you last hand power and control of job searching to a participant? Tell me more about this? Program leaders could also ask participants and employers about their experiences of working with a particular employment specialist. For example, in a study of a Swedish vocational rehabilitation program a participant commented “I feel there is a shift of power, to my advantage” (Areberg et al., 2013, p. 593).

To improve the employment specialist’s ability to develop participants’ job search skills program leaders could offer training, coaching, and mentoring to employment specialists in job search skills, including approaching employers as outlined previously. Higher performing employment specialists reported enjoying getting out and developing relationships with new employers. Higher performers also took an active interest in the employers’ business, and reported confidence in cold-
calling new employers (King & Waghorn, in press-a). Developing these skills in all employment specialists is important if the goal is improving program performance. Program leaders can prioritise this aspect of training in effective job searching and in approaching employers, and role-model how these skills can be transferred to program participants.

Personal attributes likely to be underpinning the practices of effective job matching, job securing, and enabling participants to job search, are confidence and self-efficacy. It is likely that higher performing employment specialists believe in their ability to effectively support participants into employment and are particularly confident in their ability to interact with employers. Although Taylor and Bond (2014) found employment specialists’ ratings of their own self-efficacy were not correlated with program performance, supervisor ratings of employment specialists were related to program performance. Self-efficacy and sense of personal control have been linked to human performance in a diverse range of contexts (Bandura, 1997). One caveat is that self-efficacy questions need to be both specific and unambiguous to capture any relationship to performance (Sherer et al., 1982). In addition to self-efficacy questions these attributes could be measured through behavioural interviews, or by direct observation of employment specialists in the field (Taylor & Bond, 2014).

It is also likely that employment specialists who can demonstrate confident approaches to employers have extensive experience and knowledge about individual participants and the needs of employers. For instance, an employment specialist’s underlying belief in the employability of a participant, based on a detailed knowledge of the person’s strengths and limitations, could reassure an employer (Glover & Frounfelker, 2013). An understanding that employers have different priorities and preferences when recruiting new workers could encourage an employment specialist to enquire more about the specific needs and preferences of employers. As one employment specialist stated during a research interview “The nice thing about [the agency] is that we have a great product to bring to the door” (Glover & Frounfelker, 2011, p. 204).
Provides Employment Assistance in the Workplace.

For employment specialists to provide effective assistance in the workplace to participants and employers they need job retaining knowledge and skills. These include the skills to: (a) identify work performance issues; (b) assess work performance in the workplace; (c) design training and remediation strategies; and (d) negotiate workplace accommodations with employers (Glover & Frounfelker, 2011; King & Waghorn, in press-b).

Studies have also examined the type of supports which promote job retention (Fossey & Harvey, 2010; King & Waghorn, in press-b; Williams et al., 2016). Wehman et al. (2003) argued that it is important that employment specialists can respond effectively “to both planned and unplanned job retention support needs” (p. 169). Active problem solving when obstacles and issues arise is important (Auerbach & Richardson, 2005; King & Waghorn, in press-b), along with workplace accommodations such as adjusting work hours and schedules and the identification and establishment of natural supports (Secker & Membrey, 2003; Williams et al., 2016) when this is possible. Support from supervisors and co-workers in the workplace has been found to be particularly important for promoting job retention (Corbière et al., 2014b; Huff, Rapp, & Campbell, 2008). Employment specialists with the knowledge and skills to identify and facilitate these supports are likely to help participants retain jobs for longer.

Five of the 10 items on the BAKES subscale, relationships with employers and supervisors, were associated with higher program performance. This subscale asks employment specialists to rate how often they engage in behaviours related to helping participants keep jobs (Corbière et al., 2007). For example, do you: “Work in collaboration with employers to modify job tasks when your clients encounter difficulties” (p. 12). Do you: “Facilitate direct discussions with your clients and their employers about problematic situations that may occur in the workplace” (Corbière et al., 2007, p. 6).

King and Waghorn (in press-b) found that higher performers maintained more contact with employers once the participant was working, and conducted more post-placement visits at the workplace. Higher performing employment specialists were also more likely than lower performers to assist participants to disclosure relevant health information to employers.
Having the skills to support participants to manage disclosure is an important prerequisite to being able to negotiate work accommodations and organise natural supports to sustain employment (Corbière et al., 2014b; Hielscher & Waghorn, 2015). Enhancing employment specialists’ knowledge and skills in relation to helping participants manage disclosure of their personal health and disability related information is therefore also likely to be important for improving program performance (Hielscher & Waghorn, 2015; King & Waghorn, in press-b).

Job retention requires a series of planning and decision-making processes which some participants may need help with (Vornholt et al., 2018). Employment specialists need the skills to: (a) identify work performance issues; (b) assess work performance in the workplace; (c) design training and remediation strategies; (d) help participants develop problem-solving strategies; and (e) discuss and negotiate workplace accommodations with employers (Huff et al., 2008; Wallace & Tauber, 2004; Williams et al., 2016).

Conducting situational assessments to identify any performance issues and assess current performance has been identified as an important aspect of post-employment support. In a study representing the views of 78 vocational staff working in an Australian disability employment program (Waghorn et al., 2015). Although employment specialists could collectively articulate the need for situational assessments, it remains unclear whether many employment specialists use a systematic approach, or even attempt to do this on a regular basis (Waghorn et al., 2015).

In the absence of consensus on what constitutes effective post-employment support practices, program leaders seeking to improve the provision of post-employment assistance, could make use of planning checklists like the one developed by Waghorn et al. (2015, p. 24). Their checklist was based on the views of 78 vocational staff working in Australia on what they considered was important to consider when planning and preparing post-employment assistance. Although there was no empirical evidence linking the checklist to program performance, the checklist provides a useful starting point for program leaders seeking to establish and standardise effective post-employment support practices.
Develops a Network of Employers

Proactively connecting with and building positive working relationships with employers is an essential part of securing employment and helping participants to maintain employment. These activities were linked to higher performing employment specialists and to overall program performance (Corbière et al., 2014a; Glover & Frounfelker, 2011, 2013; King & Waghorn, in press-a; Whitley et al., 2010). Higher performers approach employers differently to lower performers, have larger employer networks, and set up and attend more informal opportunities for participants to meet employers rather than relying on standard interview practices (Glover & Frounfelker, 2011; King & Waghorn, 2018, in press-a). Building relationships with employers is also an important part of preparing the ground for a successful post-employment support plan (Waghorn et al., 2015).

In a study of 78 employment specialists in Australia, King and Waghorn (in press-a) found higher performing employment specialists generated most jobs through cold-calling, targeting specific businesses based on the needs and preferences of participants on their caseload. Lower performers relied on advertised positions, whereas higher performers used advertised vacancies as a mechanism to make new employer approaches (King & Waghorn, in press-a). Underpinning the behaviours of higher performers was the view of the employer as a customer. As a result, higher performing employment specialists spent time getting to know the employers’ business and their recruitment needs, in the same way as they got to know the individual needs of people on their caseloads (King & Waghorn, in press-a).

Higher performers were more likely to disclose information about a participant’s health issues to employers in a thoughtful and constructive way than to avoid any form of mental illness disclosure (King & Waghorn, in press-a). Higher performers were more likely to attempt to educate employers to reduce any mental illness related stigma and discrimination in the workplace. Whereas lower performers preferred to either downplay or withhold information about a participant’s mental health issues (King & Waghorn, in press-a).
Glover and Frounfelker (2011, 2013) found that higher performing employment specialists were better at communicating with employers than lower performers. Interactions by the higher performers were characterised by:

(a) assertive and confident approaches;
(b) providing relevant information;
(c) allowing employers to speak without interrupting them; and
(d) demonstrating active, attentive body language, such as maintaining eye contact.

In contrast, lower performing employment specialists often lacked confidence and would provide too little or too much information to potential employers. The authors also found that although the lower performers might be able to articulate what skills they needed they were unable to perform them (Glover & Frounfelker, 2013). This latter finding emphasises how important it is for program leaders to observe employment specialists in the field. When not in the field, program leaders could ask employment specialists to recount examples of their practices with caseload participants. Program leaders could also set up systems to measure the size of employment specialists’ employer networks, the number of employer contacts per week, and the type of employer contacts. For example, cold-calling, meetings to introduce participants, attending interviews, and providing in-work support.

Builds Effective Working Relationships with a Range of Stakeholders

The communication and interpersonal skills needed to build effective relationships with a range of stakeholders relevant to individual caseload participants is likely to be an important attribute of employment specialist expertise (Corbière et al., 2007, 2014a). Employment specialists who collaborated well and who provided relevant information to the employment team and clinicians achieved better average monthly competitive employment rates than those who did not (Glover & Frounfelker, 2013). The higher performing employment specialists used communication strategies to provide and obtain relevant information.
Stakeholders can include a participant’s family, friends, and significant others who provide social support. Families who understand evidence-based practices can, through their advocacy, help to sustain vocational rehabilitation programs (Biegel, Swanson, & Kola, 2007; McFarlane et al., 2000; Swarbrick et al., 2017). Employment specialists who communicate effectively with families and friends about the value of employment, and the importance of evidence-based practices, are likely to mobilise the support of a participant’s family and friends. Family and friends then in turn provide additional assistance to individual participants to secure employment (Swarbrick et al., 2017).

Program leaders could in supervision and in performance appraisals ask employment specialists to describe how they involve clinicians and family members in the employment assistance process. Program leaders could also ask employment specialists for specific examples when they have tried unsuccessfully to involve other stakeholders. This could help program leaders to assess the communication skills of employment specialists and to improve these skills, to help the employment specialists develop better working alliances with relevant local stakeholders. Program leaders and technical assistance providers could also role-model communication skills and explore different communication strategies to assist employment specialists to learn and develop better skills through observation.

**Puts in Extra Effort to Address Participants’ Assistance Needs and Barriers**

A behavioural indicator that distinguishes higher performing employment specialists from lower performers is the additional and specific efforts they undertake to problem-solve and address participants’ employment-related barriers (King & Waghorn, 2018; McGurk & Mueser, 2006). This extra effort would appear to reflect a high degree of commitment to program goals and to each participant, but there is little empirical evidence from vocational rehabilitation to help identify the knowledge, attitudes, and other attributes underlying this level of commitment (King & Waghorn, in press-b).
One possible influence is that higher performers have positive beliefs based on real experiences about the employability of people with mental illnesses. Holding a positive belief about the recovery of people with serious mental illnesses has been identified as an important personal quality of people working in mental health services (Coursey et al., 2000; Farkas, Gagne, Anthony, & Chamberlin, 2005; Gowdy, Carlson, & Rapp, 2003). In interviews with higher performers, Glover and Frounfelker (2013) found that these employment specialists talked about the rights of people with severe mental illnesses to be employed, the links between employment and recovery, and the skills and abilities of people on their caseload. If this positive attitudinal bias is linked to performance, it is a promising quality for program leaders to identify and value in the recruitment and selection of staff, as well as through ongoing training and development.

Tilson and Simonsen (2013) examined the personal characteristics of 17 top performing vocational staff working with young people with disabilities. These higher performers expressed positive beliefs about the young person’s ability to be employed. One staff member was recorded saying: “I believe in each person’s potential” (p. 130). The authors considered that it was the underlying commitment to, and belief in participants, that were strong motivating factors for these vocational staff to work extra hours, work in a flexible way, and go that extra mile for the participants on their caseload.

Knaeps et al. (2015) examined the differences in the beliefs of three groups of vocational workers. The authors developed a self-report questionnaire to measure perceived barriers, sense of control, and self-efficacy around returning to employment. The development of the questionnaire was underpinned by the theory of planned behaviour (TPB) in which beliefs are hypothesised to shape individual self-efficacy, subjective norms, and perceived control to influence behaviours (Ajzen, 1991). The authors found that specialist vocational staff working with people with disabilities perceived fewer barriers to employment, had a greater sense of control over the process of employment assistance, and had higher levels of self-efficacy. Specialist vocational staff were compared to the public employment service case managers and to the gatekeepers of the employment service. The study did not measure
any association between the personal qualities of the vocational counsellors, such as levels of motivation, or their behaviours and practices, or the employment outcomes of the participants on their caseload. This remains a promising direction for further research.

It is also likely that employment specialists observed putting in extra effort have good knowledge of the individual assistance needs and barriers to employment, of caseload participants. A study investigating this knowledge found vocational staff generally did not have a good understanding of participants’ assistance needs (Rampton, Waghorn, Souza, & Lloyd, 2010). Knowledge was measured through the Work-Related Self-Efficacy scale (Waghorn, Chant, & King, 2005a). Both vocational staff and participants completed the scale. Ratings were provided on participants’ efficacy performing 37 core work-related tasks, such as attending appointments on time, checking instructions with the supervisor, asking an employer for information about a job. A low correlation was found between the employment specialist-rated and the participant-rated work-related self-efficacy scores ($r=0.10$) (Rampton et al., 2010). Although there does not appear to be any evidence yet that such knowledge improves outcomes, the study authors argued that increasing employment specialist understanding of individual participants’ employment assistance needs could enhance the strength of the working alliance, and in turn enhance program performance.

Vocational staff who put in extra effort in addition to understanding participants’ assistance needs and related barriers, are also likely to acquire an understanding of how to overcome these barriers. This understanding could be developed through knowledge of current systems and contracts and how these can help or hinder the vocational outcomes of caseload participants (Corbière et al., 2014b). In a study examining the core competencies of transition specialists coordinating services for young people with disabilities, Defur and Taymans (1995) found that knowledge of “agencies and system change” (p. 5) was rated the highest of all the 12 competency domains they examined. Similarly, Corbière et al. (2014b) argued that workplace and policy knowledge is particularly important for employment specialists. Although these authors (Corbière et al., 2014b) did not explicitly relate this to how this might influence an employment specialist’s behaviour, they did incorporate two related
items into the BAKES scale (Corbière et al., 2007). These scale items are: “Do you know the disability compensation programs available in your province (e.g. work integration contract, opportunities fund)?” (p. 15); and “Do you have a good understanding of the organisational structure of the mental health system and the related services?” (p. 16). Yet these were not correlated with employment outcomes (Corbière et al., 2014a, 2017). The BAKES scale also includes items covering knowledge of clinical symptoms, health or vocational services, knowledge of the workplace, and knowledge of disability laws and policies (Corbière et al., 2007). Although these items appear promising, they have not yet been found associated with employment outcomes.

Another likely important attribute underpinning committed employment specialists is the extent that employment specialists are motivated by externally imposed goals and targets and are not demotivated (Australian Government, 2001; King & Waghorn, 2018; Latham, 2015, p. 161). Vocational staff motivated by targets and goals typically feel a sense of personal responsibility for participant outcomes (King & Waghorn, 2018; Tilson & Simonsen, 2013). To assess the extent that employment specialists have greater commitment that drives extra effort, program leaders could ask employment specialists to give examples of how they learn about participants’ assistance needs and assistance priorities, and how they have previously addressed specific barriers including participants’ fears. From the responses, program leaders will get information revealing the level of effort employment specialists apply, and also their knowledge of local contracts and systems. This information could then inform strategies to improve the targeting of employment specialists’ effort.

Clear expectations for the level of commitment and effort valued can also be outlined in job descriptions as well as in performance appraisal systems. Some program leaders may try to improve commitment related behaviours through financial rewards and incentives (Latham, 2015, p. 171).

In summary, the likely attributes underpinning employment specialists who are observed going the extra mile for participants are as follows:

(a) higher expectations of success;

(b) self-efficacy for attaining the desired outcomes;
(c) perceived control over the vocational process of helping participants find and keep suitable employment;
(d) knowledge of the individual assistance needs of people on their caseload;
(e) knowledge of local and national systems and how to access support for caseload participants;
(f) the belief in the employability of caseload participants; and
(g) motivated by goals and targets, or financial rewards (see Table 26).

**Effectively Manages a Caseload of Participants**

The ability to manage and prioritise a caseload are also important skills for employment specialists (Torrey et al., 1998). Some studies have examined the way in which higher performing employment specialist’s structure and manage their time, their ability to work flexibility, to prioritise and manage unexpected demands as they occur (Australian Government, 2001; Glover & Frounfelker, 2013). Higher performers developed mechanisms for managing their caseloads and their time and were found to be more prepared and better at planning than lower performers (Glover & Frounfelker, 2013). Higher performers were more organised and innovative and used their networks and communication skills to obtain resources for participants, and used them effectively (Australian Government, 2001).

To assess this attribute, program leaders could ask employment specialists about the systems they have in place to manage their caseload. Program leaders could also check that employment specialists are managing to spend time with all participants on their caseload in relation to their different assistance needs. Where employment assistance is integrated with clinical treatment services, program leaders could utilise electronic shared record systems to understand employment specialists contacts with individual participant’s, clinicians, and family members. The range of responses to these questions and the information gathered could assist program leaders to identify whether more structured strategies are needed. These strategies could be developed in partnership with the employment specialist to ensure that no person on their caseload is disadvantaged by poor planning,
poor communication, or by poor time management. Careful monitoring of caseload management skills could help identify specific problems, such as an employment specialist who does not value the input and involvement of particular stakeholders, which could hinder the development of this attribute.

**Other Candidate Attributes for Targeted Measures**

Other potential influences on employment specialists' performance were identified in the literature reviewed for the conceptual analysis. However, the evidence was limited. This means there could be a range of other behaviours and practices of interest to program leaders that require more empirical research to establish causal links to program outcomes. This review identified two other promising behaviours (a) applies relevant evidence-based practices, and (b) engages in self-reflection and self-evaluation of role performance.

**Applies relevant evidence-based practices**

Knowledge of evidence-based practices in vocational rehabilitation has been explored through the construction of a scale designed to measure this attribute. The IPS-Q is a 30-item multiple choice quiz (Dreher et al., 2010). However, in a study of 182 employment specialists, responses to the IPS-Q were not associated with employment commencements or with job tenure (Taylor & Bond, 2014).

One possible explanation for this unexpected finding could be that applied knowledge is more important than theoretical knowledge (Taylor & Bond, 2014). For example, it is likely that knowledge of evidence-based practices and how to apply these to each participant in the current context positively influences program performance. King and Waghorn (2018) found that both lower performers and higher performers were able to describe good practices, but only the higher performers could give behavioural examples of how these practices were applied with respect to specific participants. This finding suggests that higher performers both understand evidence-based practices and can confidently execute these as part of their daily routine. This is a promising area for further research because little is known about why some practitioners can describe evidence-based practices yet fail to implement these with respect to their current caseload.
Engages in self-reflection and self-evaluation of role performance

Knowledge of evaluation methods and how to use them to improve one’s own performance was identified as one of 12 core competencies for staff working with people with severe mental illnesses (Coursey et al., 2000). Self-evaluation involves being able to systematically assess one’s own work performance and knowing how to obtain and use feedback from a range of sources. Sources could include supervisors, colleagues, peers, and family members. Self-evaluation could also involve identifying and using valid satisfaction surveys, and other evaluation measures to understand and monitor one’s performance (Coursey et al., 2000).

The capacity for self-reflection, self-monitoring, and self-performance evaluation is critical to self-reflective practice (Atkins & Murphy, 1993; Bannigan & Moores, 2009). Some authors believe that in the rehabilitation sciences this competency is more critical to success than any other practitioner attribute (Donaghy & Morss, 2000; Cropley, Hanton, Miles, & Niven, 2010). The reason given is that without this attribute practitioners are unlikely to be able to continuously learn and develop their practice in response to emerging evidence of new and better principles and practices. Looking for and incorporating in one’s practices alternative sources of knowledge, beyond one’s existing knowledge, is important in reflective practice (Bannigan & Moores, 2009). So is seeking expert guidance and support and building in opportunities for others to scrutinise one’s ideas and practices (Donaghy & Morss, 2000). Although there is limited supporting evidence, it seems reasonable to expect that knowing how to monitor and evaluate one’s performance and to continue learning, i.e. keeping oneself up to date, could be more important than an employment specialist’s level of education or qualifications attained. Program leaders could specify in employment specialist job descriptions that employment specialists should take time for self-reflective practice and participation in professional development groups.

Discussion

The unsolved problem driving the investigation in this chapter is how to improve the performance of vocational rehabilitation programs. This is particularly relevant once a sufficient
fidelity threshold has been reached in the application of evidence-based practices, yet performance remains suboptimal and below expectations.

This chapter proposes that one way forward is to go beyond measuring program implementation fidelity and measure and develop employment specialist expertise conceptualised as the knowledge, skills, attitudes, and other personal attributes which distinguish higher performing vocational staff from lower performers.

A large volume of research has investigated the competencies and qualities of staff in vocational rehabilitation programs, yet agreement has not been reached on the set of most likely attributes that are associated with higher performance. The conceptual analysis in this chapter has brought together the main empirical findings which appear useful for specifying a range of promising behavioural indicators for employment specialist expertise. It is hypothesised that these behavioural indicators can adequately represent the cluster of knowledge, skills, attitudes, and other characteristics likely to represent employment specialist expertise in vocational rehabilitation for people with severe mental illnesses. It is anticipated that the suggested item pool constructed from this analysis could assist program leaders to identify, describe, assess, and develop the expertise of vocational staff (see Table 26).

The conceptual analysis has covered three areas.

1. A description and definition of the construct of employment specialist expertise in the context of vocational rehabilitation for people with severe mental illnesses.
2. The identification of the measurable behavioural attributes that distinguish higher performing employment specialists from lower performers.
3. Suggestions of how program leaders can use this information to identify, describe, assess, and improve the expertise of employment specialists in vocational rehabilitation programs.

The review of the empirical evidence and the wider conceptual literature identified and prioritised ten behavioural indicators to describe and define employment specialist expertise. These behavioural indicators were developed from the convergence of the empirical evidence in vocational
rehabilitation. The eight with the strongest evidence represent the behaviours most likely to distinguish higher performing employment specialists from lower performers.

The process of mapping the conceptual space for the construct of employment specialist expertise drew on the wider conceptual literature and empirical evidence. This identified the likely knowledge, skills, attitudes underpinning the eight targeted behavioural indicators. There are likely to be unquantified overlaps among these attributes because they were found to converge to underpin employment specialist expertise (see Figure 12).

The included studies indicate that higher performing employment specialists build excellent working relationships with participants, as well as with employers, clinicians, family members, and other key stakeholders. They are technically proficient in all stages of providing employment assistance. They are particularly good at searching for, developing, and securing jobs, and supporting job retention by providing support in the workplace and by supporting employers. They also excel at engaging participants in active job search activities and strategies, and have well developed employer networks. Higher performing employment specialists seem to utilise vocational rehabilitation practices, local labour markets, and health and welfare systems to the benefit of each caseload participant. Higher performers take initiative, are well-organised, and are good time managers. Higher performers demonstrate belief in the employability of all the participants they support, and have confidence in their own technical ability, as well as their ability to influence employment outcomes. Higher performers are likely to value participant choice, control, and self-determination.

By logically and systematically unpacking the behavioural indicators and the likely underpinning attributes of higher performers it was possible to suggest how program leaders could identify, describe, assess, and improve these attributes in employment specialists. This can inform staff recruitment and training, and on-going workforce development. It is anticipated that developing staff in these directions will improve program performance.
Measurement

In developing an item pool for measuring the core construct employment specialist expertise, the strengths and limitations of previous attempts to measure this construct were considered. Much previous research in vocational rehabilitation has used self-report instruments to measure employment specialist expertise, when observational measures or behavioural interviews are potentially more promising methods, especially for predicting job performance (King & Waghorn, 2018; Taylor & Bond, 2014; Taylor & Small, 2002).

Table 26 summarises the findings from the conceptual analysis. The first column lists each of the identified eight behavioural targets (behavioural indicators). The second column outlines the possible underpinning attributes for each behavioural target that emerged from the wider review of the conceptual and empirical literature. The third column provides an item pool with suggested measures for each behavioural target. The fourth and fifth columns suggest how to collect the measures including sample questions to be directed to employment specialists and participants. The last (sixth) column, summarises the sources of evidence related to each candidate attribute.

Limitations

Employment specialist expertise is a complex and multi-dimensional construct. Clear conceptualisation is an important prerequisite for the construction of valid and reliable measures (Proctor et al., 2013. This investigation took a pragmatic approach using the available empirical evidence to help identify the behaviours that distinguish high performing employment specialists from lower performers, but it was more an exploratory analysis than a fully systematic review. Therefore, it is possible that some literature was missed which could have identified additional behaviours. However, the evidence collected and reviewed can be considered an important starting point to operationalise this construct.

This conceptualisation did not identify motivation as a distinct dimension of employment specialist expertise but rather considered it as an important attribute underpinning goal orientation. This
is because those not motivated by goals and targets are not likely to be suited to this role. Motivation as it relates to the observable behaviour of participant responsiveness is discussed further in Chapter 8.

Chapter Conclusions

This chapter presented the conceptual analysis of the seventh dimension of implementation quality in the IFVR, employment specialist expertise. This is a complex and multi-dimensional concept where a clear and precise conceptualisation needs to inform the development of a valid and reliable measure. It is anticipated that this partial operationalisation of employment specialist expertise could inform the further development and testing of a scale to identify, define, and measure the construct, employment specialist expertise. Once fully operationalised this measure could assist program leaders to support employment specialists identify their current approaches to vocational rehabilitation and adopt new practices to improve the performance of vocational rehabilitation programs.

The next chapter, Chapter 8, partially operationalises the eighth, and final, dimension of implementation quality in the IFVR, namely participant responsiveness.
Figure 12. Candidate knowledge, skills, and attitudes underlying the behavioural indicators of employment specialist expertise

Notes. The overlapping ellipses denote the clusters of different skills, attitudes and applied knowledge likely underpinning the behavioural indicators. 1. Employment specialist expertise is theorised to be made up of eight behavioural indicators. 1. Establishes a working alliance with caseload participants. 2. Secures jobs that align with the desires, and preferences of individual participants. 3. Develops participants’ job search skills. 4. Provides employment assistance to participants and employers in the workplace. 5. Develops a network of employers. 6. Builds effective working relationships with a range of stakeholders. 7. Puts in extra effort to address participants’ employment assistance needs and related barriers. 8. Effectively manage a caseload of participants.
Table 26. Operationalising the construct of employment specialist expertise

<table>
<thead>
<tr>
<th>ES behavioural indicator</th>
<th>Underpinning ES attributes</th>
<th>Measurement of behavioural indicator</th>
<th>Method of measurement</th>
<th>Examples of questions</th>
<th>Evidence for the influence of the attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td>Establishes a working alliance with caseload participants</td>
<td>Interpersonal, communication and engagement skills, cultural competency, valuing self-determination, choice and control.</td>
<td>Program leaders engage in the following activities. 1. Observe vocational staff interacting with caseload participants. 2. Ask vocational staff to describe examples of successful and less successful working relationships with caseload participants. 3. Ask vocational staff to give examples of how they get to know individual caseload participants e.g. personal values, traditions. 4. Conduct participant satisfaction surveys which include questions about the quality of the working relationship between the ES and participant. (Score 3 if three or more are present; 2 if two are present, 1 if one is present).</td>
<td>Interview Training and supervision records</td>
<td>Examine training records and description of training content. Ask the supervisor and the ES how they would know that they are developing good working alliances with participants. Ask about the diversity of caseloads and how well vocational staff build working relationships with participants from other cultures to their own.</td>
<td>Areberg et al., 2013; Becker &amp; Drake, 2003; Bellini, 2003; Bordin, 1979; Catty et al., 2008; Corbière &amp; Amundson, 2007; Corbière et al., 2007, 2014a, 2017; Glover &amp; Frounfelder, 2013; Fiency &amp; Lucca, 2004; King &amp; Waghorn, 2018; Kukla &amp; Bond, 2009; Poremski et al., 2016; Tilson &amp; Simonsen, 2013; Vasquez, 2007; Wehman et al., 2003; Whitley et al., 2010.</td>
</tr>
<tr>
<td>Secures jobs that align with the desires and preferences of individual participants</td>
<td>Job matching and securing skills, knowledge of evidence-based practices, of caseload participants preferences, and employer engagement techniques.</td>
<td>ES training includes the following. 1. Information on the association between job matching and employment outcomes. 2. How to identify individual preferences. 3. Field-mentoring on employer approaches. (Score 1 for each anchor present).</td>
<td>Interview Training logs</td>
<td>Ask ES about the job preferences of caseload participants. Find out how they discover participants’ job preferences and whether they do this with everyone. Review induction and training schedules.</td>
<td>Becker et al., 1996; Bond et al., 2013; Bond &amp; Kukla, 2011; Corbière et al., 2007, 2014a; Glover &amp; Frounfelder, 2011, 2013; King &amp; Waghorn, in press-a; Mueser et al., 2001a; Wehman, 2012;</td>
</tr>
<tr>
<td>ES behavioural indicator</td>
<td>Underpinning ES attributes</td>
<td>Measurement of behavioural indicator</td>
<td>Method of measurement</td>
<td>Examples of questions</td>
<td>Evidence for the influence of the attribute</td>
</tr>
<tr>
<td>--------------------------</td>
<td>----------------------------</td>
<td>---------------------------------------</td>
<td>-----------------------</td>
<td>---------------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Develops participants' job search skills</td>
<td>Engagement skills, confidence, self-efficacy, perceived control.</td>
<td>1. The ES can explain how they discover the job preferences of individual caseload participants. 2. At least one participant on the ES’ caseload is assisted to start a suitable new job each month. (Score 1 for each anchor present).</td>
<td>ES outcomes</td>
<td>Review ES outcomes.</td>
<td>Waghorn &amp; Lloyd, 2005.</td>
</tr>
<tr>
<td>Provides assistance in the workplace.</td>
<td>Job sustaining skills, knowledge of workplace accommodations, and evidence-based practices.</td>
<td>The ES carries out the following practices. 1. Field mentoring to teach job search skills to participants. 2. Helps caseload participants focus their job search activities on realistic jobs. 3. Helps caseload participants to use their family and social networks to find job openings. 4. Role-models proactive and confident employer approaches. 5. Regularly encourages caseload participants that they can get and keep employment. (Score 3 if four or more are present; 2 if three are present, 1 if one or two are present).</td>
<td>Observation Interview Program records</td>
<td>Ask ES to describe how they help participants look for jobs, to give examples of assisting caseload participants with job search activities. Ask participants how they work with the ES to look for jobs.</td>
<td>Australian Government, 2001; Corbière et al., 2014a, 2017; Glover &amp; Frounfelker, 2011; King &amp; Waghorn, 2018, in press-a; Taylor &amp; Bond, 2014; Koletski et al., 2009; Waghorn et al., 2005a; Whitley et al., 2010.</td>
</tr>
<tr>
<td>ES behavioural indicator</td>
<td>Underpinning ES attributes</td>
<td>Measurement of behavioural indicator</td>
<td>Method of measurement</td>
<td>Examples of questions</td>
<td>Evidence for the influence of the attribute</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-----------------------------</td>
<td>-------------------------------------</td>
<td>-----------------------</td>
<td>----------------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Develops a network of employers</td>
<td>Interpersonal and communication skills, knowledge of individuals, belief in employer as customer.</td>
<td>The ES carries out the following practices. 1. Cold-calls new employers on a weekly basis. 2. Has an established employer network. 3. Gets to know employers’ business needs. 4. Each week has set up at least one meeting between a caseload participant and an employer. 5. Makes frequent and on-going visits to work sites of employed caseload participants. (Score 3 if four or five are present; 2 if three are present, 1 if one or two are present).</td>
<td>Interview Program records Observation</td>
<td>work. Also ask to see in-work support plans.</td>
<td>Hierscher &amp; Waghorn, 2015; Hiff et al., 2008; Secker &amp; Membrey, 2003; Vornholt et al., 2018; Waghorn et al., 2015; Wallace &amp; Tauber, 2004; Wehman et al., 2003; Williams et al., 2016.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ES behavioural indicator</th>
<th>Underpinning ES attributes</th>
<th>Measurement of behavioural indicator</th>
<th>Method of measurement</th>
<th>Examples of questions</th>
<th>Evidence for the influence of the attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td>Builds effective working relationships with a range of stakeholders</td>
<td>Interpersonal and communication skills, knowledge of individual participants’ family and social networks.</td>
<td>The ES provide examples of positive working relationships with stakeholders. 1. Clinicians of caseload participants. 2. Family members of caseload participants. (Score 1 for each anchor present).</td>
<td>Interviews Training and supervision records</td>
<td>Ask program leaders and ES how they would know they have good working relationships with clinicians and family members. Questions might be included in supervision, or in regular meetings between the supervision and clinicians, and in stakeholder satisfaction surveys.</td>
<td>Biegel et al., 2007; Corbière et al., 2007; 2014a; Glover &amp; Froumfelder, 2013; McFarlane et al., 2000; Swarbrick et al., 2017.</td>
</tr>
<tr>
<td>Puts in extra effort to understand and address each participants’ employment assistance needs and barriers</td>
<td>Problem-solving skills, knowledge of local and national systems and individual participants assistance needs and barriers. Self-efficacy and perceived sense of control over outcomes. Believes in the employability of people on their caseload.</td>
<td>The ES carries out the following practices. 1. Demonstrates they know the assistance needs and assistance priorities of caseload participants. 2. Can provide examples of how they have used the support from local health and welfare services to help the participants on their caseload. 3. Can provide examples of how they approach conflicts and barriers in relation to helping participants get and keep a job. 4. Can provide examples of helping participants to address employment-related fears. 5. Talks about the employability of caseload participants. (Score 3 if four or more are present; 2 if three are present, 1 if one or two are present).</td>
<td>Interview Participant records/vocational assessments</td>
<td>Ask the ES to talk about caseload participants. Ask how the ES gets to know and understand their needs and priorities; what specific support is available locally from the health and welfare agencies; the types of fears and concerns that participants have about employment. Ask ES if they have anyone on their caseload who is not employable, and why.</td>
<td>Australian Government, 2001; Corbière et al., 2007, 2014a, b; Coursey et al., 2000; Defur &amp; Taymans, 1995; Farkas et al., 2005; Glover &amp; Froumfelder, 2013; Gowdy et al., 2003; Guydish et al., 2014; Knaeps et al., 2015; King &amp; Waghorn, 2018; Latham, 2015; McGurk &amp; Mueser, 2006; Rampton et al., 2010; Tillson &amp; Simonsen, 2013; Waghorn et al., 2005a.</td>
</tr>
<tr>
<td>ES behavioural indicator¹</td>
<td>Underpinning ES attributes</td>
<td>Measurement of behavioural indicator</td>
<td>Method of measurement</td>
<td>Examples of questions</td>
<td>Evidence for the influence of the attribute²</td>
</tr>
<tr>
<td>---------------------------</td>
<td>-----------------------------</td>
<td>-------------------------------------</td>
<td>-----------------------</td>
<td>----------------------</td>
<td>-------------------------------------------</td>
</tr>
<tr>
<td>Effectively manages a caseload of participants</td>
<td>Time management and organisational skills.</td>
<td>1. The ES has mechanisms in place to help them manage their caseload of participants. 2. The ES sees all participants when required. 3. The ES stays in touch with all the other stakeholders in priority order. (Score 1 for each item that is present).</td>
<td>Interview Review program records</td>
<td>Ask ES to talk about the mechanisms they have in place to plan and prioritise the participants on their caseloads. This could include diary prompts and weekly reviews of their caseloads. Ask the ES how they manage their weekly time, across participants, and other stakeholders and how they make decisions about who to contact and make time for.</td>
<td>Australian Government, 2001; Glover &amp; Frounfelker, 2011; Torrey et al., 1998.</td>
</tr>
</tbody>
</table>

Notes. ES=Employment Specialist. 1. The eight behavioural indicators were identified from a targeted review of the empirical literature which identified the attributes of higher performing vocational staff compared to lower performers. 2. Green=Category A: The evidence supports the inclusion of the attribute and its effect on program performance. Red=Category B: The evidence does not support the inclusion of the attribute and its effect on program performance. Blue=Category C: No empirical evidence was identified, but authors have written about the potential for the attribute to influence program performance. 3. This item is looking for ES actively working with concerns rather than deferring these to other providers e.g. clinical team.
Chapter 8: Operationalising Participant Responsiveness

Introduction

Chapter 4 developed the IFVR for improving the performance of evidence-based practices in vocational rehabilitation for people with severe mental illnesses. The IFVR hypothesises that there are eight dimensions of implementation quality in addition to implementation fidelity. These eight dimensions are considered candidates for further improving vocational rehabilitation program performance. This is particularly relevant once a sufficient fidelity threshold has been reached in the application of evidence-based practices, yet performance remains suboptimal and below expectations. To apply and test the framework the dimensions need further conceptualising and operationalising. Chapters 6 and 7 described the methods and findings from the conceptual analysis of seven of the additional dimensions of implementation quality.

This chapter (Chapter 8) describes the methods and findings from the conceptual analysis of the last candidate dimension of implementation quality, namely participant responsiveness. This construct is a dimension of implementation quality that focuses on the behaviours of individual participants in vocational rehabilitation programs. These behaviours are directly observable and can potentially be objectively described and defined, once measures are constructed through a logical and systematic process of conceptual analysis. Therefore the conceptual analysis of this construct draws on theories of motivation and behaviour considered relevant by the existing vocational rehabilitation literature and the broader literature concerning the science of psychosocial program implementation.

Participant responsiveness, like employment specialist expertise operationalised in the previous chapter (Chapter 7), is a complex and multi-dimensional phenomenon. To support a logical and systematic unpacking of this concept, and measurable attributes making up the construct, the conceptual analysis of participant responsiveness was guided by four questions.

1. What is participant responsiveness in the context of vocational rehabilitation for people with severe mental illnesses?
2. How is the construct participant responsiveness related to participant motivation?
3. What are the conditions or influences which could improve participant responsiveness in the context of vocational rehabilitation for people with severe mental illnesses?

4. What measures are already available to identify and measure relevant conditions and influences on participant responsiveness in vocational rehabilitation programs?

The conceptual analysis aims to develop an operational specification for participant responsiveness through mapping each potential attribute considered essential to adequately represent the construct. Once further developed and tested, the operationalisation of this dimension of implementation quality is intended to lead to the development of practical tools to improve program performance. The use of such tools can then explore and quantify the influences on participant responsiveness, with the ultimate goal being to optimise participant responsiveness in vocational rehabilitation programs.

Chapter overview

This chapter starts by briefly describing the analytical methods. Then the first question (What is participant responsiveness in the context of vocational rehabilitation?), is explored both conceptually and in terms of any established empirical associations between measures of participant responsiveness and program performance. The conceptual analysis begins by first considering the hypothetical question of what would be observable if substantial levels of participant responsiveness were present or absent in vocational rehabilitation programs.

Previous research in vocational rehabilitation has largely used program attrition as a proxy measure of participant responsiveness. The limitations of that approach to operationalisation are discussed next. A rationale for exploring previous approaches to conceptualisation and measurement is then provided, in terms of achieving conceptual clarity. This is important because a lack of conceptual clarity appears to have derailed previous attempts to define and measure participant responsiveness. In particular, because previous conceptualisations have not delineated the construct, from potentially confounding concepts of motivation for employment.
Next, in line with the second question guiding the analysis, a broad construct specification for participant responsiveness in the context of vocational rehabilitation is proposed. The specification recognises a theoretical link between participants' responsiveness and participants’ motivation for employment. By recognising this link, the discussion then describes how previous research attempted to elucidate the multi-dimensional phenomenon of employment motivation. Different conceptualisations of employment motivation have led to different measures being developed to examine participant responsiveness in vocational rehabilitation programs. Clear conceptualisation of these two related constructs does not appear to have been achieved previously. Yet, this is a prerequisite to exploring the third question in this conceptual analysis namely: What are the conditions and influences which could improve participant responsiveness in the context of vocational rehabilitation programs? In this part of the conceptual analysis five candidate sources for potentially measurable attributes for participant responsiveness are examined.

1. Participants’ general interest in employment as a current personal goal or a personal goal for the near future.

2. Participants' perceptions of the vocational rehabilitation being offered and the match to their own perceptions about their employment goals, employment capabilities, and perceived barriers to employment.

3. Other individual participant characteristics such as age, gender, work history, educational attainment, communication skills, and the nature of health conditions and impairments.

4. Individual circumstances that can facilitate or hinder participants’ availability for employment, and their availability for participation in the program.

5. Characteristics of the vocational rehabilitation program that can influence participant engagement.

To address the fourth question, the chapter maps the conceptual space for each of the five potential sources of measurable attributes. This mapping includes descriptions and specifications of any potentially applicable measures, and the strengths and limitations of previous research relevant to this
Suggestions for new, or alternative measures for candidate attributes of participant responsiveness are also outlined.

The chapter concludes with a discussion on the practical implications arising from the conceptual analysis. A set of targeted statements which could help to further operationalise the construct of participant responsiveness are developed. This partial operationalisation is expected to assist leaders of vocational rehabilitation programs to better identify, measure, and potentially improve or optimise participant responsiveness. The expectation is that delineating and measuring the attributes that best represent the influences on participant responsiveness could potentially improve performance of vocational rehabilitation programs, particularly for people with severe mental illnesses.

**Methods**

The conceptual analysis of the construct of participant responsiveness followed the same seven-step process in the conceptual analysis for the other dimensions of implementation quality, as outlined in Chapter 6.

1. Development of the construct specification.
2. Identification of measurable attributes of the construct.
3. Identification of existing relevant measures.
4. Refinement of the construct specification and measurable attributes.
5. Generation of targeted statements for each attribute.
6. Rationalisation of statements to increase parsimony, remove redundancy, and to improve construct representativeness.
7. Development of a measurement item pool.

To identify the relevant literature the following initial search terms were used: ‘responsiveness, engagement, retention, motivation’. These were combined with the terms 'program, treatment, vocational rehabilitation’, and ‘supported employment’. These terms were then combined with the terms 'engagement, responsiveness, outcomes, performance’, and ‘employment'. As articles were sourced and the operational specification developed the literature search continued in a second iteration.
using terms relevant to newly identified measurable attributes. The terms used included: ‘motivation, motivation for employment, self-efficacy, individual characteristics, program characteristics, health status, program relevance, flexibility’, and ‘cultural congruence’. These terms were then combined with the previously searched terms, i.e. ‘engagement, outcomes, performance’.

Google Scholar was the primary database used to search for literature supplemented with searches of Medline, PubMed, and PsycINFO. As relevant articles were identified and reviewed their reference lists were searched to identify further potentially relevant literature. Searches were then extended to find articles citing the relevant articles that had already been identified. As with the conceptual analysis for the other seven dimensions of implementation quality, this iterative process prioritised literature from vocational rehabilitation first, and next considered the literature from the wider science of program implementation.

**Developing the Construct Specification**

The development of the operational construct specification for participant responsiveness was an iterative process informed by an examination of three specific areas of the literature.

1. The conceptualisation of participant responsiveness developed from the broader science of psychosocial program implementation and how this conceptualisation of participant responsiveness relates to program performance.

2. The relationship between one commonly used measure of participant responsiveness, namely program attrition and program performance.

3. The relationship between participant responsiveness and a widely used access criterion, that participants consistently say that they want to be employed. This criterion has been widely used in controlled trials examining the efficacy and effectiveness of IPS programs.

**Participant responsiveness and program performance**

Increasing the level of participant responsiveness among those participants who appear less than ideally engaged in the program is expected to positively influence program performance. Higher levels
of participant responsiveness are likely to lead to participants staying in a program longer, and participating more actively and effectively (Berkel et al., 2011; Dane & Schneider, 1998; Durlak & DuPre, 2008). Active and effective participation are likely to reduce attrition and increase the prospects for the desired program goal being attained (Southwick, 2014). For these reasons, the level of participant responsiveness is a potential source of previously unexplained variance in the performance of vocational rehabilitation programs. Participant responsiveness has been identified as an important candidate dimension of implementation quality (Berkel et al., 2011; Durlak & DuPre, 2008; Lockett et al., 2018a).

Indicators of higher levels of participant responsiveness could include signs of participant engagement, enthusiasm, satisfaction, and expressed interest in the program as demonstrated by active participation both in and between meetings, greater contributions in sessions, turning up on time, and regular attendance (Corbière et al., 2017; Dane & Schneider, 1998; Domitrovich & Greenberg, 2000). Another potential marker of responsiveness is the language participants use. Engaged participants used the word “we” more often than the phrase “them and us” (Bright, Kayes, Worrall, & McPherson et al., 2014, p. 650). At a program level, a lower level of participant responsiveness is expected to be marked by poor attendance, failure to complete assigned tasks, and leaving a program prior to gaining an outcome (program attrition) (Baydar, Reid, & Webster-Stratton, 2003).

Program attrition and program performance

Program attrition has been examined in studies of vocational rehabilitation program implementation (Southwick, 2014). The number of participants who drop out of the program before securing an employment outcome is one possible program measure of levels of participant responsiveness. In a three-arm controlled trial, Mueser et al. (2004) compared IPS, psychosocial rehabilitation, and usual care. The extent of program attrition differed significantly across each study arm with higher attrition in the lower performing sites. Mueser et al. (2004) did not control for attrition in the study arguing that attrition is closely linked to program type and to program effectiveness. Whilst it is likely that attrition is linked to program characteristics, it is also possible that attrition involves an
effect independent of program type. If attrition is an independent effect, the differences in attrition at a site level observed by Mueser et al. (2004) may represent an interaction between program type and participant responsiveness. One way to explore these issues further is to conceptualise attrition as an independent effect, then explore the extent it is influenced by participant responsiveness, program type, and program effectiveness. Developing a measure of participant responsiveness would provide the means to explore the drivers of attrition that can be so influential on outcomes in vocational rehabilitation (Drebing et al., 2012).

Trials examining vocational rehabilitation interventions often report research participation attrition and sometimes report vocational program attrition, yet reasons for attrition are rarely reported. The differences between those who leave a program early and those who do not, are not usually examined in any detail. An exception was a European multi-site RCT comparing IPS to several different forms of vocational rehabilitation. The number of participants dropping out and the reasons for dropping out were collected at each of the six study sites (Burns et al., 2006). The number of drop-outs in the control services were significantly higher than in the IPS sites, across all sites, except Rimini in Italy. In Rimini, the control group had fewer drop-outs than all the other control groups, and attrition did not differ significantly in the control group from the IPS intervention.

Further data analysis of individual-level predictors of withdrawing early from the program revealed that two influences were significant. 1. Participant’s satisfaction with their living situation. 2. The number of non-vocational needs met including information, transport, access to a telephone, and welfare benefits. Contrary to expectations, increased satisfaction with participant’s own living arrangements increased the chances of the participant withdrawing from the vocational program. Similarly, where a non-vocational need like transport was met, this doubled the likelihood of dropping out of the program (Burns et al., 2006, p. 281). Non-vocational needs were measured through the Camberwell Assessment of Needs scale (Phelan et al., 1995). Satisfaction with living situation was measured through the Lancashire Quality of Life Profile (Oliver, Huxley, Priebe, & Kaiser, 1997). The authors could only speculate that positive changes in personal circumstances reduced the motivation of
participants for competitive employment, and for remaining in the program. However, the Burns et al. (2006) study was not designed to explore such counter-intuitive relationships.

This finding highlights the potential importance of attrition in vocational programs, which can have a greater effect size than the intervention itself. An understanding of the influences on attrition, and how it can best be managed in vocational rehabilitation, is urgently needed before the relative effectiveness of different vocational rehabilitation programs can be appreciated. This is particularly important when unequal attrition is present in an RCT. If attrition is a significant and non-random effect, then it is possible that uneven attrition in a controlled trial could be fatal to any conclusions the authors wish to form about relative program effectiveness (Amico, 2009).

An Australian study provides a further illustration of the need for a better understanding of the links between participant responsiveness, program attrition, and program performance. Parletta, Waghorn, and Dias (2017) compared the employment outcomes of participants with obligations to look for and secure employment (non-volunteers), with participants with no obligations (volunteers). The two groups were participating in the same IPS program. Employment commencements were higher for the group with obligations (non-volunteers) and this was considered at least partly due to attrition being lower for this group. Fourteen of the 31 (56.0%) volunteers left the program early with no outcome, compared to 11 of 37 (44.0%) of the non-volunteers.

**Entry criteria applied by IPS trials**

The majority of RCTs examining the efficacy of IPS in the USA have required candidate participants to attend one or two research meetings prior to being accepted into the research trial. Attendance at these meetings is often used as a marker of participants’ motivation to participate in vocational rehabilitation (Macias et al., 2001). An important exception was the RCT conducted by Lehman et al. (2002). All participants who were eligible were included in the study, they did not have to volunteer to participate in the program. In this study the IPS intervention outperformed the usual vocational service, as measured by the proportion of all eligible participants commencing competitive employment (41.6% in the IPS intervention compared to 11.3% in the control, $p<.001$). However,
attrition from the vocational services was unequal across the RCT. Of the IPS group, 92.9% (105/113) stayed in the program and received vocational services, compared to only 33.0% (35/106) of those enrolled in the control group vocational service. There was no difference between the groups in attrition from the clinical services (Lehman et al., 2002). High and uneven attrition across the two groups limits the findings, because the outcome difference may represent a confounding of program type and program attrition effects.

A study by Macias et al. (2001) provides a parallel example to the Lehman et al. (2002) RCT. An expressed desire for employment was not a requirement of participating in the vocational rehabilitation program. Whilst desire increased the chances of starting a job, 14 of 49 (28.6%) participants who had no initial interest in employment commenced employment. These findings highlight the need for clear conceptual clarity and to consider participant responsiveness in the context of a broader range of influences and program characteristics.

The Need for Conceptual Clarity

Participant responsiveness and participant motivation

The empirical findings outlined above indicate that clear conceptualisation of participant responsiveness is particularly important. There is also an added complication to clear conceptualisation because of a potential link between participant responsiveness and participant motivation.

Motivation describes an internal mental state that is not directly observable and has been surrounded by conceptual confusion (Drieschner, Lammers, & van der Staak, 2004). This conceptual confusion may have contaminated the utility and validity studies of instruments designed to measure participant responsiveness and related concepts such as treatment engagement, as well as participant motivation (Drieschner et al., 2004).

Researchers in implementation science have conceptualised participant responsiveness as “levels of participation and enthusiasm” (Dane & Schneider 1998, p. 45), and the degree to which the program “stimulates the interest or holds the attention of the participants” (Durlak & DuPre, 2008, p. 250).
Berkel et al. (2011) distinguished between the behaviours of participants (participant responsiveness), and the behaviours of program facilitators. They interpreted participant responsiveness as “the participants’ level of enthusiasm for and participation in an intervention” (p. 23). Research examining participant engagement in vocational rehabilitation programs has interpreted the construct of participant engagement as simply composed of three factors. 1. Program attendance. 2. Expected contribution to the sessions. 3. Homework completed outside of sessions (Southwick, 2014).

These conceptualisations have guided researchers to operationalise participant responsiveness by the number of meetings attended, how active participants are in treatment sessions, participant satisfaction levels, and the amount of completed homework (Berkel et al., 2011; Southwick, 2014). This is problematic because this conceptualisation of participant responsiveness and participant engagement appears to fall into what has been described as “conceptual quicksand”, where the concept is defined by the very behaviours it is theorised to predict (Bandura, 1997; Drieschner et al., 2004, p. 1119).

This type of conceptual confusion typically arises from a failure to specify a logical and causal map of the entity being examined and the means for its measurement (Kanfer et al., 2008, p. 21). In this case there is a failure to differentiate the internal attitudinal states of motivation from (a) the proxy indicators of motivation, (b) the engagement process itself, and (c) the various activities and behaviours that signify higher engagement. Researchers have previously failed to account for the various internal and external factors which are hypothesised to influence participants’ motivation to engage. To remedy this, Drieschner et al. (2004) proposed a logical framework for conceptualising treatment motivation. To achieve this, the observable behaviours of treatment engagement were differentiated from the following:

(a) the motivation to engage in the treatment (an unobserved psychological state theorised to cause the behaviour);

(b) other internal states that influence a person’s motivation to engage; and
(c) the circumstances and individual characteristics that influence these internal states (Drieschner et al., 2004).

This conceptualisation also recognised that there can be circumstances when someone has substantial levels of motivation to engage, but because of other individual limitations or circumstances, such as cognitive or neuropsychological limitations, or a participants’ non-availability, this may not always result in high levels of engagement (Drieschner et al., 2004).

Re-conceptualising participant responsiveness in vocational rehabilitation programs

The Drieschner et al. (2004) conceptual framework appears relevant to the practical problem of defining participant responsiveness in vocational rehabilitation for people with severe mental illnesses. The framework suggests how to re-conceptualise and re-operationalise participant responsiveness in the context of vocational rehabilitation programs.

Participant responsiveness can be conceptualised as encompassing relevant program characteristics, individual characteristics, individual circumstances, and participants’ perceptions of themselves and their perceptions of the vocational rehabilitation program (see Figure 13). These candidates all have the potential to influence participants’ employment motivation, and in turn influence relevant behaviours, namely their engagement with and participation in the program, once participants commence program activity.

Participant responsiveness in the context of vocational rehabilitation therefore appears to be a complex phenomenon where responsiveness is theoretically influenced by participants’ underlying motivation for employment. Motivation for employment is another complex phenomenon influenced by participant perceptions, participants’ general interest in employment, relevant program characteristics, participants’ circumstances, and participant characteristics, such as age, gender, ethnicity, communication skills, and educational attainment (see Figure 13).
Figure 13. Re-conceptualising the construct participant responsiveness

Notes. 1. The dotted circle denotes the eight dimensions of implementation quality in the Implementation Framework for Vocational Rehabilitation (IFVR), in addition to participant responsiveness, which are also hypothesised to have a direct relationship on improving program performance (Lockett, Waghorn, & Kydd, 2018a). 1. Fidelity to the original program. 2. Program intensity. 3. Program delivery. 4. Complementary programs. 5. Removal of non-evidence-based practices. 6. Technical assistance. 7. Program evaluation. 8. Employment specialist expertise.

This re-conceptualisation recognises that participant responsiveness may be linked with the participants’ motivational state and with employment motivation in particular. This recognition introduces the diverse and inconclusive literature on theories of motivation and motivation for employment (Kanfer et al., 2008, p. 7). Some of these theories are yet to be tested in the context of vocational rehabilitation. It remains unclear whether they have informed, confounded, or possibly derailed the science and practice of vocational rehabilitation (Kanfer et al., 2008, p. 578; Lam, Wiley, Siu, & Emmett, 2010).

However, the main purpose of this conceptual analysis is to develop measures based on directly observed behaviours which enable vocational rehabilitation leaders to differentiate low levels of participant responsiveness from adequate to high levels. The expectation is that reducing the incidence of low levels of participant responsiveness could improve program performance by both increasing program outcomes and by reducing program attrition.

Related forms of this re-conceptualisation of participant responsiveness have been proposed by other researchers. In a review of the literature on engagement in healthcare, Bright et al. (2014)
identified 31 studies investigating the concept of engagement in adult healthcare, or studies that measured engagement in healthcare. The authors concluded that participant engagement is commonly viewed and measured solely through participants' behaviours ignoring other potentially important influences. The authors argued that engagement is a more complex concept with the nature of the activity and the people delivering the activity playing an important role. As a result of the review, the authors conceptualised participant responsiveness as a dynamic process rather than a static behaviour (Bright et al., 2014). Participant responsiveness was thought to also include the way the delivery system interacts with and influences participants (Hermann & Provost, 2003).

The wider conceptualisation of participant responsiveness proposed in this conceptual analysis has important implications for identifying ways to measure this construct. This in turn, has implications for improving program performance. A wider conceptualisation can potentially capture how program leaders can facilitate participant responsiveness rather than ignoring it as an intrinsic participant-owned problem. This is important because participant characteristics conceived as fixed traits can also be determined to be outside the program’s responsibility or influence (Bright et al., 2014; Drieschner et al., 2004). A wider conceptualisation also implies that measuring program attrition, whilst important, is not a representative or sensitive enough measure on its own to understand and enhance individual-levels of participant responsiveness. Attrition, as measured as part of routine program evaluation, is a useful but limited proxy indicator for monitoring the results of efforts to improve participant responsiveness.

**Participants’ Employment Motivation**

It is often assumed that levels of participant responsiveness in a vocational program are the consequence of participants’ underlying motivation for employment (Lam et al., 2010; Sasaki, Sato, Yamaguchi, Shimodaira, & Kawakami, 2018). For this reason, further exploration of the concept of employment motivation and previous attempts to describe, define, and measure this form of motivation, is warranted in this conceptual analysis.
There is conceptual ambiguity surrounding motivation for employment in the vocational rehabilitation literature. This is because researchers have used the term to infer two different entities. One is a dynamic and fluctuating psychological state. The other is a fixed trait-like individual characteristic. Constructs of employment motivation are then defined differently without researchers being explicit about the assumptions concerning the nature of the underlying motivation entity. This ambiguity is then often compounded by developing measures of motivation that lack construct validity because the underlying motivational entity was not described or defined (Kanfer et al., 2008, p. 21).

In this conceptual analysis of participant responsiveness in the IFVR, employment motivation is conceptualised as primarily a dynamic psychological state which is shaped by five different sources of influence (see Figure 13).

1. Participants’ general interest in employment as a current personal goal or a personal goal for the near future.
2. Participants' perceptions of the vocational rehabilitation being offered and the match to their own perceptions about their employment goals, employment capabilities, and perceived barriers to employment.
3. Other individual participant characteristics such as age, gender, work history, educational attainment, communication skills, and the nature of health conditions and impairments.
4. Individual circumstances that can facilitate or hinder participants’ availability for employment and their availability for participation in the program.
5. Characteristics of the vocational rehabilitation program that can influence participant engagement.

Individual perceptions are expected to mediate the influence of program characteristics, situational circumstances, and individual characteristics on participants' motivation for employment.

This conceptualisation assumes that employment motivation is a complex internal state resulting from the combined influences of internal cognitive and emotional processes with external factors. Employment motivation cannot be observed directly by rehabilitation professionals but can be
inferred from individual behaviours (Drieschner et al., 2004; Kanfer et al., 2008, p. 4; Villotti, Corbière, Zaniboni, Lecomte, & Fraccaroli., 2015).

In the vocational rehabilitation literature, employment motivation has been previously conceptualised as both a state and a trait. This duality is briefly discussed, to explain why an exploration of employment motivation and its relationship to participant responsiveness is needed to advance the science and practice of vocational rehabilitation.

**Motivation as a psychological state**

As an internal state, employment motivation is considered necessary to explain the forces that cause a participant to act, meet a goal, or satisfy a particular need (Villotti et al., 2015). It has been described as a psychological state that precedes, and influences goal-directed behaviour (Saperstein, Fiszdon, & Bell, 2011). As a fluctuating internal state rather than a fixed personality trait, motivation is expected to fluctuate in response to external influences and to a person’s health and social circumstances (Drieschner et al., 2004).

Observable employment-directed behaviours are likely to include, but not be limited to, active participation in a vocational rehabilitation program. Motivation for employment for people with severe mental illnesses has been considered to result from both a conscious and an unconscious assessment of the advantages and disadvantages of employment (Honey, 2004). If the advantages outweigh the disadvantages, then employment-directed behaviours are thought more likely to appear. This idea that participants continually weigh up the pros and cons of employment implies a process that fluctuates over time, differs by participants, and recognises “the waxing and waning of the psychological state of motivation to engage” (Drieschner et al., 2004, p. 1122). This approach also reflects the possibility that among people with disabilities and severe mental illnesses in particular, motivation for employment is constantly changing and is not a fixed attribute of a person (Kanfer et al., 2008, p. 6; Kinn, Holgersen, Aas, & Davidson, 2014). As an individual’s underlying employment motivation fluctuates, it is anticipated that so will their participation in and responsiveness to the vocational rehabilitation program.
Conceptualised in this way, employment motivation can be enhanced by both external and internal forces (Helfrich, Kielhofner, & Mattingly, 1994). Ajzen (1991) in a theory of planned behaviour contended that to understand human behaviour it is important to identify the causes of that behaviour. Drieschner (2005) adds to this argument highlighting the need to identify influences on motivation, not just to understand behaviours, but to learn how to further enhance the target motivation.

**Motivation as a fixed individual trait**

In vocational rehabilitation, a participant’s level of motivation for employment has also been conceptualised as a fixed trait. The presence or absence of this trait can then be used as the main criterion to assess whether or not employment assistance should be provided (Sasaki, et al., 2018). This has implications for national disability and employment programs because the size of the target group varies considerably depending on how sufficient motivation is defined and where the threshold is set for eligibility assessments of individuals in scope for the program. Describing, defining, and measuring employment motivation could therefore be a critical ingredient in public employment services to both estimate the size of the target group and to help match the most intensive services to those most in need of those services.

In a study of the job search behaviours of people with severe mental illnesses, Alverson, Carpenter, and Drake (2006) found that even though each participant expressed the desire for employment when commencing the program, 10 of 25 participants (40.0%) did not undertake any active job search activities. Desire for employment was assumed to be a fixed, trait-like characteristic, and whilst this was not explicitly acknowledged, was also used as a proxy indicator for employment motivation (Alverson et al., 2006). Since desire for employment did not result in the goal-directed behaviours of looking for employment in 40.0% of cases, this study supports a construct of employment motivation that is more dynamic than the expression of a fixed trait consisting of the desire to get a job.
Measuring employment motivation

Table 27 provides an overview of research studies attempting to examine employment motivation and employment outcomes. Various proxy indicators for employment motivation are shown such as work values, perceived barriers, employment-related self-esteem, self-efficacy, and general interest in work. These previous studies were informed by different underlying theories of motivation and therefore have different conceptualisations of employment motivation. Both state-based constructs and trait-based assumptions are presented. Researchers taking a trait-based approach are frequently informed by the popular stages of change theory of motivation (Prochaska & DiClemente, 1982), and the stage of ‘readiness for change’, which has led to the concept of ‘readiness for employment’ (Lam et al., 2010). These different conceptualisations have generated an array of measures targeting employment motivation. These measures include: the Motivation to Find a Job Scale (Villotti et al., 2015); the Work Values Questionnaire (Zaniboni, Corbière, Fraccaroli, & Perron, 2010); the Barriers to Employment and Coping Efficacy Scale (Corbière, Mercier, & Lesage, 2004); and the Work-Related Self-Efficacy scale (Waghorn et al., 2005a).

Whilst each study has its strengths and limitations, the use of different underlying theories, conceptualisations, and measures of employment motivation has inevitably led to mixed and inconsistent findings. Consequently, no clear understanding has emerged of what constitutes or influences employment motivation. In some cases, the binary assumption is made that motivation is either completely present or completely absent. This assumption can sometimes lead to the exclusion of people from employment assistance. Whilst caution is needed with the application in vocational rehabilitation of any motivational theory which rigidly classifies participants, useful practice developments have emerged from trait-like conceptualisations of employment motivation. For example, the practice of motivational interviewing, which was developed from the stages of change theory (Miller & Rollnick, 2014), is considered a useful approach to supporting participants with vocational decision-making (Lloyd, Tse, Waghorn, & Hennessy, 2008).
The discussion that follows adopts a state-based approach to describing and defining employment motivation. Using this re-conceptualisation of participant responsiveness, the following sections bring together what is known about the likely conditions and influences on participant responsiveness. These conditions and influences are considered measurable attributes of the construct participant responsiveness in the context of vocational rehabilitation programs for people with severe mental illnesses. Each condition or influence is discussed in turn, along with examples of how other researchers have attempted to measure these attributes.
<table>
<thead>
<tr>
<th>First author, date</th>
<th>Country</th>
<th>Overview of study</th>
<th>Concept of employment motivation</th>
<th>Measurement of employment motivation</th>
<th>Summary of study findings</th>
<th>Study strengths</th>
<th>Study limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Westcott, 2017</td>
<td>Australia</td>
<td>Surveyed 254 participants with schizophrenia and schizoaffective disorder living in the community. Investigated the influence of different individual characteristics on interest in employment and employment status.</td>
<td>State-like</td>
<td>A single question: “Are you interested in employment as a goal in the future?” Answer coded yes or no.</td>
<td>Individual characteristics, age, employment history, and severity of current hallucinations affect employment interest. However, age and employment history have less influence on interest in employment than they have on employment status. No other clinical factors (severity, course pattern of illness, or functioning) had any impact on employment interest.</td>
<td>Distinguished between interest in employment and employment status.</td>
<td>Single question to measure motivation, unclear whether this is sufficient.</td>
</tr>
<tr>
<td>Hielscher, 2017</td>
<td>Australia</td>
<td>Investigated employment fears and employment values using two scales. 1. Employment Fears Scale. 2 Employment Values Scale. Study sample was 25 participants with severe mental illnesses.</td>
<td>State-like</td>
<td>Employment Fears Scale (9-items) and Employment Values Scale (10-items).</td>
<td>Employment fears are closely related to self-stigma. Employment values are independent of past experiences, self-stigma, work history, and employment fears. Employment values appear more trait-like, employment fears more state-like. Both fears and values appear needed to understand current engagement and motivation for employment.</td>
<td>Both scales have promising psychometric properties.</td>
<td>Small sample size. Larger, longitudinal studies are needed to validate scales and understand predictive validity.</td>
</tr>
<tr>
<td>Parletta, 2017</td>
<td>Australia</td>
<td>Examined the influence of welfare obligations on employment outcomes in 68 participants with mental illnesses in a VR program.</td>
<td>State-like</td>
<td>Asked participants to rate their motivation on a scale of 1 to 10.</td>
<td>External pressure (welfare obligations) led to increased employment outcomes through reducing program attrition (Non-volunteers had lower attrition).</td>
<td>The effects of welfare obligations on program</td>
<td>Used a convenient non-randomised sample, also</td>
</tr>
<tr>
<td>First author, date</td>
<td>Country</td>
<td>Overview of study</td>
<td>Concept of employment motivation</td>
<td>Measurement of employment motivation</td>
<td>Summary of study findings</td>
<td>Study strengths</td>
<td>Study limitations</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------</td>
<td>-------------------</td>
<td>----------------------------------</td>
<td>--------------------------------------</td>
<td>--------------------------</td>
<td>----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Villotti, 2015</td>
<td>Canada</td>
<td>Study of 366 participants with mental illnesses participating in different VR programs.</td>
<td>State-like</td>
<td>MTFJ (7-items)</td>
<td>Non-volunteers were not less motivated than volunteers.</td>
<td>Outcomes were identified.</td>
<td>Sample size was small.</td>
</tr>
<tr>
<td>Corbière, 2011</td>
<td>Canada</td>
<td>Path analysis to investigate the influence of individual characteristics and individual perceptions on job search behaviours. Sample of 281 participants with mental illnesses in VR programs.</td>
<td>State-like</td>
<td>MTFJ (14 items), BECES, Attitudes towards employment (10-item Likert scale)</td>
<td>Social encouragement, participants' perceptions of their job search efficacy, attitudes to work, and perceived barriers to employment influenced intention to engage in job search and actual job search behaviours.</td>
<td>Large sample size.</td>
<td>Did not control for fidelity, or other qualities of VR program.</td>
</tr>
<tr>
<td>Zaniboni, 2010</td>
<td>Canada</td>
<td>Validation of the Work Values Questionnaire for use with people with mental illnesses Classified 254 participants on the basis of their work values and compared people on work outcomes and other personal characteristics.</td>
<td>Trait-like</td>
<td>Work Values Questionnaire (58-item and shortened 30-items)</td>
<td>Higher work values were associated with commencing employment, as well as higher levels of employment-related self-efficacy, and greater coping skills to overcome barriers to employment. No relationship was found between work values and gender, age, education level, diagnosis, and work history. Work values are likely to be more trait-like.</td>
<td>Developed a scale to attempt to measure work values. Large sample size</td>
<td>Early stages of psychometric testing of the new measure.</td>
</tr>
<tr>
<td>First author, date</td>
<td>Country</td>
<td>Overview of study</td>
<td>Concept of employment motivation</td>
<td>Measurement of employment motivation</td>
<td>Summary of study findings</td>
<td>Study strengths</td>
<td>Study limitations</td>
</tr>
<tr>
<td>-------------------</td>
<td>------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>----------------------------------</td>
<td>-------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>----------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Corbière, 2009</td>
<td>Canada</td>
<td>Examination of participants’ self-esteem when they commenced employment. Sample was 47 participants. Study also sought to validate the self-esteem as worker scale.</td>
<td>State-like</td>
<td>Self-esteem as a worker scale&lt;sup&gt;2&lt;/sup&gt;</td>
<td>Individual’s self-esteem as worker increased when in employment, and this was independent of primary diagnosis. Social self-esteem did not change. Self-esteem as worker was also correlated with satisfaction with the job and with the supervisor.</td>
<td>Developed a domain specific measure of self-esteem.</td>
<td>Short follow-up period, small sample size.</td>
</tr>
<tr>
<td>Catty, 2008</td>
<td>Six European countries</td>
<td>Examination of the association between self-esteem and employment outcomes using data from a six-site RCT of IPS and TVR.</td>
<td>Trait-like</td>
<td>Rosenberg self-esteem scale</td>
<td>Self-esteem did not predict employment outcomes.</td>
<td>General measure of self-esteem used.</td>
<td></td>
</tr>
<tr>
<td>Waghorn, 2005a, 2007</td>
<td>Australia</td>
<td>Community sample of 104 people with schizophrenia and schizoaffective disorder. This was a 12-month longitudinal study examining work-related self-efficacy and work-related subjective experiences.</td>
<td>State-like</td>
<td>Work-related self-efficacy scale. Work-related subjective experiences scale</td>
<td>Higher scores on work-related self-efficacy scale were associated with current employment status. Work-related self-efficacy varied across time and were not predictive of future employment status. It is possible that work-related self-efficacy is more important as a result of being in employment, rather than as a predictor of employment. Subjective experiences were relatively independent of other variables.</td>
<td>Recognised that self-efficacy needs measuring specifically at the task level.</td>
<td>Early stages of psychometric testing of the new measures.</td>
</tr>
<tr>
<td>First author, Country, Date</td>
<td>Overview of study</td>
<td>Concept of employment motivation</td>
<td>Measurement of employment motivation</td>
<td>Summary of study findings</td>
<td>Study strengths</td>
<td>Study limitations</td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------------------</td>
<td>---------------------------------</td>
<td>------------------------------------</td>
<td>---------------------------</td>
<td>----------------</td>
<td>------------------</td>
<td></td>
</tr>
<tr>
<td>Alverson, USA 2006</td>
<td>Ethnographic study observing job search behaviours in 25 people with mental illnesses. Considered that motivation is a behavioural concept and that true motivation is discovered through behaviours, not through self-report. To discover the employment, education, and training goals of people using mental health services. Sample size was 149 participants.</td>
<td>Trait-like</td>
<td>Self-reported expressed desire for employment</td>
<td>Even though group all said they had a desire to find employment, 60% were active job seekers and 40% were passive or resistant. Active job seekers were motivated to work to earn money, and through social support from families. 141 (90%) of participants had an interest in work now or in the future. 132 were interested in paid employment. Of those with an interest, 116 (82%) wanted assistance to get work.</td>
<td>Desire for employment does not necessarily predict program participation.</td>
<td>Small sample size.</td>
<td></td>
</tr>
<tr>
<td>Seebohm, England 2005</td>
<td>To discover the to discover the employment, education, and training goals of people using mental health services. Sample size was 149 participants.</td>
<td>State-like</td>
<td>Self report interest in any work. Answer coded yes, maybe, or no.</td>
<td>Subjective experiences were correlated with current employment status but not with having a current employment goal or with work history. Having a current vocational goal was correlated with work history.</td>
<td>Interviews conducted by peer researchers.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Waghorn, Australia 2005b, 2007</td>
<td>Community sample of 111 people with schizophrenia and schizoaffective disorder to investigate work-related subjective experiences, employment goals, and current employment status.</td>
<td>State-like</td>
<td>Work-related subjective experiences scale</td>
<td>The two scales showed promising psychometric properties of validity and reliability for people with mental illnesses. Both scales were correlated with the Rosenberg self-esteem scale. Levels of coping efficacy were related to length of absence from the workplace.</td>
<td>Developed and tested a scale for measuring participants subjective experiences.</td>
<td>Early stages of psychometric testing of the new measures</td>
<td></td>
</tr>
<tr>
<td>Corbière, Canada 2004</td>
<td>The study examined the psychometric properties of two scales in a sample of 254 participants with mental illnesses in VR programs. 1 The Barriers to Employment, Coping Efficacy scale (BECES). 2. The Career Search Efficacy scale (CSES).</td>
<td>State-like</td>
<td>BECES CSES</td>
<td>The two scales showed promising psychometric properties of validity and reliability for people with mental illnesses. Both scales were correlated with the Rosenberg self-esteem scale. Levels of coping efficacy were related to length of absence from the workplace.</td>
<td>Developed and tested measures for the influences on employment motivation.</td>
<td>The study did not investigate the predictive validity of the two scales.</td>
<td></td>
</tr>
<tr>
<td>First author, date</td>
<td>Country</td>
<td>Overview of study</td>
<td>Concept of employment motivation</td>
<td>Measurement of employment motivation</td>
<td>Summary of study findings</td>
<td>Study strengths</td>
<td>Study limitations</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------</td>
<td>-------------------</td>
<td>----------------------------------</td>
<td>--------------------------------------</td>
<td>---------------------------</td>
<td>----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Mueser, 2001b</td>
<td>USA</td>
<td>528 participants with mental illnesses were followed for two years to identify predictors of competitive employment.</td>
<td>Trait-like</td>
<td>The authors do not report on method of measurement, just described as ‘the expressed desire to work’.</td>
<td>Participants who expressed a desire for employment were more likely to be employed in the future. People who dropped out did not differ on desire for employment.</td>
<td>Desire for employment does not necessarily predict participant behaviours.</td>
<td>No information provided on the measure of desire to work.</td>
</tr>
<tr>
<td>Macias 2001</td>
<td>USA</td>
<td>166 participants with mental illnesses, 86 were enrolled in PACT vocational rehabilitation program and 80 in a Clubhouse vocational rehabilitation program. At entry 70% of participants were interested in getting a job, 30% were not (most gave a definite answer of no, rather than stating they were uncertain).</td>
<td>Trait-like</td>
<td>Single question: Are you currently interested in working? Answer coded: Yes or no. Interviewer clarified that it was paid employment.</td>
<td>With the exception of ethnicity, the participant groups interested in employment, and those who were not, did not differ on individual characteristics. Interest effected engagement in VR but even those not initially interested, 55 engaged with VR and 16 commenced employment. Stated interest in work and receipt of VR predicted whether a person was employed and how long it took to get a job.</td>
<td>Desire for employment does not necessarily predict participant behaviours.</td>
<td>Transitional employment jobs were counted as employment outcomes.</td>
</tr>
</tbody>
</table>

Notes. VR=Vocational Rehabilitation. CE=Competitive Employment. RCT=Randomised Controlled Trial. IPS=Individual Placement and Support. TVR=Traditional Vocational Rehabilitation. 1. Participants were asked: “What was your motivation to work at the commencement of the IPS service” and “What was your motivation to work at the completion of the IPS trial” on a scale of 1 to 10 (Parletta, Waghorn, & Dias, 2017, p. 118). 2. MTFJ=Motivation to Find a Job Scale. A 7-item self-report measure. The scale asks questions about a participant’s self-efficacy to overcome barriers to employment, the participant’s intentions, and the participant’s attitudes, and values towards work. Participants are asked to rate on a seven-point Likert scale from 1 (completely disagree) to 7 (completely agree), the statements: “Right now, getting a job is one of my main objectives”; “I really feel motivated to find a job”; “I am willing to put in the necessary efforts in order to get a job”; and “I would be very disappointed if I were not able to get a job in weeks to come” (Villotti et al., 2015, p. 61). 3. The authors increased the specificity of the 10-item Rosenberg self-esteem scale to the rehabilitation context.
The Influences on Participant Responsiveness

This re-conceptualisation of participant responsiveness, and of participant employment motivation, suggests there are some conditions and influences which could, if modified, improve participant responsiveness. The conceptual space for each of the five measurable attributes thought to make up the multi-dimensional construct of participant responsiveness are discussed in turn.

1. Participants’ general interest in employment as a current personal goal or a personal goal for the near future.
2. Participants’ perceptions of the vocational rehabilitation being offered and the match to their own perceptions about their employment goals, employment capabilities, and perceived barriers to employment.
3. Other individual participant characteristics such as age, gender, work history, educational attainment, communication skills, and the nature of health conditions and impairments.
4. Individual circumstances that can facilitate or hinder participants’ availability for employment, and their availability for participation in the program.
5. Characteristics of the vocational rehabilitation program that can influence participant engagement.

Participants’ general interest in employment

Participants’ general interest in employment is likely to be the first sign of employment motivation and may be more enduring than an internal state of employment motivation (Westcott, Waghorn, McLean, Statham, & Mowry, 2017). However, there is limited empirical research examining how interest in employment develops into employment motivation and leads to seeking employment assistance and participating in vocational rehabilitation programs (Westcott et al., 2015). Asking participants about their interest in employment as a current personal recovery goal has the advantage of being less potentially stigmatising with implications of less social desirability bias, than asking a person directly about their current motivation for employment (Fisher, 1993).
Having an interest in employment improves the likelihood that a participant will in future commence competitive employment (Macias et al., 2001; Mueser, Salyers, & Mueser, 2001b). Even when participants have no expressed desire for employment, they can be supported to gain employment (Lehman et al., 2002; Macias et al., 2001). This is likely because interest is based on knowledge and previous experiences (Bettman & Park, 1980). No expressed interest may reflect little knowledge of what it takes to be employed or no relevant experiences of employment. The utility of asking participants about their interest in employment lies in exploring why they are not interested. This could be due to a range of factors including no employment experience, lack of availability due to other commitments, perceived lack of skills, or concerns about the implications of their health condition on employment.

The way interest in employment is measured can also impact on the responses obtained. Using a simple dichotomous question, “Are you interested in employment as a goal for the future”, Westcott et al. (2015) found that 85.1%, in a sample of 255 of participants with a schizophrenia spectrum diagnosis living in the community, were either already employed or had an interest in employment as a goal in the future. Seebohm and Secker (2005) recruited and trained interviewers with a diagnosis of a severe mental illness. In this study participants were asked if they were interested in obtaining work of any kind including voluntary work. Answers were coded, “yes”, “maybe in the future” and “no” (p. 12). Only 10.0% of participants responded with no interest. Of the 141 participants who were interested, 132 (95.7%) were interested in paid employment. Furthermore, 116 participants (82.3%) wanted assistance to obtain work.

These two studies indicate that interest in employment in people with severe mental illnesses might be greater than previous research has identified. Considering this research, a better approach may be for employment programs to assume that most participants are generally interested and to explore how much participants value or fear employment, why they want employment, and what kind of help they think is important. Exploring these underlying reasons, especially when a participant is not
engaging in the program, could assist vocational staff to tailor employment assistance and improve participant responsiveness.

**Participants’ perceptions**

Participants’ perceptions of the vocational rehabilitation program and the match with their perceptions of themselves, are likely to influence their responsiveness when offered vocational assistance. Previous research has examined participants’ perceptions including:

(a) employment-related self-efficacy (Waghorn, Chant, & King, 2005a, 2007) and self-esteem (Corbière, Lanctôt, Sanquirgo, & Lecomte, 2009);

(b) perceived barriers to employment (Corbière et al., 2004; Honey, 2004);

(c) employment values and fears (Hielscher & Waghorn, 2017); and

(d) perceived relevance of the program and how valued program outcomes align to the goals of individual participants and participants’ expectations of success (Corbière et al., 2011).

The influence of participants’ beliefs about whether the program can help them achieve their goal of competitive employment, or how relevant they believe the program is to them, is likely to be important (Carroll et al., 2007). For example, Besse, Poremski, Laliberte, and Latimer (2017) investigated participants’ perceptions of stress in returning to employment. The study found that the way participants perceived the support from the vocational rehabilitation program was a positive influence on their levels of perceived stress.

**Employment-related self-efficacy and self-esteem**

Self-efficacy is considered an important influence on participants’ behaviours, particularly program engagement and participation (Corbière et al., 2011). Self-efficacy refers to a person’s reported confidence in successfully performing specific tasks. Self-efficacy can vary across a wide range of situations and types of task. Employment-related self-efficacy has been conceptualised and measured with respect to a range of tasks relevant to participating in vocational rehabilitation and to being employed (Corbière et al., 2011; Waghorn et al., 2005a). People with severe mental illnesses can also have poor or fluctuating self-esteem, and in particular, poor self-esteem as worker (Corbière et al.,
2009). Low levels of self-esteem can increase vulnerability to social stigma and internalising negative beliefs about one’s own job capability. This could also reduce employment-related self-efficacy (Corbière et al., 2011).

For example, Corbière et al. (2011) studied 281 participants with severe mental illnesses. Those with strong beliefs about their ability to effectively job search, who perceived less barriers to employment, who placed a high importance on employment in their lives, and who had encouragement from others, were more likely to actively engage in job searching (Corbière et al., 2011). Negative past experiences in the workplace and when seeking employment are likely to have a negative impact on employment-related self-efficacy (Peterson et al., 2017).

Waghorn et al. (2005a, 2007) developed and tested a measure of employment-related self-efficacy. The authors noted that self-efficacy is best measured at the specific task level. This is because generic measures of self-efficacy have been shown to be poor predictors of human behaviour (Waghorn et al., 2005a). The scale asked participants to rate themselves in relation to four dimensions reflecting career planning skills, job securing skills, work-related social skills, and general work skills. The authors found a moderate positive (r=0.41, p<.001) association between participant’s employment-related self-efficacy and their current employment status (Waghorn et al., 2007). Understanding employment-related self-efficacy may be important. Participants with lower levels of employment self-efficacy may need more assistance in vocational rehabilitation than those with higher levels of employment self-efficacy to optimise their engagement and participation in the program.

**Perceived barriers to employment**

To improve participant responsiveness program leaders and vocational staff could also explore participants perceived barriers to employment. Whilst perceived barriers vary among individuals, research can identify the common barriers reported by people with severe mental illnesses. These can then be discussed with individual participants, particularly those showing signs of low levels of engagement and participation, to check whether they perceive similar barriers to employment.
This type of exploratory research is useful and informative. In a study testing the Barriers to Employment and Coping Efficacy Scale (BECES), Corbière et al. (2004) sampled the views of 254 participants with severe mental illnesses in vocational rehabilitation programs in Quebec. All 43 items in the scale were relevant to participants while 10 types of issues were commonly reported as barriers to employment:

(a) prolonged absence from the workplace;
(b) stressful events (bereavement, break up);
(c) pressure related to the job;
(d) lack of self-confidence;
(e) stress related to job search;
(f) employers’ prejudices;
(g) competition in the workplace;
(h) anxiety and fears;
(i) lack of employer flexibility; and
(j) asserting oneself with co-workers (Corbière et al., 2004, p. 468).

These identified barriers could offer a useful starting point for vocational staff to explore perceived barriers to employment with caseload participants.

**Employment fears**

Another potential influence on participant responsiveness is participants’ fears about starting employment, or returning to employment (Peterson et al., 2017). These fears are likely to be influenced by past negative experiences at work such as losing a job, unfair discrimination, and even self-stigma around the likelihood of getting a job (Hielscher & Waghorn, 2017). Unacknowledged or unaddressed fears are likely to result in a lack of engagement in program meetings, job search activities, or exiting the program completely.

Participants in a European vocational rehabilitation program reported multiple employment fears. These included fears that they wouldn’t be able to do the job properly or that they would have
problems with colleagues (Koletsi et al., 2009). Participant fears also centred on being seen only as a psychiatric patient by potential employers. These fears had a negative influence on their job search behaviours. Working with employment specialists to address these employment-related fears was reported as helpful. Participants also reported fearing social disapproval, and for some participants receiving welfare benefits was a major contributor to employment-related fears (Koletsi et al., 2009).

Similarly, in a study examining the information needs of participants, clinicians, and employment specialists, participants reported how self-doubt and lack of confidence combined with fears of stigma and discrimination and fear of losing welfare benefits. These combined into a negative impact on their views of employment (King et al., 2011). Information provided by the vocational rehabilitation program can help to normalise and overcome these fears by countering incorrect beliefs, supporting a person to manage personal information, and providing financial guidance. Relevant information may also be important to address the employment-related fears held by family members and clinicians (King et al., 2011).

Re-conceptualising participant responsiveness in this way shows that there is much rehabilitation staff and employment specialists can do to reduce the negative impact of participants’ perceptions, beliefs, and concerns. Interventions can be designed for each person based on a more detailed exploration of the issues involved and the impact on the person. For example, if an individual participant reports ambivalence about employment as a personal goal, rehabilitation professionals could use specific questions to explore the participant's perceptions of relevant situations which may discourage their participation in future employment and influence their engagement in the program. Program leaders could develop staff competencies in exploring, and helping to address, participants’ employment fears. The Employment Fears Scale, constructed by Hielscher and Waghorn (2017), provides some prompts for employment specialists to use. For example, the items in the scale asked about fears around (a) being treated differently at work or discriminated against, (b) losing government welfare benefits, (c) employers focusing on limitations not abilities, (d) being marginalised in the workplace, and (e) being unable to work at times due to poor or fluctuating health.
Deeper exploration of participants’ perceptions by vocational rehabilitation staff could lead to better directed employment assistance and improved participant responsiveness (Lloyd, Tse, Waghorn, & Hennessy, 2008). Exploring relevant issues such as past experiences in the workplace, perceptions of barriers to employment, employment fears, employment values, and a participant’s confidence in their ability to work, can provide program leaders with tailored information to modify the delivery of the program in order to improve participation and retention. This has a stronger rationale than focusing on fixed individual characteristics that are challenging to influence (Schroder, Sellman, Frampton, & Deering, 2009; Waghorn et al., 2007).

**Perceptions of program relevance**

To improve participant responsiveness, it is also important that programs are perceived as relevant, i.e. that program goals align with participants’ goals (Waynor & Pratt, 2011). Program materials and activities also need to be designed to reflect and meet the needs and circumstances of participants (Schroder et al., 2009; Small, Cooney, & O’Connor, 2009). If programs are not perceived as relevant to participants then participant dis-engagement is likely (Carroll et al., 2007; Blank et al., 2011). Relevance includes delivering supports at the right time as well as in the right way (Brun & Rapp, 2001; Small et al., 2009). This might involve increasing the intensity of program assistance when a participant is about to start a job, or when a participant is working but there are significant changes in their workplace. Program leaders can improve program alignment and relevance by conducting regular surveys to gather participant perceptions (Kirkpatrick & Kirkpatrick, 2009, p. 10). This can be done immediately following program activities by checking participants’ experiences at the end of sessions using Likert response rating scales (Pereira & Marques-Pinto, 2017).

**Other participant characteristics**

Many investigations in vocational rehabilitation have attempted to determine the individual traits most likely to influence employment outcomes (Burke-Miller et al., 2006; Tsang, Lam, Ng, & Leung, 2000). Individual trait-like characteristics investigated include medical and psychiatric diagnoses, level of functioning, age, gender, ethnicity, education levels, and co-occurring substance use
disorders and physical health conditions (Bond & Drake, 2008; Burke-Miller et al., 2006; Campbell, Bond, & Drake, 2009; Corbière et al., 2011, 2017; Jonsdottir & Waghorn, 2015; Metcalfe, Drake & Bond, 2017b; Taylor & Bond, 2014).

Individual characteristics such as age, employment restrictions (Waghorn, Chant, & Harris, 2009a) and work history (Tsang et al., 2000), have consistently been found to be associated with employment status in population-level studies. Investigations of participants in supported employment programs have found that the nature of the assistance offered can override the disadvantages of many individual demographic and clinical characteristics (Bond & Drake, 2008; Campbell et al., 2009; Metcalfe et al., 2017b; Wescott et al., 2017). Also, the individual characteristics most likely to influence employment outcomes, may be different to those that influence employment interests (Westcott et al., 2017). For example, age may influence interest in employment, with younger people (aged 18-34 years) often reporting a stronger interest. Similarly, a good work history may help sustain interest in employment. Westcott et al. (2017) found that whilst interest in employment declined with age, it did not decline with severity of clinical symptoms. Whilst age influenced interest, the authors found that 64.8% of participants aged 45-65 years retained an interest in employment. The only clinical indicator negatively associated with employment interest was the presence of current active hallucinations (Westcott et al., 2017). Corbière et al. (2011) found that the length of time a participant had been unemployed negatively influenced job search behaviours and commencing employment, and negatively influenced participants’ attitudes to work.

These findings from the vocational rehabilitation literature are consistent with research from the wider health literature on treatment retention. In a sample of young people attending alcohol and drug treatment in New Zealand, Brorson, Arnevik, Rand-Hendriksen, and Duckert (2013) found that individual characteristics like age, gender, ethnicity, and diagnoses were not associated with program retention. Instead, dynamic individual characteristics such as interest in the treatment, and expectations of treatment success, coupled with program characteristics, were associated with retention in the program (Schroder et al., 2009). Whilst some demographic characteristics can reduce participants’
probability of successfully commencing employment, such individual characteristics need not be a basis for exclusion, but instead can be used for matching individuals to an appropriate intensity of employment assistance.

**Employment values**

One individual characteristic which has received less investigation is participants’ employment values. Employment values are likely to include the level of importance participants place on employment, their work preferences, and the anticipated satisfaction that results from being employed (Zaniboni et al., 2010). A study in Australia explored the employment values of 25 people with severe mental illnesses to test the psychometric properties of a brief Employment Values Scale (Hielscher & Waghorn, 2017). In contrast to participants’ fears, which appeared to fluctuate based on past experiences participants’, values were relatively independent of past experiences, self-stigma, work history, and participants’ employment fears (Hielscher & Waghorn, 2017).

Zaniboni et al. (2010) developed a measure for employment values, the Work Values Questionnaire. There is a long version (58-item) and a shortened version (30-items) of the measure. Both versions of the scale were tested with 254 participants with severe mental illnesses in a vocational rehabilitation program in Quebec. Higher scores on the Work Values Questionnaire were correlated with commencing employment. The study also found an association between work values, employment-related self-efficacy, and the coping mechanisms to overcome barriers to employment. Participants with higher work values were more able to cope with barriers to employment and had increased employment-related self-efficacy (Zaniboni et al., 2010). These findings suggest that employment values may be more trait-like characteristics which remain stable over time and are independent of several other individual characteristics and personal attributes.

Work values and work ethics can be shaped through the cultural transmission of values, beliefs, norms, and behaviours across and within generations. This has been identified as a particular issue for participants who have grown up in households with both parents in receipt of government welfare payments (Baron, Cobb-Clark, & Erkal, 2008). This suggests that low work ethics and low value for
employment could be more enduring trait-like quality than some of the other contributors to employment motivation like employment fears and employment-related self-efficacy. Understanding and exploring participants past experiences, and in particular asking about family circumstances, could help to identify if a participant is likely to need more or less intensive employment assistance, with time to explore pathways to break inter-generational cycles.

In summary, vocational rehabilitation is intended for people for whom returning to employment is unlikely to happen without individual employment assistance (Waddell et al., 2008, p. 8). Most participants will have had previous negative experiences of employment, or no experience of employment, which will have shaped their employment-related self-efficacy and perceived control over the process of gaining employment. Many will also have fears about commencing employment and face actual and perceived barriers. Others will have different values attached to the role and importance of employment in their lives (Peterson et al., 2017). Bandura (1989) argued that self-efficacy can be developed and increased using four strategies.

1. Improving a person’s mastery experience, that is helping them to achieve simple tasks, and then lead on to more complex ones.
2. Social modelling by providing role-models with whom participants can identify who demonstrate the targeted behaviours.
3. Improving a person’s physical and emotional state, helping participants to feel relaxed before attempting something new.
4. Verbal persuasion, that is providing encouragement for a person around task completion and specific achievements.

These general strategies provide a useful starting point which program leaders can apply to modify the employment environment. Bandura’s (1989) suggestions can also be applied when vocational staff identify participants who are not yet engaging in the vocational rehabilitation program (McAlister, Perry, & Parcel, 2008, p. 182).
Participants’ circumstances

The personal circumstances of individual participants can help or hinder their availability for employment and their participation in a program (Bond & Drake, 2008; McQuaid & Lindsay, 2005). These include any circumstances that limit attendance at appointments or limit attendance at work. These could include participants’:

(a) current health status, co-occurring health conditions and impairments;

(b) health treatment arrangements;

(c) other social roles and responsibilities;

(d) housing situation;

(e) means of transport;

(f) financial circumstances;

(g) other obligations, such as to welfare benefit providers, or to other family members;

(h) access to social support; and

(i) any limits on their availability to participate in job searching, and on their ability to be in employment for essential periods (McQuaid & Lindsay, 2005; Waynor & Pratt, 2011).

The internationally agreed definition of unemployment is that the person is “part of the working-age population, without a job, and is actively seeking and available for paid work, or the person has a new job start within the next four weeks” (Statistics NZ, 2017). Simply asking people about their availability for employment offers a neutral way to start exploring individual circumstances. Rather than asking about everything else in a person’s life which may not matter if availability is low, the rehabilitation practitioner or employment specialist can begin by asking about any restrictions on availability. Further research can identify the common restrictions on availability reported by people with severe mental illnesses.

Health status and co-occurring health conditions and impairments

Some researchers have explored participant self-reports of their health status as a predictor of future employment status. A study examining the association between course pattern of illness reported
over the past few years by 380 participants with a diagnosis of schizophrenia or schizoaffective disorder, found that those reporting chronic illness with residual symptoms between episodes were 10 times less likely to be employed than those reporting a single episode with good recovery since that episode (Waghorn, Chant, & Whiteford, 2003). Course pattern of illness was a more potent predictor than many other clinical symptoms. Similarly, in a 24-month study of 1273 people with severe mental illness participating in a vocational rehabilitation program participants’ course pattern of illness and functioning were strong predictors of likely employment status (Razzano et al., 2005).

People with severe mental illnesses with co-occurring substance use disorders and physical health conditions have more negative prospects for attaining employment (Waghorn, Chant, & Jonsdottir, 2011a; Waghorn, Lloyd, Abraham, Silvester, & Chant, 2008). Co-occurring health and substance use issues are also likely to impact on participants’ participation in vocational rehabilitation programs. Whilst this area has not been directly investigated, Waghorn, Hielscher, Atveo, and Saha (2016) examined data from participants in the Australian National Survey of High Impact Psychosis (n=1825). The authors found that many participants with psychotic disorders and long-term substance use disorders were not accessing any employment assistance, and for those who were, it was more likely to be mainstream employment services, rather than more intensive and more suitable vocational rehabilitation (Waghorn et al., 2016). This cross-sectional study could not determine whether this was because they were being excluded from the programs or whether they were less inclined to seek out assistance.

**Other possible influences on availability**

Other possible influences, both positive and negative, on a participants’ availability for employment and for participation in vocational rehabilitation programs include transport (Becker & Drake, 1994), social supports, and social and family networks (Alverson et al., 2006; Corbière et al., 2011; Honey, 2003, 2004). However, in the absence of supporting research to guide a specific approach vocational rehabilitation professionals can explore participants’ interest, values, availability, and employment goals. Exploring and understanding any restrictions on participants’ availability for
employment seems a logical place to start. Understanding this could make a difference to participants’
engagement in the program, once their other commitments were fully understood and accepted.

**Relevant program characteristics that influence participant engagement**

There are characteristics of vocational rehabilitation programs that are likely to facilitate or
hinder participant responsiveness (Blank, Harries, & Reynolds, 2011; Macias et al., 2001). Program
characteristics can impact on participants’ perceptions of the suitability of the program in relation to
their individual needs, and participants’ expectations of success (Drieschner et al., 2004; Honey, 2004).

Potential candidates for relevant program characteristics are as follows.

1. Flexibility, the extent that the program is flexible and can be tailored to individual needs and
preferences for assistance.
2. Cultural alignment of the program to the cultural identity and needs of target participants.
3. Program engagement, how actively the program tracks and reaches out to prioritise and
support participants who do not have an outcome yet.
4. The program’s track record of success.

**Program flexibility**

Programs with mandatory components that are not perceived as helpful by an individual are
likely to impact negatively on participant responsiveness (Johnson et al., 2009; Koletsi et al., 2009).

Any relevant program characteristics that directly influence participation or responsiveness can be
considered. For instance, a program that: (a) has a standard profile of assistance offered to all
participants; (b) does not identify individual preferences for the type of assistance to be provided; and
(c) does not respond by tailoring standard forms of assistance to individual preferences, is unlikely to
induce participant responsiveness (Areberg et al., 2013; Johnson et al., 2009; Brun & Rapp, 2001).

Such a program is more likely to degrade responsiveness, disengage participants, or cause early
attrition from the program (Lauder et al., 2009; Padgett, Henwood, Abrams, & Davis, 2008).

The provision of flexible support was an important feature of employment assistance identified
by Areberg et al. (2013) in their interviews with 17 participations in a vocational rehabilitation
program. Participants wanted to meet at times and in places that suited them. They also had preferred methods of communication, i.e. by text rather than face-to-face (Areberg et al., 2013). Lack of program flexibility has been raised as an issue by participants in vocational rehabilitation programs who did not commence employment (Koletsi et al., 2009).

**Cultural alignment**

Being culturally congruent with participants’ culture or ethnic background may be relevant to improving engagement and retention of participants, particularly when involving participants from indigenous and minority cultures (Beach et al., 2005; Cram, Smith, & Johnstone, 2003; Priest & Lockett, under review). This requires programs to take into account participants’ diverse cultural, spiritual, and linguistic needs (Aggarwal et al., 2016; Brach & Fraserirector, 2000). Evidence for the potential contribution of cultural congruence to participant responsiveness can be found in the disparities in populations, by ethnicity, in terms of who does and does not access health services and treatment programs. Evidence can also be found in disparities by ethnic groups among those who leave treatment programs early (Aggarwal et al., 2016; Beach et al., 2005; Guerrero & Andrews, 2011).

Strategies to improve cultural congruence include having staff that reflect the populations the program is targeted to, particularly indigenous and minority groups. Programs leaders can also search out and include indigenous participants and people from minority groups in order to improve the cultural competency of all staff working in the program (Beach et al., 2005).

Cultural congruency means more than having policies and procedures in place and the right knowledge and attitudes in staff, it is about being able to effectively apply these in cross-cultural situations (Brach & Fraserirector, 2000; Cram et al., 2003). Leaders of culturally competent programs have an ongoing commitment to be culturally congruent. This commitment can include: (a) availability of interpreter services; (b) the recruitment and retention of staff from indigenous and from minority groups; (c) cultural competency training programs for all staff; (d) coordinating services with families and community members; and (e) ensuring the administrative and organisational environments are welcoming and inclusive of different cultural groups (Brach & Fraserirector, 2000).
Embedding cultural competency techniques in programs can improve communication, increase trust, and improve knowledge about effective ways to work with indigenous and minority groups. This in turn, can improve the engagement and retention rates of indigenous and minority populations (Beach et al., 2005; Guerrero & Andrews, 2011). To understand whether cultural competency strategies are having the intended effect program leaders can continually monitor referrals, attrition, and outcomes by different cultural and ethnic groups (Priest & Lockett, under review).

**Program engagement**

Another relevant characteristic is program engagement. This could involve program leaders setting up a proactive strategy to prevent attrition, including a system which can regularly identify participants in the program who have yet to make visible progress. Even in good evidence-based programs, it can take a long time for all participants to gain an observable benefit from their perspective (Metcalfe et al., 2018a). Failure of the program to deliver in a timeframe that matches participant’s expectations is also likely to be a major source of attrition (Amico, 2009).

Metcalfe et al. (2018a) undertook secondary analysis of 2,055 participants in a USA multi-site trial of IPS, examining the likelihood of commencing employment during each quarter over 24 months. Whilst the likelihood reduced over time, participants in the IPS intervention group and in the control group were still obtaining their first job up to 18 months after being accepted onto the program.

A proactive strategy to prevent attrition would involve program leaders establishing tracking and support systems, so that participants who have yet to obtain an outcome can be identified and prioritised. The type and intensity of employment assistance may need changing or an additional vocational staff member could double up on providing employment assistance. A proactive engagement strategy should mean that participants are supported at a pace that matches their needs and aspirations even if this changes, and that variations in the activity and engagement levels of individual participants are picked up early, so that strategies are put in place to minimise those leaving the program without an outcome.
The effectiveness of a vocational rehabilitation program is likely to influence participants’ self-efficacy and expectations of success. Yet few programs evaluate their outcomes sufficiently and even fewer try to communicate the results of evaluations to current, past, and future program participants.

Expectations of success can be enhanced through participants having direct contact with peers who have successfully commenced employment (Corbière & Lanctôt, 2011; Kirsh, 2016). Whilst empirical evidence on the role of peers on improving participant responsiveness in vocational rehabilitation programs is limited, the role of peers, and peer leadership, has been found to improve engagement of participants in other health programs, for example in self-management programs (Druss et al., 2010). Bandura (1989) argued that self-efficacy can be improved through vicarious experience, by judging one’s capabilities in comparison with performance of others with whom they identify and through exposure to others who are successful. To enhance employment-related self-efficacy and expectations of success, program leaders could involve peers in modelling job-search activities and provide examples of how peers previously engaged in the program to commence and sustain employment.

It is suggested that program leaders find ways to communicate their track record. This could be done through marketing materials, through peer workers, or through peer support groups (Kirsh, 2016). In behavioural economics small changes in the environment that are easy and inexpensive can influence participants’ behaviours. Whilst these techniques have been examined in healthcare settings there is limited evidence yet of the value of these small changes in the environment around the delivery of vocational rehabilitation (Drebing et al., 2007).

Measuring Participant Responsiveness

The final stage of the conceptual analysis was the construction of targeted statements and an item pool for measuring these statements. Once further developed and tested, new measures are expected to assist program leaders and vocational staff to identify lower levels of participant
responsiveness, to explore the reasons underlying participant engagement and participation, to match the employment assistance to individual perceptions, circumstances, and characteristics, and to improve responsiveness to the vocational rehabilitation program (see Table 28). Once such measures are further developed and tested, it is anticipated that they will assist program leaders to describe, define, modify, and improve the influences and conditions to enhance participant responsiveness in vocational rehabilitation programs.
<table>
<thead>
<tr>
<th>Measurable attribute</th>
<th>Targeted statements</th>
<th>Measurement item pool</th>
<th>Method of measurement</th>
<th>Examples of questions</th>
<th>Evidence for attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construct specification: Participant responsiveness is how relevant program characteristics and individual characteristics and circumstances combine with participants’ perceptions of themselves, and the vocational rehabilitation program, to influence participants’ employment motivation, and engage and hold their involvement and interest, once they are accepted into the program and commence participation.</td>
<td>Goal-directed behaviours</td>
<td>Identifying responsiveness</td>
<td>The program collects information on participant’s employment-directed behaviours. 1. Job search activities outside meetings. 2. Attendance at meetings. 3. Participation in meetings. (Score one for each anchor that is present).</td>
<td>Interview</td>
<td>Do you have any processes in place where you regularly observe, and collect information on participants’ behaviours once they have commenced assistance in the program?</td>
</tr>
<tr>
<td>Interest in employment</td>
<td>Exploring general interest in employment</td>
<td>VR staff explore individual participant’s interest in employment. (Score one if the anchor is present).</td>
<td>Interview Document review</td>
<td>Do program staff explore how much participants want employment, why they want employment, and what kind of assistance they want? This item seeks to gauge whether interest is explored to tailor the assistance needed rather than to exit someone from the program.</td>
<td>Lehman et al., 2002; Macias et al., 2001; Mueser et al., 2001b; Seebomh &amp; Seeker, 2005; Westcott et al., 2015.</td>
</tr>
<tr>
<td>Participants’ perceptions</td>
<td>Exploring influences on responsiveness</td>
<td>When participants are not engaging or participating in the VR program, program staff actively explore potential influences. 1. Employment-related self-efficacy. 2. Perceived barriers to employment. 3. Coping mechanisms. 4. Employment fears. 5. Program relevance and expectations of success. (Score three if three or more are present, score two if two are present, score one if one is present).</td>
<td>Interview Observation</td>
<td>Ask program staff when the last time they asked participants on their current caseload questions like: How confident are you in your job search abilities? What might get in the way of you getting, or keeping a job? How important is work to you, relative to other things in your life at the moment? Do program staff have a good understanding of participants employment fears, past work experiences?</td>
<td>Corbière et al., 2004, 2011; Hielscher &amp; Waghorn, 2017; Honey, 2003, 2004; Waghorn et al., 2005a; 2007.</td>
</tr>
</tbody>
</table>
| Measurable attribute | Targeted statements | Measurement item pool | Method of measurement | Examples of questions | Evidence for attribute
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived relevance</td>
<td>The alignment between individual characteristics, desires and needs and program delivery</td>
<td>1. The program identifies the goals for each participant. 2. A plan of assistance which aligns with those goals is set up for each participant. 3. Program staff regularly check with participants how the participant experienced the session, i.e. if the session was useful to and relevant for them. (Score one for each anchor that is present).</td>
<td>Program records Interview</td>
<td>Ask program staff how they know whether the program is relevant to participants. Examples might include conducting participants satisfaction surveys, or through holding regular conversations with individual participants, where participant goals are re-visited, and alignment checked</td>
<td>Areberg et al., 2012; Berkel et al., 2011; Carroll et al., 2007; Fukui et al., 2012; Kirkpatrick &amp; Kirkpatrick, 2009; Pereira &amp; Marques-Pinto, 2017; Rapp 1998; Small et al., 2009; Waynor &amp; Pratt, 2011.</td>
</tr>
<tr>
<td>Individual characteristics</td>
<td>Matching assistance to individual needs and individual's perceptions</td>
<td>1. Additional employment assistance is offered to some participants. For example, if participants are older, or if participants have been out of employment for longer periods, or participants have greater fears or greater perceived barriers. 2. VR staff explore the work values and preferences of individuals. (Score one for each anchor that is present).</td>
<td>Interview Document review</td>
<td>Seek documentation on participants work history, and age, is this recorded, and how familiar are VR staff with the work history and age, of participants on their caseload. There is evidence that additional assistance is provided based on individual characteristics and circumstances. How do staff tailor the assistance they provide around individual participants’ concerns, fears, perceived barriers? Ask staff to give examples.</td>
<td>Corbière et al., 2011; Lloyd et al., 2008; Metcalfe et al., 2017b; Westcott et al., 2017; Zaniboni et al., 2010.</td>
</tr>
<tr>
<td>Personal circumstances</td>
<td>Participants’ availability to engage and participate</td>
<td>Vocational staff explore possible restrictions on participants’ availability. 1. To participate in the vocational rehabilitation program. 2. To look for and take up employment. (Score one for each anchor that is present).</td>
<td>Interview Observation</td>
<td>Ask vocational staff about how they would know about individual circumstances and the effect of these circumstances on their participation in the program and in employment. This might include childcare or other caring responsibilities, lack of transportation, regular medical appointments, and ongoing health issues.</td>
<td>Alverson et al., 2006; Corbière et al., 2011; Honey 2003, 2004; McQuaid &amp; Lindsay, 2005; Waghorn et al., 2003 Waynor &amp; Pratt, 2011.</td>
</tr>
<tr>
<td>Measurable attribute</td>
<td>Targeted statements</td>
<td>Measurement item pool</td>
<td>Method of measurement</td>
<td>Examples of questions</td>
<td>Evidence for attribute¹</td>
</tr>
<tr>
<td>---------------------------</td>
<td>-------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------</td>
<td>-----------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>Program flexibility</td>
<td>The extent the program can be tailored to individual needs, circumstances and preferences for assistance</td>
<td>The program takes a flexible approach. 1. Meeting locations. 2. Times of appointments. 3. Length of meetings. 4. Intensity of employment assistance. 5. Program flexibility is communicated to participants (Score three if three or more are present, score two if two are present, score one if one is present).</td>
<td>Interview Review program documentation</td>
<td>Ask program staff about the way they operate, where and when they meet with people, how often, and for how long. Ask them to give examples. Ask what they do when participants cannot attend meetings at the program offices, want to have longer appointments, or need the intensity of support adjusting. Review program marketing materials.</td>
<td>Areberg et al., 2012; Blank et al., 2011; Johnson et al., 2009; Koletsis et al., 2009; Laudet et al., 2009; Rapp 1998.</td>
</tr>
<tr>
<td>Program cultural alignment</td>
<td>The extent to which the program is culturally congruent</td>
<td>1. The program has a formal strategy which outlines how it will meet its commitment to the cultural needs of program participants. 2. The program staff are representative of the cultural needs of the target participants. 3. All program staff have attended cultural competency training. 4. The program has access to interpreter services. 5. Program collateral is available in different languages. 6. The program engages with participants’ family members, if appropriate. 7. The program monitors referrals, attrition, and outcomes by cultural group. (Score three if three or more are present, score two if two are present, score one if one is present).</td>
<td>Interview Review program documentation</td>
<td>Review cultural strategy and policies if available, staff demographics and how this matches the local demographics as well as the target population for the program. Ask staff and participants what types of cultural needs participants have, and how these have or have not been accommodated. Review data reporting to see if information is provided by ethnicity and whether staff are aware of this information.</td>
<td>Aggarwal et al., 2016; Beach et al., 2005; Brach &amp; Fraserirector, 2000; Cram et al., 2003; Guerrero &amp; Andrews, 2011 Cram et al., 2003.</td>
</tr>
<tr>
<td>Measurable attribute</td>
<td>Targeted statements</td>
<td>Measurement item pool</td>
<td>Method of measurement</td>
<td>Examples of questions</td>
<td>Evidence for attribute¹</td>
</tr>
<tr>
<td>----------------------------</td>
<td>------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------</td>
</tr>
</tbody>
</table>
| Program engagement         | Identify participants who have not yet made visible progress                        | 1. Program leaders regularly identify participants who remain in the program with no job outcome.  
2. Program staff know which of their caseload participants do not have an outcome yet.  
3. Participants who have been in the program and have no job outcome are prioritised for employment assistance.  
(Score one for each anchor that is present).                                                                 | Interview Review program documentation | Ask staff whether they know the participants on their caseload who do not have a job outcome yet.  
Ask staff how they are responding to these participants.  
Review minutes of staff meetings and supervision, to see if participants with no job outcomes yet are discussed and prioritised for assistance.                                                                 | Amico, 2009; Metaille et al., 2018a. |
| Track record of success    | The track record of the program is well-known                                       | 1. The program actively promotes stories of people who have been successful in the program.  
2. The program uses peer support workers or peer workers.  
(Score one for each anchor that is present).                                                                 | Interview Review marketing material | Ask program staff how potential and existing participants and family would know about what the program achieves.  
Does the program employ any peer workers?                                                                                                                                          | Bandura, 1989; Corbière & Lanctôt, 2011; Kirsh, 2016; Druss et al., 2010. |

Notes. VR=Vocational Rehabilitation. ¹Green=Category A: The evidence supports the inclusion of the attribute and its effect on program performance. Red=Category B: The evidence does not support the inclusion of the attribute and its effect on program performance. Blue=Category C: No empirical evidence was identified, but authors have written about the potential for the attribute to influence program performance.
Discussion

The conceptual analysis for the dimension of implementation quality participant responsiveness was guided by four questions.

1. What is participant responsiveness in the context of vocational rehabilitation for people with severe mental illnesses?
2. How is the construct participant responsiveness related to participant motivation?
3. What are the conditions or influences which could improve participant responsiveness in the context of vocational rehabilitation for people with severe mental illnesses?
4. What measures are already available to vocational rehabilitation programs to identify and measure relevant conditions and influences on participant responsiveness in vocational rehabilitation programs?

Through a logical and systematic unpacking of the concept and the related concept of employment motivation, this analysis has led to a re-conceptualisation of the construct participant responsiveness. The operational specification of participant responsiveness involves several elements. That is, how relevant program characteristics and individual characteristics and circumstances combine with participants’ perceptions of themselves and their perceptions of the vocational rehabilitation program. How these program characteristics and individual characteristics, circumstances, and perceptions influence participants’ employment motivation, and in turn, how this influences participants’ engagement and participation in the program, once they are accepted into the program and commence participation (see Figure 14).

This operational specification recognises a theoretical link between participant responsiveness and employment motivation. The conceptual analysis drew on empirical and conceptual evidence from vocational rehabilitation, theories of human motivation and behaviour, and the broader science of psychosocial program implementation to identify a set of conditions and influences that, if modified, could enhance the levels of participant responsiveness in vocational rehabilitation. The investigation identified five specific sources of measurable attributes.
1. Participants’ general interest in employment as a current personal goal or a personal goal for the near future.
2. Participants' perceptions of the vocational rehabilitation being offered and the match to their own perceptions about their employment goals, employment capabilities, and perceived barriers to employment.
3. Other individual participant characteristics such as age, gender, work history, educational attainment, communication skills, and the nature of health conditions and impairments.
4. Individual circumstances that can facilitate or hinder participants’ availability for employment, and their availability for participation in the program.
5. Characteristics of the vocational rehabilitation program that can influence participant engagement.

Using an iterative process, candidates for operationalised measurement were identified from all five sources (see Figure 14). These linked the conceptual space for each source to the available literature, identifying specific candidate for influences on participants’ responsiveness. These candidates are recommended for further investigation of their potential to enhance participant responsiveness. Improved levels of participant responsiveness are in turn, expected to improve program performance. This approach could be most useful to programs that have focused first on establishing sufficient fidelity in the application of evidence-based practices, yet their performance remains suboptimal and below program leader expectations.
Figure 14. The influences on participant responsiveness in vocational rehabilitation programs

Notes. The strengths and directions of possible relationships remain unclear and are yet to be investigated. The lines between the boxes do not indicate any more than hypothesised relationships. In addition, possible mediators and moderators have not been identified, this can only be done through empirical research.
Limitations

This investigation took a pragmatic approach using the available empirical and conceptual evidence to help identify the theoretical influences and conditions most likely to improve levels of participant responsiveness. The investigation was more of an exploratory theoretical analysis than a systematic review. Therefore, it is possible that some literature was missed which could have identified other important influences and conditions on behaviours. However, the evidence collected and reviewed can be considered an important starting point to operationalise this construct.

There are conceptual overlaps between participant responsiveness and other dimensions of implementation quality in the conceptual framework. For example, quality of program delivery attempts to measure the resources (premises, offices, staff, supervision, facilities), quality assurance processes, performance frameworks, and health and disability standards. The look and feel of the waiting room and reception area from participants’ perspectives can be particularly important (Nepe et al., 2011). Whereas in participant responsiveness, the target becomes specific qualities of the program considered most relevant to participant perceptions of the suitability of the program to their individual circumstances. Program characteristics considered likely to influence participant responsiveness could extend to include cultural congruence and the alignment of the program's goals with a participant’s personal rehabilitation goals.

This conceptualisation of participant responsiveness does not include the working alliance between the participant and the employment specialist. This influence was included in the conceptualisation of employment specialist expertise discussed in Chapter 7. Participant responsiveness in this section is focused on the wider range of behaviours of participants in vocational rehabilitation programs. Maintaining this distinction between the development of the working alliance, and participant responsiveness, also allows research to further examine the relationship between these two related constructs (Southwick, 2014).
Chapter Conclusions

This chapter presented a re-conceptualisation and re-operationalisation of the eighth dimension of implementation quality in the IFVR, namely participant responsiveness. The systematic unpacking of this complex and multi-dimensional concept could assist program leaders to better describe, define, and modify the conditions and influences that can improve levels of participant responsiveness. Doing so is expected to improve the performance of vocational rehabilitation programs.

The targeted statements resulting from this conceptual analysis, once further developed and tested, could form a subscale to measure participant responsiveness. This subscale combined with subscales for the other seven dimensions of implementation quality in the IFVR, are expected to form a new general measure of program implementation and program delivery. This new general measure would complement the existing fidelity scales.

The application of this new measure is considered in Chapter 9, as well as the implications and limitations of the findings of this wider investigation which sought to improve the effectiveness of evidence-based practices in vocational rehabilitation programs for people with severe mental illnesses.
Chapter 9: Discussion

Introduction

The labour force participation of people with severe mental illnesses remains well below the personal vocational aspirations and capabilities of affected individuals (Drake et al., 2016; Jonsdottir & Waghorn, 2015; Morgan et al., 2016; Peterson et al., 2017). To reduce the personal, social, and economic burdens associated with low labour force participation there is an urgent need to improve the effectiveness of vocational rehabilitation programs designed for this population. This investigation shows how this can be done.

The body of research in this thesis generated new knowledge with the potential to improve the effectiveness of all forms of vocational rehabilitation being offered to people with severe mental illnesses. By taking a broad conceptualisation of implementation quality, beyond adherence to known evidence-based principles and practices, other candidate influences for improving the performance of vocational rehabilitation programs were identified, described, defined, and partially operationalised.

The search for new ways to improve program performance is not unique to vocational rehabilitation. Improving the effectiveness of all psychosocial programs is important across the health and social sciences (Durlak & DuPre, 2008; Sederer, 2009). It is therefore anticipated that the findings from this investigation, once further developed and tested, will have wider application across the broader science and practice of psychosocial program implementation.

This final chapter (Chapter 9) discusses the implications of what has been discovered for improving the effectiveness of vocational rehabilitation programs, and how this may practically be achieved. First the key findings are summarised, compared, and contrasted with previous research. Next, the utility and generalisability of the findings are considered. This includes how the proposed measures of the eight additional dimensions of implementation quality can be used in conjunction with existing fidelity scales to monitor and improve program performance. Also relevant are issues concerning whether the findings of this thesis can be generalised (a) to other countries, (b) to other vocational rehabilitation programs with different populations, and (c) to other psychosocial programs.
more generally. The limitations of the research are outlined next, including the need for further
development of and empirical testing of a new scale. Suggestions for future research directions also
arise from this body of research follow.

The chapter concludes by highlighting how conceptualising implementation quality, beyond
implementation fidelity, could improve the performance of evidence-based practices in vocational
rehabilitation programs. This broader conceptualisation utilises existing knowledge from the broader
science of psychosocial program implementation. The conceptualisation represents an alternative
approach that complements persisting with optimising implementation fidelity, and trialling new
innovative practices other than those currently recommended, to improve the performance of vocational
rehabilitation programs.

Key Findings

The body of research outlined in this thesis has generated knowledge which promises to
advance the science and practice of vocational rehabilitation. This knowledge also has potential
applications across a range of intervention programs in the broader health and social sciences. The
investigation has systematically addressed each of the four research questions outlined in Chapter 1.
The key findings can be summarised as follows.

1. Delivery of a program with adherence to the original principles and practices of the
intervention, implementation fidelity, is important. However, implementation fidelity cannot
guarantee acceptable program performance. Even though low fidelity implies poor program
outcomes, a substantial proportion of high fidelity programs do not achieve acceptable
performance. This is because program fidelity explains only a small proportion (16.8%) of the
total variance in program performance, as measured by the proportion of participants
commencing competitive employment. There is also an apparent threshold of 80% fidelity,
where scores over that threshold do not further predict program performance (Lockett et al.,
2016).
2. A broader conceptual implementation framework beyond implementation fidelity, can assist with the identification and prioritisation of other factors and influences. These other factors and influences, when described, defined, and measured, could further improve the performance of vocational rehabilitation programs once a threshold of good fidelity has been reached (Lockett et al., 2018a).

3. Systematic examination of the local socio-economic policy context leads to a better understanding of how the wider context is helping or hindering the availability and implementation of evidence-based practices. This knowledge can then inform ways to improve the implementation of evidence-based practices in a particular program delivery context (Lockett et al., 2018b).

4. The eight dimensions of implementation quality identified in the new implementation framework for vocational rehabilitation (IFVR) can be sufficiently conceptualised and operationalised to construct a new general tool for measuring the quality of program implementation and program delivery. The tool could complement existing measures of program fidelity.

The association between program performance and program fidelity

This investigation began with Chapter 2 reviewing the evolution of research in vocational rehabilitation since the early 1990s. Substantial progress in the practice and theory of vocational rehabilitation is apparent. One approach to vocational rehabilitation, Individual Placement and Support (IPS) has defined and constructed a set of standard evidence-based practices that can be implemented and measured internationally. A similar strength of relationship between these standardised practices and program effectiveness has been found across different contexts (Drake & Bond, 2017). This is a significant achievement. However, even in this highly developed program, current implementation quality remains suboptimal overall and there is considerable variation in program performance both within and between countries.
Examination of the empirical evidence on IPS enhancements over the last decade, found that none of the enhancements examined have since evolved into new evidence-based practices. Furthermore, whilst factors potentially explaining variation in program performance were identified, they have not been prioritised, or sufficiently described and defined to measure their relative contribution to program performance. The current conceptual approach for implementing evidence-based practices in vocational rehabilitation has an almost exclusive focus on improving program performance through improving program fidelity, as measured by one of two IPS fidelity scales (Drake et al., 2012, p. 112). Yet this conceptual approach could be too narrow. It is possible that further improvements could be made through improving the quality of implementation and program delivery more broadly (Durlak & DuPre, 2008; Sederer, 2009).

This exploration of the evolution of theory and practices in vocational rehabilitation described in Chapter 2 led to Study 1 reported in Chapter 3. Study 1 is a systematic review and meta-analysis of empirical evidence on IPS implementation internationally. The association was assessed between adherence to evidence-based practices as measured through current fidelity scales, and program performance as measured by the proportion of participants commencing competitive employment. Implementation fidelity was found to be a robust and important predictor of program performance across and within international contexts. Fidelity, as measured by the existing IPS fidelity scales, accounted for 16.8% of the variation in program performance (Lockett et al., 2016). This finding is consistent with previous studies of the predictive validity of the current IPS fidelity scales (Bond et al., 2011, 2012c; Kim et al., 2015). Whilst fidelity was a consistent predictor of program performance, the majority of variance remains unexplained. This implies that factors other than fidelity exist. If these other factors can be identified, described, defined, and measured they could be utilised to further improve program performance.

Previous research on the predictive validity of IPS fidelity scales suggest a monotonic relationship between fidelity scores and program outcomes (Kim et al., 2015). However, the systematic review and meta-analysis in Chapter 3 found a more nuanced relationship. Low fidelity scores were,
with only one exception (Morris et al., 2014) associated with low performance. Yet, 37.5% of high-fidelity programs also reported low performance. Some of these low performers also had very high fidelity scores (Lockett et al., 2016). This implies that program fidelity is necessary, but it is not sufficient to expect acceptable program performance. An implication of this finding, is that trying to increase fidelity beyond a minimum threshold to avoid low performance does not necessarily further improve performance (Durlak & DuPre, 2008). The findings reported in Chapter 3 indicate the threshold for the IPS scales is 80% of the maximum fidelity score (Lockett et al., 2016). Below this threshold the risk of poor performance is so high, that program leaders should prioritise improving practices to attain a minimum of 80% fidelity to known evidence-based principles and practices.

The findings from Chapter 3 could differ to previous research because of the way program performance was measured. The primary outcome variable examined was the proportion of participants who commenced employment, with respect to all program participants as the denominator. This analysis only included studies reporting this outcome using individual-level data from a defined cohort over a defined period.

Previous research studies in this field have reported outcomes from undefined cohorts not constituted by individual-level data from defined cohorts. Instead these studies reported point-in-time (usually quarterly) caseload level employment outcomes (Becker et al., 2011a; Bond et al., 2011, 2012c; Kim et al., 2015). Unlike individual-level data from a defined cohort, caseload point-in-time measures do not need to report the same cohort at each time point. In addition, attrition and new participants are not identified or reported. Caseload point-in-time measures cannot differentiate programs with high attrition from those with low attrition, nor do they identify or control the length of time each participant has been exposed to the program. Individual-level data derived from a defined cohort is a more rigorous and reliable measure of program performance (Mann, 2003). For this reason individual-level data derived from a defined cohort became an essential inclusion criterion for selecting studies for the meta-analysis in Chapter 3.
A framework to search for unexplained variance in program performance

Next the focus shifted to how researchers, practitioners, program administrators, and policymakers could explore other factors which could influence program outcomes and explain variation in program performance. The search for such candidates led to Study 2 in Chapter 4. In this review of the broader science of psychosocial program implementation, other potential candidates that could help explain variance in program performance were systematically considered. The review examined conceptual and empirical literature from 21 previous reviews and 22 single studies. The synthesis of this literature identified 23 contextual factors and eight dimensions of implementation quality, all in addition to the influences and practices that constitute implementation fidelity. Candidates were selected and prioritised based on their relevance to vocational rehabilitation, the weight of empirical evidence, and their potential for operationalisation.

These candidate influences were further synthesised to construct a new conceptual implementation framework for improving the performance of vocational rehabilitation programs, named the Implementation Framework for Vocational Rehabilitation (IFVR) (Lockett et al., 2018a). The new framework proposes and defines a broader concept of implementation quality beyond implementation fidelity by taking into account the results of this investigation and that of previous research.

Other researchers have proposed candidate influences for improving the performance of vocational rehabilitation programs (Drake & Bond, 2011; Drebing et al., 2012; Kukla & Bond, 2012; Marshall et al., 2014). These reports have explored some of the influences identified in this analysis namely: employment specialist’s competencies, participant characteristics, society features, cultural issues, labour laws, health and disability policies, government regulations (Drake & Bond, 2011), discontinuing non-evidence-based practices, and providing technical assistance (Kukla & Bond, 2012). Other candidates have included building support from stakeholders, improving job development strategies, and addressing participants’ motivation to work (Kukla & Bond, 2012). Whilst research in vocational rehabilitation has generated suggestions for possible influences, a logical and systematic
mechanism for identifying and classifying potential candidates has been absent. To the author’s knowledge, the IFVR is the first attempt to identify, classify, and prioritise these influences in a systematic way.

Other researchers have also recently recognised the need for a logical and systematic unpacking of contextual and implementation factors. Metcalfe et al. (2018b) examined the relative influence of participant, program, and environmental characteristics on employment outcomes using data from a large-scale RCT of IPS in the USA (Drake et al., 2013). Whilst participant work history, local unemployment rate, and length of time a participant was on a welfare benefit, were significant predictors of commencing employment, their full regression model explained less than 1% of the total outcome variance. The authors, two of whom were involved in developing the IPS approach, called for more research that focuses on “deconstructing prominent client-specific characteristics such as work history and on measuring client and environmental factors that have yet to be studied” (Metcalfe et al., 2018b, p. 89).

This recommendation is consistent with the aims of the proposed new implementation framework, the IFVR. The IFVR hypothesises that there are eight other dimensions of implementation quality in addition to program fidelity, which may directly influence program performance. In addition, 23 candidate contextual factors were identified as likely to have an indirect influence on program performance (Lockett et al., 2018a).

**The wider socio-economic context and program implementation**

The proposed new IFVR is currently only partially operationalised. It needs further development and testing. One particular area for consideration is the impact of the service delivery context on the uptake and implementation of evidence-based programs, which is expected to be context dependent. Understanding the delivery context could enable local programs to identify opportunities and constraints to target effort and improve implementation quality. It can also help policymakers to improve the service delivery context so that any downstream constraints on program effectiveness are
removed. The identification of how the wider environment affects delivery is an important precursor to the development of effective implementation strategies (Killackey & Waghorn, 2008).

This rationale led to Study 3 reported in Chapter 5. This is a systematic examination of the wider socio-economic policy context in New Zealand. This country was chosen as the specific delivery context to investigate because the author now resides in New Zealand. Compared to other countries, New Zealand has a relatively straightforward administrative system with a central government and no state or federal governments, which limits the scope of the relevant documentation needing examination. A policy analysis was conducted which examined how the economic and social context helps or hinders the implementation of current IPS principles and practices. The analysis found that the wider economic and social policy context is currently impeding the availability of IPS programs in New Zealand. In addition, even where IPS programs are available, the context is limiting the ease and quality of implementation (Lockett et al., 2018b).

Previous research has examined the influence of the wider country-specific policy context on the implementation of evidence-based practices in vocational rehabilitation. King et al. (2006) examined the Australian policy and service environment. The authors found that the environment was supporting the provision of employment assistance services separate from mental health clinical services, rather than the practices of integrated employment and mental health treatment and care services. In the USA, Hogan et al. (2014) examined the sources of funding, and mechanisms for financing evidence-based practices in vocational rehabilitation, in particular the implications of the new Affordable Care Act. They argued that a national campaign was needed to lever the potential of these legislative changes on the financing of employment assistance for people with severe mental illnesses.

The influence of the wider economic and social policy has also been investigated using data from IPS implementation internationally. Metcalfe et al. (2017a) examined 30 sites in 21 RCTs in 12 countries, to understand the effects of the wider economic and social context on differences in performance between the IPS interventions and the control sites. Weaker employment protection legislation, as measured through the OECD Employment Protection Legislation Index (Nicoletti,
Scarpetta, & Boylau, 1999), increased the relative advantage of IPS performance. Whereas, more generous disability benefits weakened the relative effects of IPS compared to comparison employment services. Whilst this analysis did not identify the mechanisms by which the economic and social policy context impacted on performance, i.e. through lower fidelity, it does highlight the impact of the delivery context on the relative performance of IPS programs, compared to other vocational rehabilitation programs.

The potential importance of country-specific influences on IPS implementations is also highlighted in a recent Norwegian RCT of an IPS intervention. The authors provided a detailed description of the study context, particularly welfare and labour market legislation, and how these may have constrained the outcomes achieved (Reme et al., 2018). For example, compared to other countries implementing IPS programs, Norway has relatively generous welfare payments and stronger employment protection legislation (p. 7). Both these contextual factors have been found to negatively influence the relative effectiveness of vocational rehabilitation interventions (Metcalf et al., 2017a).

**Operationalising the new implementation framework**

The final stage of the thesis was captured by Study 4 reported in Chapters 6, 7, and 8. This is a conceptual analysis of the eight dimensions of implementation quality in the new conceptual framework, the IFVR. This resulted in the partial operationalisation of each construct, ready for further refinement of the item pools and empirical testing. The constructs in the framework are complex and multi-dimensional. It is therefore important that for each one a precise conceptualisation is developed based on sound underpinning by theoretical principles. It is anticipated that this process will help maximise the construct validity of the scales and subscales as they are further developed and tested (Clark & Watson, 1995; Kanfer et al., 2008, p. 22).

The conceptual analysis involved a seven-step process, starting with an in-depth review of the wider health and social sciences literature combined with vocational rehabilitation literature. The process of conceptual analysis led to the identification and prioritisation of the measurable attributes and targeted statements considered to best represent the conceptual space for each of the eight
additional dimensions of implementation quality in the IFVR. To the author's knowledge, this is the first-time a broader set of dimensions of implementation quality beyond implementation fidelity, have been conceptualised and partially operationalised for the purposes of improving the effectiveness of vocational rehabilitation programs for people with severe mental illnesses.

The conceptual analysis of the dimensions of implementation quality, employment specialist expertise and participant responsiveness, results in a different description and definition, and different candidate measures for these constructs, compared to previous research (Corbière et al., 2011; Glover & Frounfelker, 2011, 2013; Southwick, 2014; Taylor & Bond, 2014). The operationalisation of employment specialist expertise used the available empirical evidence to synthesise and propose a set of eight behavioural indicators that distinguish higher performing vocational staff from lower performers. The likely knowledge, skills, and attitudes underpinning these behaviours are described and discussed, and a targeted item pool constructed to help program leaders identify, describe, assess, and develop the performance of vocational staff. Most, but not all, previous research focused on identifying employment specialist competencies, rather than the attributes that differentiate higher performing vocational staff from lower performers. Also, research has not always clearly distinguished between those attributes that are likely to be trait-like, and less amenable to training, from those that are state-like and could be developed and improved. Previous research has also relied on self-report, which has limitations due to social desirability response bias (Breitenstein et al., 2010).

The re-conceptualisation of participant responsiveness recognises a likely relationship between participant motivation and participant responsiveness, and makes a conceptual distinction between the two. The focus however, remains on the more observable behaviour of responsiveness, not the more elusive and less observable internal state of motivation. The analysis identifies, describes, and prioritises a set of conditions and potential influences that program leaders can explore to directly target levels of participant responsiveness in vocational rehabilitation programs.

By distinguishing between the concepts of participant responsiveness and participant motivation, this analysis contrasts with previous definitions in the literature. Previous research has
sometimes conceptualised motivation for employment as an intrinsic and therefore potentially unchangeable personal characteristics of individual participants (Alverson et al., 2006; Macias et al., 2001; Mueser et al., 2001b). This can lead to the development of measures which exclude individuals from employment assistance. An assessment of high employment motivation can also lead to the assistance needs of individuals being underestimated, and consequently a lower intensity of employment assistance being provided (Alverson et al., 2006; Lam et al., 2010). In contrast, the approach taken in this conceptual analysis recognises and identifies the influences of the program itself on participant responsiveness, making suggestions for how program leaders can improve these influences.

Implications and Applications

The utility of the new assessment tool

The construction of the new IFVR is motivated by the knowledge that good fidelity is necessary but not sufficient for good program performance (Lockett et al., 2016). In the context of the IPS program, program leaders first need to measure adherence to IPS principles and practices using one of the two validated fidelity scales. Where the total score is less than 61 out of a maximum score of 75 on the IPS-15 scale, and less than 100 out of a maximum score of 125 on the IPS-25 scale (80% fidelity score), it is reasonable for the primary focus for program improvement to remain on implementation fidelity. However, once these thresholds are attained, the focus should be broadened. This is because higher fidelity scores over this threshold were not correlated with higher program outcomes in a large systematic review (Lockett et al., 2016). Instead of continuing to drive fidelity scores further above the threshold, a more productive strategy for program leaders could be to review their programs with respect to the other eight dimensions of implementation quality identified and prioritised in this thesis (Lockett et al., 2018a).

Program performance also needs to be reliably measured. It is therefore also important that cohort-based outcome measurement is implemented, along with recording attrition and new program
starters. Although many IPS programs monitor performance in terms of the quarterly proportion of the current caseload in employment, this variable is suboptimal because it does not represent individual-level data. Although there may be a correlation between caseload data slices and individual outcomes from a defined cohort, the former lacks sufficient precision for tracking program outcomes when investigating the influence of program fidelity, and the possible influences of other dimensions of program implementation quality, on program performance.

It is expected that measures of the eight other dimensions of program implementation quality will complement, rather than replace, existing fidelity scales. It is also anticipated that once fully developed and tested, each of the eight underlying dimensions of implementation quality will be sufficiently distinct to support use as independent subscales. However, it is highly likely there will be some overlap and even redundancy, which only further empirical testing can assess and eliminate.

Since most good measures in the social sciences require years of development and testing (Clark & Watson, 1995), the focus of this thesis has been on clear conceptualisation and the generation of sample items, so that the next phases of item pool generation and empirical testing can be advanced by other independent researchers.

**Generalisability**

Generalisability is an important property for measures in vocational rehabilitation. It is not enough that a measure establishes properties of reliability and validity in one country, or in one specific vocational setting. While many background influences in vocational rehabilitation are context-dependent, it is important that the core properties of reliability, construct validity, and predictive validity in particular, can be established in multiple contexts, in other languages, and in a range of countries. Further development and empirical testing of the scale by other independent researchers can explore the generalisability of the IFVR by replicating studies in different countries and comparing the reliability and validity indices obtained. The main limitation to exploring generalisability is generated by the nature of the primary outcome variable. That is competitive employment in the open labour market. This means that only countries with developed market economies (e.g. OECD countries) that
have an open labour market are suitable for studies of generalisability. Less suitable are developing countries that do not yet have a market economy, or countries that exclude the participation of people with health conditions or disabilities from the open labour market.

Another facet of generalisability is the extent a measure is valid and reliable when applied to different populations of program participants, with differing characteristics. However, it is rare for the properties of a measure to be re-established for each new sub-population. For example, IPS principles and practices have now been implemented for a range of different population groups (Drake & Bond, 2017; Reme et al., 2018). The same IPS fidelity scales measure implementation fidelity in these programs. One of the dimensions of implementation quality proposed in the IFVR, program intensity, can vary markedly across different participant group in vocational rehabilitation programs (Waddell et al., 2008, p. 46). Therefore, as part of the ongoing development of the new measures proposed it would be useful to test the validity and reliability indices with different populations of people who use vocational rehabilitation programs. Testing across population groups can help determine whether and how to take account of these differences in the IFVR.

This thesis considered developments in the wider science of psychosocial program implementation. Therefore, the findings could have broader application than vocational rehabilitation programs and may help improve the performance of other psychosocial programs. The investigation proceeded in the direction of a general theory of program implementation quality in the social sciences. This involved a systematic approach to conceptualising and measuring the wider context. This in turn led to a method and language for identifying, describing, defining, and measuring the quality of implementation beyond implementation fidelity.

For example, this approach may be applicable to housing assistance programs. A housing assistance program for homeless people with severe mental illnesses, known as Housing First, has developed a fidelity scale. The scale measures the adherence of program practices to the principles of the original Housing First model (Tsemberis, 2011). Initial testing of the fidelity scale identified promising psychometric properties (Stefancic et al., 2013). Further research is needed to identify the
nature of the influence of fidelity on housing program performance. Once a reliable fidelity scale is established, a program can explore the complementary strategies of increasing fidelity, alongside a broader approach applying this new conceptual framework as conceptualised for vocational rehabilitation programs. The dimensions may need re-conceptualising and re-operationalising in relation to housing assistance, but the framework itself could be a useful starting point to design this type of research, as well as in other psychosocial programs.

**Limitations and Areas for Future Research**

**Limitations**

This body of research has limitations which have been identified throughout the chapters. Firstly, the absence of sufficient implementation studies measuring fidelity using the latest 25-item version of the IPS fidelity scale, and the lack of published IPS implementation studies with low fidelity, limited the utility of the meta-analysis reported in Chapter 3. The analysis also relied on a single measure of program performance, the proportion of a defined cohort who commenced employment during a defined period. Whilst the proportion commencing employment is the most frequently used performance measure, alternative performance indicators might include: time to first job, duration of longest job, duration of all jobs in a year, mean hours worked, and mean earnings per week (Bond et al., 2012a; Drebing et al., 2012). Despite these limitations, the findings of the meta-analysis align with previous studies in vocational rehabilitation and the wider health and social sciences. Fidelity at best accounts for an important minority of the variance in program performance. Further studies are needed to better understand the non-linear nature of the relationship, which suggests that the predictive value of IPS fidelity scales occurs only below, and not above, a particular threshold.

Whilst the review of literature, presented in Chapter 4, to construct the IFVR drew on a large body of literature, it is possible that some relevant literature was missed. Consequently, other important candidate dimensions of implementation and contextual factors may not have been identified. Furthermore, the author made an a-priori decision on the available evidence, to reduce the relative
importance of two candidate dimensions of implementation, ‘program reach’ and ‘program adaptation’ (Durlak & DuPre, 2008). Future empirical research could check the validity of this decision, by examining the extent that these dimensions of implementation quality influence vocational rehabilitation program performance.

The background policy analysis presented in Chapter 5 was limited to a demonstration of how this can be applied to one country, namely New Zealand. The method of this policy analysis was in part determined by the nature of the policy context in New Zealand. This means that while the methods were appropriate for a country like New Zealand, with a relatively straightforward administrative and political system, a more nuanced method of policy analysis may be needed for countries like the USA and Australia. These countries have more complex economic and social policy frameworks.

The final part of this thesis, reported in Chapters 6, 7, and 8, focused on developing a clear operational construct specification for each of the eight dimensions of implementation quality. Whilst this was an important process to form valid constructs, the proposed measures need further development and empirical testing, to explore reliability and construct validity in particular. In addition, a conceptual analysis for each of the 23 candidate contextual factors is suggested so that these candidates can be better identified, described, defined, and measured. Once this process of conceptual analysis is completed, the hypothesised influence of these contextual factors on implementation quality, and directly on program performance, can be explored.

Whilst this investigation has limitations it represents an attempt at a thorough and systematic approach to understanding the current performance of evidence-based practices in vocational rehabilitation for people with severe mental illnesses, and the reasons for variation in program performance both within and across countries. The IFVR represents a logical framework for prioritising those factors which could improve the performance of vocational programs. It is anticipated that the IFVR could further improve performance once the optimal fidelity threshold has been obtained. In addition, the IFVR may also be used by program leaders seeking to improve the overall quality of implementation at the same time as working on improving program fidelity. Once the proposed
dimensions of implementation quality in the IFVR have been fully operationalised and empirically tested, it is anticipated that their measurement will help to advance the science and practice of vocational rehabilitation, as well as other psychosocial programs.

**Areas for further research**

There are other areas for further research which arise from this investigation.

1. A repeat of the methods used in the systematic review and meta-analysis examining the predictive validity of total fidelity score and item-level fidelity scores, once sufficient cohorts measuring fidelity using the 25-item fidelity scale are available.

2. Further development and refinement of the item pools for each dimension of implementation quality in the IFVR.

3. Empirical testing to quantify the key psychometric properties of the scale, including utility, construct validity, concurrent validity, structural validity, and predictive validity. Studies of inter-rater reliability and internal consistency will also contribute to understanding the strengths and limitations of the measures being developed.

4. Application of the new scale in different countries operating programs aligned to IPS principles and practices and supporting people with severe mental illnesses.

5. Further conceptual analysis and testing of the 23 contextual factors in the IFVR to examine the relative strengths of these constructs as potential indirect influences on program performance.
Conclusions

This thesis began developing and operationalising a new guiding implementation framework the IFVR, to advance the science and practice of vocational rehabilitation. With further development, testing, and validation the resulting measurement tools from the IFVR promise to assist practitioners, researchers, policy makers, and program administrators to better describe, define, and measure dimensions of implementation quality and program delivery more broadly. It is anticipated that this will further improve the performance of vocational rehabilitation programs once sufficient fidelity to basic program principles, as measured by the IPS fidelity scales, has been obtained. However, the measurement tool could also improve the overall quality of program implementation, in parallel to a focus on improving implementation fidelity.

This investigation proposes a new approach to conceptualising vocational rehabilitation implementation. This approach uses a broader definition of implementation quality beyond simply attaining good implementation fidelity. In doing so it has shown how the implementation of a program might be enhanced, not by further optimising fidelity or by adding new innovative practices, but by applying what is already known from the more general science of program implementation. By taking this approach in vocational rehabilitation it is hoped that more people with disabilities generally, and more people with severe mental illnesses in particular, can be assisted to obtain and maintain employment, and to realise their vocational aspirations. Meeting these aspirations is long overdue.
References


Baydar, N., Reid, M. J., & Webster-Stratton, C. (2003). The role of mental health factors and program
engagement in the effectiveness of a preventive parenting program for Head Start mothers.  

Child Development, 74(5), 1433-1453. doi:10.1111/1467-8624.00616


Medical Care, 43(4), 356-373. doi:10.1097/01.mlr.0000156861.58905.96


Psychiatric Services, 52(3), 351-357. doi:10.1176/appi.ps.52.3.351


Zealand. In W. Gaebel, W. Rossler, & N. Sartorius (Eds.), *The stigma of mental illness-end of the story?* (pp. 263-287). doi:10.1007/978-3-319-27839-1_15


Fairhurst, K., & Huby, G. (1998). From the trial data to practical knowledge: Qualitative study of how general practitioners have accessed and used evidence about statin drugs in their management of hypercholesterolemia. *British Medical Journal, 317*(7166), 1130-1134. doi:10.1136/bmj.317.7166.1130


doi:10.1093/oxfordjournals.schbul.a006930


doi:10.1176/appi.ps.53.3.281


doi:10.1176/appi.ps.54.5.655


doi:10.1037/prj0000127


doi:10.1177/0308022617712199


doi:10.1111/j.1600-0447.2011.01780.x


Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports, 3*(3), 635-694. doi:10.2466/pr0.1957.3.3.635


Minimum Wage Act 1983 s. 6 (New Zealand).


doi:10.1017/sjp.2016.104


doi:10.1192/bjp.167.5.589

doi:10.1111/j.1365-2524.2010.00929.x


mental illness: Findings from the employment intervention demonstration program. *The Journal of Nervous and Mental Disease, 193*(11), 705-713. doi:10.1097/01.nmd.0000185939.11282.3e


Schroder, R., Sellman, D., Frampton, C., & Deering, D. (2009). Youth retention: Factors associated...


doi:10.4276/030802211X13099513661153


doi:10.1177/1078390310394844


doi:10.1177/10442073030140030601


