The development, reliability and validity testing of an observation tool designed to measure recovery orientation in inpatient recovery and rehabilitation services.

Christian Williams

Clinical Supervisor: Dr Andrew Vidgen and Professor Jason Davies
Academic Supervisor: Dr Andrew Vidgen

2016

Thesis submitted is partial fulfilment of the requirement for the degree of Doctor of Clinical Psychology
at Cardiff University
and South Wales Doctoral Course in Clinical Psychology

DECLARATION

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed ……………………………………… (candidate) Date ……………………….

STATEMENT 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of Doctorate in Clinical Psychology

Signed ……………………………………… (candidate) Date ……………………….

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated.
Other sources are acknowledged by explicit references. The views expressed are my own.

Signed ……………………………………… (candidate) Date ……………………….

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ……………………………………… (candidate) Date ……………………….

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loans **after expiry of a bar on access previously approved by the Academic Standards & Quality Committee.**

Signed ……………………………………… (candidate) Date ……………………….

SUMMARY OF THESIS: POSTGRADUATE RESEARCH DEGREES

Please return the completed form to:
School Postgraduate Research Office

Please TYPE or write in BLACK ink and use BLOCK capitals

SECTION A: TO BE COMPLETED BY THE CANDIDATE AND SUBMITTED WITH THE THESIS

Student ID Number:	**0835102**
Title:	**Please circle appropriate value** **Miss**
Surname:	**WILLIAMS**
First Names:	**CHRISTIAN**
School:	**PSYCHOLOGY**
Title of Degree:	**Please circle appropriate value** **DClinPsy,**
Full Title of Thesis	THE DEVELOPMENT, RELIABILITY AND VALDITY TESTING OF AN OBSERVATIONAL TOOL DESIGNED TO MEASURE RECOVERY ORIENTATION IN INPATIENT RECOVERY AND REHABILITATION SERVICES.

Student ID Number:	**0835102**

Summary of Thesis:

AIM: This project aimed to design an observational tool for assessing the recovery orientation of a service, and to test its reliability and validity.

METHOD: From the recovery literature over-arching themes in the recovery process were identified. From these, descriptions of types of interactions between patients and staff members, termed ‘codes’, were developed. Using a Delphi method, experts working in inpatient Rehabilitation services agreed which codes should be included in the tool.

Staff teams in seven wards completed a standardised self-report measure of Recovery. The wards with the highest and lowest mean scores were used to pilot the new tool. Time sampling was used to capture interactions between staff and patients on both wards. These were coded using the new tool. Interactions coded as supporting recovery were rated positive; those coded as hindering recovery were rated negative. It was hypothesised that more positive interactions would be observed on the ward with the highest mean score on the standardised measure. Reliability was assessed by comparing percentage agreement between two researchers independently coding 20% of the interactions. Validity was assessed by comparing the number of positive and negative interactions on the two wards.

FINDINGS: 23 of the 30 codes developed were included in the final tool. The reliability of the tool was good, with 93.8% agreement on positive/negative ratings. However, the validity was poor, with Chi-square showing no significant difference between wards in the number of positive and negative interactions.

CONCLUSION: The evidence base for the tool is sound. Its reliability is encouraging. The poor validity may be because the two test wards did not differ greatly on the standardised measure of recovery used. Further development of both the codes used to construct the tool and the method in which it is used could provide a clinically useful tool for Services.

Acknowledgements

With thanks to both Dr Clare Sandford and Professor David Jason, for all your very selfless hard work.

A special thanks to my academic supervisor, Dr Andrew Vidgen, for seeing this through to the bitter end.

I would like to especially thank all the wards involved; the ward managers for giving up their time, which allowed me to compete this project, to the members of staff who took part, and to the clients for letting me work on their wards. This would have been literally impossible without you.

And finally, to my friends and family, for their love, care and support. I promise I will be more fun now. A special mention goes to my Mum and Dad (Anne and Mike) for their blind faith in me, which never faltered despite the (overwhelming) evidence suggesting they were wrong. To my sisters Emma, Lucy and Caroline, for being eternally positive. And to James and Nat, who have been the ultimate warriors in this war of attrition.

Abstract

AIM: Recovery orientation is important to inpatient mental health services. However, it has proved difficult to measure and hence to implement. Using an observational method, this project aimed to design a tool for assessing the recovery orientation of a service, and to test its reliability and validity.

METHOD: From the recovery literature over-arching themes in the recovery process were identified. From these, descriptions of types of interactions between patients and staff members, termed ‘codes’, were developed. Using a Delphi method, experts working in inpatient Rehabilitation services agreed which codes should be included in the tool.

Staff teams in seven wards completed a standardised measure of Recovery (Recovery-Self Assessment [RSA]). The wards with the highest and lowest mean scores were used to pilot the new tool. Time sampling was used to capture interactions between staff and patients on both wards. These were coded using the new tool. Interactions coded as supporting recovery were rated positive; those coded as hindering recovery were rated negative. It was hypothesised that more positively coded interactions would be observed on the ward with the highest RSA score. Reliability was assessed by comparing percentage agreement between two researchers independently coding 20% of the observed interactions. Validity was assessed by comparing the number of positive and negative interactions on the two wards.

FINDINGS: 23 of the 30 codes developed were included in the final tool. The reliability of the tool was good with 93.8% agreement on positive/negative ratings (Kappa 0.827). However, the validity was poor, with Chi-square showing no significant difference between wards in the number of positive and negative interactions.

CONCLUSION: The evidence base for the tool is sound. Its reliability is encouraging. The poor validity may be because the two test wards did not differ greatly on the RSA. Further development of both the codes used to construct the tool and the method in which it is used could provide a clinically useful tool for Services.

Contents

List of Figures, Tables, Text Boxes and Appendices.

Figures

List of Tables

Text boxes

Appendixes

Chapter 1: Introduction

1.1 Introduction

The aim of this piece of research was to create a tool to assess the recovery orientation of inpatient rehabilitation mental health services, and to assess the new tool's reliability and validity. This has been done by using the evidence base into recovery for the initial development the tool, utilising a Delphi method to create a consensus on the content of the tool, and then piloting the new tool in two wards which were differentiated, based on a standardised measure of recovery.

The introduction chapter will familiarise the readers with layout of the project, and what can be expected from the following chapters. Terms used within this project will be specified. Following this, a review of the evidence into the recovery orientation will be presented, including consideration of models of personal recovery and examination of current methods available for measuring recovery, highlighting the difficulties which are faced. The current state of inpatient care will then be discussed, including an examination of the modes by which care is delivered, with a view to establishing how this information can be best used to enhance services' ability to support personal recovery. Finally, a rational drawing together of the evidence discussed will be presented. This will detail how the evidence has been used to support the development of the tool, which this project has evolved and tested.

Chapter 2 will then provide a systematic review of the evidence of the role of ward atmosphere and the influence it has within inpatient mental health services. This offers an in-depth analysis of a central component of inpatient services. Chapter 3 provides the method, which was developed and followed in light of the evidence which has been reviewed in the first two chapters, leading to the development and testing of the new tool. The result from both the development of the tool and the reliability and validity testing will be presented in Chapter 4. Finally, the results will be discussed and critiqued in Chapter 5, highlighting methodological issues which may have impacted upon the results found. This will also include a discussion of the clinical implications for the use of such a tool and future research.

1.1.1 Terms and key concepts

Recovery: This is a key concept within this thesis, and will be explored and explained in much greater detail within this chapter (see section 1.2). Unless specifically stated, recovery refers to the orientation or model, rather than the process of someone 'getting better.'

Ward atmosphere: There are several terms for this which are used interchangeably across the literature (e.g. ward environment, therapeutic milieu etc). In this project, this concept will be referred to as ward atmosphere. This was decided upon because it was the term most consistently used in the literature (see Chapter 2).

Patients: Terms used to describe people who are utilising mental health services is a debated issue, with different bodies using various terms. For example, the British Psychological Society (in Code of Ethics and Conduct, 2006) and the College of Occupational Therapists both refer to clients (British Psychological Society, 2006 College of Occupational Therapists, 2005). The National Service Framework for Mental Health refers to service user 48 times, patient 12 times and client twice (Department of Health (DoH) 1999). The Royal College of Psychiatrists, in its report ‘Mental Illness: Stigmatisation and Discrimination within the Medical Profession’, refers to patient 77 times, service user 7 times and user (independently of service user) twice (Royal College of Psychiatrists, 2001). Within this project, those within inpatient services will be referred to as patients. This is because firstly there is empirical evidence to support this choice; Simmons, Hawley, Gale and Sivakumaran (2010) study used questionnaires to establish the preferred term receivers of mental health services wish to be known by, from a choice of; service user, patient, client, user or survivor, when being addressed by either; psychiatrist, nurse, psychologist, social worker or occupational therapist. The findings of this study reported evidence-based terminology should be used in referring to ‘patient’ or, in some groups, ‘patient or client’ in preference to ‘service user’. Secondly, within this project consultation from service users who had used inpatient services was utilised to develop certain aspects of the materials. During this consultation, part of what emerged was that ‘patient’ was seen to be the preferred term for those who were currently within an inpatient unit.

Inpatient services: A range of inpatient services exist to meet the need of various populations, such as older adults, adults, children and young people. Adult mental health inpatient services include Acute and Rehabilitation services. Acute services facilitate short term care in periods when people are acutely unwell to a point where community support is inadequate. The aim of such services is to ensure an individual’s safety, provide thorough assessment and develop care plans which aim to set individuals on a trajectory of recovery leading to their needing less intensive services (The Joint Commissioning Panel for Mental Health, May 2013); Within acute services, there are also inpatient Rehabilitation services, which, whilst coming under the Acute umbrella, have some noticeable differences. They cater for a small population of people with high levels of need, they are recognised as being resource heavy. People using such services often require longer periods of stay within inpatient care than might be expected of those utilising acute services, for example. The Joint Commissioning Panel for Mental Health, (November, 2012) clearly states that rehabilitation services are not ‘recovery’ services; as a ‘recovery orientation’ is expected to be central to all health and social care service provisions for people with mental health problems, not limited to rehabilitation services.

Care was taken within this project to ensure that where applicable these differences were respected, for example, the panel used within the Delphi method reflected this (see section 3.2.3.1). However, this introduction aims to examine issues which transcend differences between inpatient mental health services (such as ward atmosphere and the use of a recovery orientation) meaning such differences are not central to the discussion, and as such unless specified, inpatient services refer to any working adult-aged inpatient mental health service.

1.2 Recovery Orientation

Mental health services have seen a growing focus on and interest in, recovery as model or orientation (Slade, 2009). Indeed, in most Anglophone countries, the requirement for mental health services to adopt a recovery orientation is standard policy; e.g. DoH, UK, 2012; Mental Health Foundation of New Zealand, 2008; Department of Health and Ageing Australia, 2009; Mental Health Commission of Canada, 2012; Department of Health, Social Services and Public Safety, Northern Ireland, 2010; U.S. Department of Health and Human Services, 2003). For example, it states in the mental health plan for England 2009-2019:

’…expectation that services to treat and care for people with mental health problems will be… based on the best available evidence and focused on recovery, as defined in discussion with service users.’ (DoH, 2009, pp.7)

The origins of recovery as a vision or concept within mental health have been linked to the survivor movement (Tew, Ramon, Slade, Bird, Melton, & Le Boutillier, 2012), the physical disability movement and the deinstitutionalisation within psychiatry in the United States (Bonney & Stickley, 2008). Presented as conceptually distinct from medical definitions of remission of symptoms, recovery emphasises the re-building of a worthwhile and meaningful life, irrespective of the continued experience of symptoms or not (Social Care Institute for Excellence SCIE *et al.,* 2007). The more traditional model used with mental health is the medical model, which conceptualises mental illness as a brain disease, biochemical or genetic in nature, for which treatments are largely based on medication, and for which the prognosis is one of long-term maintenance (Donnelly *et al.,* 2011).

In contrast, the recovery model has grown out of the lived experiences of people who have been diagnosed with mental illness, and via trial and error have learnt what helps and what hinders them (Davidson, Lawless, & Leary, 2006). Campbell-Orde, Chamberlin, Carpenter, Jeff, (2005) argued that a key difference between these two models is the distribution of power; in the medical model professionals set the direction and tone of the system and the primary role for patients is to be compliant and accept the treatments that are offered. Whereas, a recovery orientation stresses the idiosyncratic nature of the individual's journey and goals. The system is seen to be scaffolding to support the direction the individual wishes to pursue, putting far more emphasis on peer support and on daily life than on medical treatment per se; although it is recognised that such treatment is an important element for some people (Campbell-Orde, *et al.,* 2005).

Despite recovery not being a new concept within mental health, the prevalence it holds as an orientation, and the position afforded it by policy, what 'recovery' is, is a debated topic, with a multitude of meanings being assigned: recovery as an idea, a movement, a paradigm, a philosophy, a set of values, policy or a doctrine for change (Turner, 2002). In part, this may be because at its heart, personal recovery is a subjective experience (Bellack & Drapalski, 2012). It is acknowledged that there may be overlap between individuals' experiences, but there will be many subjective definitions of recovery, because of its idiosyncratic nature and because an individual's understanding of his/her own recovery may change over time (Slade, Amering, Farkas, Hamilton, O'Hagan, Panther, Perkins et al., 2014). Therefore, when exploring the concepts and considering how services can best adopt such an orientation, it is important to bear in mind that currently, recovery is not an easily defined or measured concept.

When looking for definitions of recovery, an important distinction is offered by Slade (2009) between two concepts within a recovery orientation; that of clinical recovery and personal recovery. Clinical recovery is a concept that has emerged from professionals and it involves reducing symptoms and increasing functioning, whereas personal recovery is a concept that has emerged from people who have experienced mental health difficulties. As such, 'personal' recovery contrasts with traditional clinical recovery. The most commonly cited definition of personal recovery (Slade & Davidson, 2011) is from Anthony (1993) who defines it as:

‘A deeply personal, unique process of changing one’s attitudes, values, feelings, goals, skills, and/or roles. It is a way of living a satisfying, hopeful, and contributing life even within the limitations caused by illness. Recovery involves the development of new meaning and purpose in one’s life as one grows beyond the catastrophic effects of mental illness.’ (Anthony, 1993, pp. 11)

The distinction between personal and clinical recovery has also been referred to as; recovery ‘from’ versus recovery ‘in’ (Davidson *et al.*, 2008); scientific versus consumer models of recovery (Bellack, 2006); and service-based recovery versus user-based recovery (Shrank & Slade, 2007). The distinction between the two can be helpful when reviewing steps which have, or are, being taken towards enhancing recovery; and whether the focus is on clinical or personal recovery.

1.2.1 Models of Recovery

A number of models of recovery have been postulated (NIMHE 2004; Barker & Buchanan-Barker 2005) and in recent years, guidelines for developing recovery-orientated practice have become available (Davidson, Tondora, Lawless, O’Connell & Rowe 2009; Slade, 2009). However, implications for working practice are still unclear and reviews of the literature suggest that conceptual clarification on recovery is necessary (Silverstein & Bellack, 2008; Warner, 2009). This has therefore been a driver for research, with work continuing to develop both in terms of models of recovery and into how services can better support the orientation. One such piece of work is the REFOCUS project.

The aim of the REFOCUS project was to develop methods of increasing how recovery-orientated community-based adult mental health services in England are. It was a five year, NHS National Institute for Health Research (Programme Grants for Applied Research), project, which ran from 2009 to 2014 at King's College London.

The project created the REFOCUS Intervention, which was based upon the REFOCUS Model. The intervention’s aim was to increase the extent to which workers supported the recovery of people using mental health services, via their relationships and working practices, which are the two components targeted by the intervention (Bird, Leamy, Le Boutillier, Williams & Slade, 2014). Details of the intervention followed the Medical Research

Council Framework for the evaluation of complex health interventions. Part of the evidence base, which was developed to create the REFOCUS model and intervention included a systematic review of published descriptions and models of personal recovery. 97 papers were included, comprising of qualitative studies, narrative literature reviews, book chapters, consultation documents reporting the use of consensus methods, opinion pieces or editorials, quantitative studies, combining narrative literature reviews with personal opinion, and elaborations of other identified papers. The 97 papers described studies conducted in 13 countries. In two thirds of the papers, their quality was reviewed using RATS (relevance, appropriateness, transparency, soundness) qualitative research review guidelines (Godlee, Jefferson & Clark, 2003). Using this framework, 16 papers which scored 15 or above, (indicating high quality on the RATS score). These papers were used to develop a preliminary synthesis. A modified narrative synthesis approach (Popay *et al.,* 2006) comprising three stages was employed;

- Developing a preliminary synthesis;
- Exploring relationships within and between studies;
- Assessing the robustness of the synthesis.

From this, a conceptual framework of personal recovery was developed. This comprised of three inter-linked, super- ordinate categories: Characteristics of the Recovery Journey; Recovery Processes; and Recovery Stages (Leamy, Bird, Le Boutillier, Williams & Slade, 2011). For the recovery journey, thirteen characteristics were identified (shown in Box i).

Recovery is an active process
Individual and unique process
Non-linear process
Recovery as a journey
Recovery as stages or phases
Recovery as a struggle
Multi-dimensional process
Recovery is a gradual process
Recovery as a life-changing experience
Recovery without cure
Recovery is aided by supportive and healing environment
Recovery can occur without professional intervention
Trial and error process

Box (i) Thirteen Characteristics of the Recovery Journey

For The Recovery Process, five interlinking recovery processes were identified, these processes are;

- Connectedness;
- Hope and optimism about the future;
- Identity;
- Meaning in life;
- Empowerment.

This was captured by the acronym CHIME. Lastly, there is the Recovery Stages framework, which is comprised of thirteen stages which can be mapped onto the five-stage Trans-theoretical Model of Change: Pre-contemplation, Contemplation, Preparation, Action and Maintenance & growth (Prochaska & DiClemente, 1982).

The validity of the framework was assessed with individuals who were currently utilising services, via 7 focus groups in three NHS trusts across England (Bird, Leamy, Tew, Le Boutillier, Williams & Slade, 2014). The data was analysed using deductive and inductive thematic analysis, applying a constant comparison approach. The aim of which was to explore the validity of the categories within the conceptual framework, and to highlight any areas of difference between the conceptual framework and the themes generated from new data collected. It was reported that the super-ordinate categories were evident in the analysis in both the inductive and deductive analysis. However, three areas of difference were reported to be apparent in the inductive analysis. The study indicates that this conceptual framework of personal recovery provides a ‘defensible theoretical base for clinical and research purposes which is valid for use with current consumers’. (Bird *et al.,* 2104, pp 644). Bird *et al.,* (2014) argue that the three areas of difference illustrate the individual nature of recovery, highlighting a need for an understanding of the population and context under investigation.

The validly of the five processes of recovery identified in the model, is also supported by the fact that it is consistent with other research. Bonney and Stickley (2008), completed a systematic review of the UK literature into recovery, which focused on ensuring that voices from service users, policy makers and services were captured, and analysed both as subgroups, and as a whole body of literature. They used thematic analysis and identified six main themes (see box (ii)). These themes show marked similarities to those identified by Leamy *et al.,(*2011):

- Identity
- The service provision agenda
- The social domain
- Power and control
- Hope and optimism
- Risk and responsibility

Box (ii) Six themes identified for the Recovery literature by Bonney and Stickley (2008)

The consistency between the themes from Leamy *et al.,* (2011) and Bonney and Stickley (2008) (Connectedness/The social domain; Hope and optimism about the future/Hope and optimism; Identity/Identity; Empowerment/Power and control) provides a useful basis from which to consider what services should be aiming to embody, even if the definition of recovery remains elusive.

The model presented by Leamy *et al.,* (2011), particularly the recovery processes identified, and emerging themes from Bonney and Stickley, (2008) undoubtedly further our knowledge and understanding of recovery, as they provide a consistent and seemingly agreed upon set of themes with regards the recovery process. However, currently there is limited information available for how services can embody such concepts and how, indeed if at all, such themes can work alongside the aims of services. For example, the experience of being hospitalised has been suggested to be a factor which could hinder recovery if individuals experience services as coercive or paternalistic. It is suggested that such experiences may undermine peoples' sense of their own ability to exercise any form of competent agency, thus increasing a sense of powerlessness, which has been linked to hindering recovery (Hughes, Haywood & Finlay, 2009).

However, risk management is often a primary purpose of hospitalisation (Bowers *et al.,* 2005), and individuals can be compulsorily detained under the Mental Health Act (2007). Slade *et al.,* (2014) identifies the concept of compulsory detention and treatment as one of the seven 'abuses' of the recovery concept. Therefore, for a recovery orientation to be truly embedded at all levels of the mental health system, such involuntary detentions would have to be stopped, or an evidence base for practice developed to help inform professionals of how interactions could be completed to avoid the possible negative impact of involuntary hospitalisation.

1.2.2 Utilising Recovery in Services

Despite pressure from policy, and a growth in support for a recovery orientation in services (Slade *et al.*, 2014) it has proved challenging to develop recovery orientation services. Indeed, it has been argued that such policy drivers, aimed at supporting personal recovery, have been developed ahead of a clear evidence base of what makes a service recovery oriented (Slade & Hayward 2007). Campbell-Orde *et al.,* describes the system within the USA as;

'….an academic one, based on professional expertise, with limited roles for consumers and family members (despite much rhetoric to the contrary).' (Campbell-Orde, *et al.,* (2005) pp. 19).

Furthermore, some commentators suggest the concept of recovery has been 'hijacked' (Wellesley Institute, 2009) by professionals and has thus lost the essence of what a recovery orientation should entail. This may be most helpfully understood if couched in Slade's (2009) definitions of clinical and personal recovery, as a way on conceptualising the nature of the criticisms. It is postulated that there need to be changes made at every level of service to support personal recovery. Slade *et al.,* (2014) argue that the whole mental health system will need to move away from a system where there is a dominance of institutional responses, drug treatments and coercive interventions.

Implementing recovery-oriented practice should be person-centred with a focus on helping individuals live a meaningful life (Farkas, Gagne, Anthony & Chamberlin 2005), which is seen to be in contrast to setting clinical goals that are largely dictated by professionals (Andresen, Oades & Caputi 2005). It is argued that this requires a shift to practice built on equal partnership, hope-promoting and facilitating self-determination (Slade *et al.*, 2014). While treatment is recognised as potentially helping personal recovery, how and when it is offered needs to be addressed, as it can also hinder it, especially if it is the dominant response and is associated with coercive practices (Slade *et al.*, 2014).

Despite these difficulties, progress is being made, such as, the consistent themes presented by research (e.g. Bonney & Stickley, 2008; Leamy *et al.,* 2011) which are indicated as key in the recovery process. Furthermore, the recovery model postulated by Leamy *et al.,* (2011) stated that the recovery processes with the most proximal relevance to clinical practice are: connectedness; hope and optimism about the future; identity; meaning in life and empowerment. As such, it is these themes which should be the focus of clinical practice. However, it is recognised that there are not currently appropriate tools to measure this and that increase in the focus on recovery outcomes and associated concepts such as well-being, will need to be supported into routine clinical practice (Slade, 2002; Slade, 2010). Indeed, the importance of having valid measures of recovery increases as greater emphasis is placed upon trying to better support the orientation. Whilst there are numerous measures available, there are some difficulties with these.

1.3 Measures of Recovery

The need for valid measurement tools of recovery has been identified (Care Service Improvement Partnership (CSIP, 2007)) and the importance of being able to assess both an individual's recovery journey and how services can support such a journey, has been indicated by numerous sources. Donnelly *et al.,* (2011) identified the need for regular measures of personal recovery to be fed-back into services as one of their key policy implications from their review into the recovery literature. There are compendia available which have aimed to collate and evaluate these measures, such as The Human Services Research Institute (HSRI) compendia (Campbell-Orde, *et al.,* 2005), which includes measures of recovery orientation available in the U.S., Burgess, Pirkis, Coombs and Rosen (2011) who published a review of recovery measures, again both for individuals and services, as part of the process of incorporating a recovery approach into the Australian mental health system and Donnelly *et al.,* (2011), who produced a review of methods for reviewing recovery in Northern Ireland. Work into both assessing measures currently available and identifying issues with such, have highlighted some common difficulties.

Burgess *et al.,* (2011) identified several issues with the measures currently available, including:

- The need for different measures for looking at service orientation and an individual
- The need for individual measures to be relevant across the lifespan, through phases of illness and care, across different illnesses, and across cultural diversity
- The need for clarity regarding the definition of recovery.

The tools which do exist for use in assessing the recovery orientation within inpatient services include measures of both personal recovery and of service level recovery orientation, (sometimes termed measures of Recovery Promoting Environments). Indeed, a number of the measures provide versions for various stake holders, such as service providers, patients and family members.

Using these compendia to consider the measures available, each measure varies in terms of the participant(s) it has been designed for, the number of items, and dimensions which it claims to measure. They also differ in terms of the level of validity and reliability testing each has received. Having said this, there are a number of similarities, both in terms of conceptual difficulties they face and in terms of the methodologies used.

With regards conceptual difficulties, what is highlighted by such compendia is that the conceptual basis of the measures is diverse; this reinforces the point previously made, about the difficulty in capturing what is an idiosyncratic process which encapsulates a concept which is not well defined. Opinions within consumer literature about recovery are diverse, and cannot be uniformly characterised (Donnelly *et al.,* 2011). This is reflected in the lack of consistency seen across definitions and understandings of recovery. The multiplicity of perspectives is in itself information which mental health systems can use to shape services, especially when considering how best to measure recovery; namely that no one approach works for everyone. Donnelly *et al.,* (2011,) argue that this highlights a need for caution about the universal applicability of any measure of recovery, where it does not incorporate personal values and goals. It calls into question what should be measured by tools assessing the recovery orientation of a service. One way of doing this would be to use a recovery framework, such as the CHIME framework discussed earlier. In a bid to address the difficulty in measuring an idiosyncratic process head on, Shanks, Leamy, Bird, Le Boutillier and Slade (2013), carried out a systematic review of personal measure of recovery and compared the tools identified to establish how well they assessed the five domains of recovery identified by the CHIME framework discussed (see section1.2.1). However, they reported that while several measures had good psychometric properties, none incorporated all of the key themes within the CHIME framework.

When considering the methodological similarities, they all either measure domains of recovery, in both the personal and service recovery measures, or assess various outcome domains of personal recovery. This does not tackle the difficult issue of needing a measure which assesses a personal process, where people may be at different stages and needing different things from services.

Secondly, all tools currently available are self-report measures, asking participants to rate items on Likert scales. Self-report measures are widely used within a variety of care settings as a method of collecting health related data (Bhandari & Wagner, 2006), and offer numerous benefits such as their speed, cost and simplicity (Coolican, 2009; Hawkshead, & Krousel-Wood, 2007; Paterson, Potoski, & Capitano, 2002). However, there are well established limitations. Self-report questionnaires rely on the honesty of the participants, which is influenced by a number of factors, such as demand characteristics, response bias and image management (Coolican, 2009; Gagné & Gaston 2005; Hawkshead, & Krousel-Wood, 2007). Influences such as these are reported to be more significant when participants feel there is a large power differentiate between researcher and participant (Mitchell, 2011). This is of note when considering measures used in inpatient settings due to the reported power imbalance between patients and professionals (Campbell-Orde *et al.,* 2005). This is also true for members of staff, as it has been postulated that it may be harder for staff in health settings to express honest opinions, due to the clear hierarchy within such settings (Boulkedid *et al.,* 2011). Questionnaires can be carried out anonymously, which may enable participants to be more honest. However, even if participants desire to be honest, such measures require individuals to have adequate introspective ability to provide an accurate response to a question. Furthermore, the level of anonymity participants perceive depends upon participants having a level of trust in the investigators. Anonymous completion may limit who can take part within a self-report measure, as it requires individuals to have adequate reading and writing skills to complete the measures unaided, and correctly to interpret each item as it was intended to be understood (Nind, 2008).

Using a rating scale (as the measures with the compendia do), allows people to provide more nuanced responses than just yes/no. However, using such a scale means participants are required to use the scales as intended by the measures and similarly to others completing the measure. There is research which suggests that people have different ways of filling out ratings scales which produces differences in scores between participants that reflect something other than that which the questionnaire was designed to measure (Austin *et al,* 1998). Such scales will also be influenced by individuals' responding habits; some individuals show tendencies to use the edges of the scales, whereas others constantly use the midpoints and rarely use the most outer points (Austin *et al,* 1998).

Despite the recognition of the importance of such measures, their development and use has highlighted a number of conceptual, measurement and methodological issues. All measures for this area encounter the difficulty of attempting to assess an idiosyncratic process.

There appears to be three main areas of weakness in the tools currently available: first the conceptual understanding of what they are measuring is inconstant; secondly, what is measured to assess recovery is an area of weakness; processes which go on which support or hinder the recovery journey are not closely scrutinised; lastly, the method of collecting data is limited to self-report measures. There appears to be an opening for a measure which assesses the processes or modes by which a recovery orientation might be supported within services. To support services in constructing a recovery orientated environment, considering other methods of assessment may be of benefit. An observational tool is one such method which is already used in health care settings to assess concepts that are difficult to quantify.

1.3.1 Alternative methods of assessment

Of the current tools which utilise observations to assess quality of care within a health setting, the most well developed system is Dementia Care Mapping (DCM©), which is currently in its 8th addition (Bradford Dementia Group, 2005). It is designed to assess and improve the quality of person-centred care received by people with a dementia in residential settings. It is based upon Kitwood's work into person-centred care (Kitwood, 1997). It utilises the premise that person-centred care aims to value people as individuals, understand their perspective, and provide an environment of supportive social psychology, as the social world that surrounds the person can have a positive or negative effect on well-being (Kitwood,1997). The tool aims to evaluate the quality of care from the patient's perspective, whilst recognising the struggle people with cognitive impairments may have in recognising or expressing their needs.

DCM© is a structured observational tool, which involves observing ('mapping') an individual or individuals, in communal areas of dementia wards and periodically recording their behaviour. One of 23 behaviour category codes (BCC) are then used to categorise the observed behaviour. This is used to determine their level of mood and engagement (ME values) during the observed activity. The measure also records any significant interactions with members of staff. The mean of the ME values over the period mapped is used as an indicator of that person's state of well-being for that time period. This is referred to as a WIB score (Well – Ill-Being: WIB). In addition, the percentage of time spent engaging in behaviours that have potential for the individual to reach high levels of well-being can be calculated as potential for positive engagement (PPE).

DCM© has been demonstrated to be an effective (Brooker & Surr, 2005) and valid (Fossey, Lee, & Ballard, 2002), method of both measuring and improving person-centred care within initialisation settings. It has been found to demonstrate good internal consistency, test–retest and inter-rater reliabilities, as well as correlating with other measures of quality of life (Brooker, 2005; Fossey, Lee, & Ballard, 2002) and has been recommended by the National Institute for Health and Clinical Excellence and the Social Care Institute for Excellence (2007). It is well established in a number of different countries, as well as the UK, including Australia and the United States.

DCM© was designed specifically for people with a dementia in residential care settings, however investigations have been made into using DCM© and its modified versions for other populations. It has been implemented in learning disability residential services (Persaud & Jaycock, 2001) and in hospital wards for patients with a variety of physical health problems (Woolley, Young, Green, & Brooker, 2008). Work has also been carried out into its utility within neuro-rehabilitation wards (McIntosh *et al.,* 2012; Westbrook *et al.,* 2013; O'Hanlon *et al.,* 2014; Leigh *et al.,* 2014). Whilst not being designed for these settings or patient groups, researchers have reported it to be both useful and effective in measuring person-centred care, well-being and, of particular interest to this study, commented on the use of an observational tool, in illustrating the activities in those care settings (Persaud & Jaycock, 2001; Woolley, Young, Green, & Brook 2008).

The diversification of DCM© has all been with populations in which individuals may lack the cognitive function to accurately verbalise their needs. Whilst this is not the case in the majority of people with inpatient adult mental health services, the success which the DCM© has demonstrated in measuring difficult to capture concepts highlights the validity of using observations as a measure of an aspect of the quality of care within healthcare settings. The evidence suggests that using observations and a set of predetermined categories to code (or map in DCM©) is a methodology which can produce useful data, both for considering the activities of those within a care setting and for usefully assessing a process, such as the delivery of person-centred care.

The growing interest in and implantation of a recovery orientation within mental health services, and the recognition for the need for other measures to assess the recovery orientation, has occurred alongside a more general change in the ethos and aims of inpatient mental health services. These contextual changes are important to understand the roles and aims of inpatient services, and the back drop against which the recovery agenda is being introduced.

1.4 Challenges in Inpatient Provision

Mental health services have changed dramatically over the past 30 years, arguably more than any other aspects of health care in the UK (Gilbert *et al.,* 2014). The roles and aims of inpatient mental health services have seen considerable change over an even longer period of time, stretching back to the 1950's. In western societies, mental health inpatient services were first established at the beginning of the nineteenth century, in the form of large asylums, and grew in number until the mid-twentieth century (Fakhoury & Priebe, 2007). The number of inpatient beds reached a peak in the UK in the 1950s, when 154,000 beds were available (Davidge, Elias & Hayes 1993). By comparison, there are now reported to be 18,919 overnight inpatient mental health beds in England (18 August 2016, KH03 report, NHS England).

This decrease has been put down to a process of deinstitutionalisation since the 1950s, which has seen the closure of hospitals and a reduction in inpatient bed numbers in the UK and indeed across the industrialised world (Fakhoury & Priebe, 2000). With the reduction of the number of beds, the role of inpatient services has also been seen to change; from the default care provision for people with acute mental illness, to being used in circumstances, and only for periods of time, when support at home or in less restrictive residential settings cannot be managed (Department of Health, 2005).

Deinstitutionalisation has seen an increase in community based support and a change in how, and what, inpatient services are delivering. However, it is also well recognised that for individuals with enduring and complex mental health difficulties, inpatient facilities play a key role in their recovery, and access to such is recommended as part of a comprehensive package of care (Joint Commissioning Panel for Mental Health, November, 2012). Whilst the importance of inpatient care is recognised, changes in both the role of inpatient services and what they are trying to deliver has created new challenges. Such cultural shifts in inpatient care, have led to some researchers claiming that the purpose of acute admission has also been insufficiently defined (DoH 2005, Bowers *et al.* 2009) and little is now known about the modes of operation in acute inpatient wards (Muijen 1999; Quirk & Lelliot 2001; Department of Health 2005). Bowers *et al.,* (2005) concludes that:

‘over the last few decades, acute inpatient psychiatric care has rather lost its way.’ (Bowers *et al.,* 2005: pp.231).

Understanding both the means of operation by which care is provided, and how services are experienced by patients offers greater scope for both the evaluation of services and introduction of alternative orientations of care, such as a recovery orientation.

1.4.1 Model of Inpatient Services

In an attempt to counter the suggested confusion around the role of inpatient services, Bowers, Chaplin, Quirk and Lelliott (2009), developed a theoretical conceptual model of the aims and functions of acute inpatient psychiatry. It was developed from a literature review of reasons for admission to inpatient services (Bowers 2005), and a qualitative study, using semi-structured interviews which were conducted with 47 members of multidisciplinary staff including, Ward Managers, F Grade nurses, Occupational Therapists and Consultant Psychiatrists, from across 14 acute psychiatric wards in an inner city area. (Bowers *et al.,* 2005). The interviews focused on rationales for admission, their care and treatment philosophy, and the roles of different professionals. Qualitative analysis was performed by three members of the research team, where the finest categories were used at first, then were progressively merged into broader domains as commonalities became apparent. This was completed for all three areas of interest. (Bowers *et al.,* 2005).

The model presented by Bowers *et al.,* (2009) provides three conceptually different categories:

1. The factors involved in determining who gets admitted (including the symptoms and problems prompting admission).
2. The function of the admission, including more than the primary purpose of the admitting clinician.
3. The means by which acute inpatient psychiatry functions.

Within this model, Bowers *et al.,* (2009) identified four distinct modes of operation which are evident in inpatient care. Bowers *et al.,* (2009,) postulate that these modes made inpatient services special, and provided the reason they can be differentiated from community alternatives which do not have the same capabilities. The modes were:

Legitimate authority. This authority is sometimes explicit i.e. those detained under mental health legislation, or derived from the social context of being a patient in a hospital staffed by doctors and nurses. When required, this authority can be backed up by large numbers of staff, and is visible to all patients when treatment is obligatory.

1. Containment of the patient, which produces a reduction in risk and social disruption. Within the concept of containment there are three overlapping ideas and inpatient care tries to deliver on all three, these are:

- Intrusion: the breaking of the normal bounds of privacy, personal space, or bodily integrity
- Separation: from people or objects
- Restriction: of freedom of physical movement.

2. Treatment and management. A range of active patient management tasks are used to deliver the primary admission tasks. Traditionally, such tasks are carried out by nurses and occupational therapists and they include: the provision of activities suitable for different patients, a daily routine, monitoring progress and listening to patients, etc.

3. ‘Presence plus (+)’; This refers to continuous staff presence. This is one of the modes via which inpatient services are able to provide these elements.

Presence + has been highlighted as useful for a number of reasons; firstly, the model postulates that tasks within such an environment are less dictated by staff availability, and opportunities for interventions can be capitalised upon whenever they naturally arise. Secondly, the constant staff presence and close proximity of staff and patients is postulated to allow time to be spent with patients in a non-goal orientated way, facilitating development of relationships, which potentially aid assessment and engagement in treatment, and, which can be therapeutic in their own right (Bowers *et al.,* 2009). This highlights the possible importance which the relationships between staff and clients can play in an individual’s experience of inpatient services.

The role of the relationship between patients and members of staff has also been indicated in playing a role in how personal recovery can be supported. Interactions between staff and patients and the relationships that develop are similarly important when considering how inpatient settings can support the recovery process; Mezzina *et al.,* (2006) consider relationships vital to recovery due to their impact on shaping identity and contributing to, or hindering, well-being, which is a significant factor in the recovery journey. Tew *et al.,* (2012) identified the significance of relationships due to their impact in either helping or hindering recovery, and Spaniol, Wewiorski, Grange and Anthony, (2002), found that having one (or more) relationships which provides hope, can be a critical factor in attaining recovery. The role of staff members and patient interactions is likely to be an important mode by which a recovery orientation can be delivered. This can be seen via the REFOCUS intervention, within which one of the two components focus on staff relationships with the patients with whom they work. Furthermore, the role of the therapeutic relationship within inpatient services has been demonstrated to be a key domain of ward atmosphere, which in its turn also correlates with the perceived services recovery orientation (Livingston *et al.,* 2013).

This model provides a framework by which to understand what inpatient services are trying to achieve and the modes by which this occurs. When considering how best to align services with principals of recovery, using such a framework may be beneficial, firstly to consider if the aims of inpatient services need reviewing in light of the recovery movement and secondly, to understand the modes by which care is currently delivered and whether these offer methods of exploring how practice could be changed to support personal recovery. The model presented is evidence based, using a literature review and qualitative research. However, it is an untested model, which has not gone through a validation process since being developed.

With this in mind, when considering how inpatient wards operate, a well-established concept worth reviewing is that of ward atmosphere. Ward atmosphere has been demonstrated to have links to a number of aspects of patients' experience, including satisfaction (e.g. Røssberg, Melle, Opjordsmoen & Friis 2006; Jörgensen, Römma & Rundmo 2009) levels of symptoms (Oshima Mino, & Inomata 2003 and 2005; Beazley & Gudjonsson, 2011) and perceptions about the nature of the recovery orientation of the service (Livingston, Nijdam-Jones & Brink 2013). As with the model presented by Bowers *et al.,* (2009), there is also evidence that indicates that the role of relationships in services has a significant impact on ward atmosphere, as will be explored.

1.4.2 Ward Atmosphere

Ward atmosphere, sometimes termed ward climate, ward environment or therapeutic milieu, has long been recognised within inpatient services and discussed with regards its potential relevance, as an underlying factor in the efficacy of inpatient psychiatric care (World Health

Organisation [WHO], 1953). It was first examined in general psychiatric settings by Moos and Houts, who described it as the:

'…perception of the social-cultural environment, although it also taps [the] prevailing philosophy and value system to some extent' (Moos & Houts, 1964: pp. 604).

This definition would suggest that the ward atmosphere may be a good measure of how aligned with recovery principles services are, as it reflects the prevailing philosophy and value system in use. Ward atmosphere is most commonly assessed using the Ward Atmosphere Scale (WAS). This is made up of ten subscales, covering three dimensions: the Relationship-dimension, Personal growth-dimension and System maintenance-dimension. Other tools, such as the Essen Climate Evaluation Schema (EssenCES) a newly developed shorter measure designed specifically for forensic services, uses similar domains: Therapeutic Hold, 'Patients' Cohesion and Mutual Support' and 'Safety', as these have been identified as key areas influencing the ward atmosphere (Schalast, Redies, Collins, Stacey & Howells 2008). Methods and measures used to capture ward atmopshere are covered in greater depth in the Systematic review in chapter 2 (see section 2.4.5.1 for details). In both cases relationships are seen as a key domain of the ward atmosphere (i.e. the Relationship-dimension, and Therapeutic hold). Since the introduction of the concept of ward atmosphere, numerous studies have examined both the impact ward atmosphere has upon patients and services and the factors which impact upon ward atmosphere. It has been linked to patient satisfaction in both mental health (Røssberg, *et al.,* 2006; Gjerden, & Moen 2001; Middelbøe Schjødt, Byrsting & Gjerris, 2001; Jörgensen, *et al.,* 2009) and forensic services (Bressington, Stewart, Beer, & MacInnes, 2011), outcome (Jörgensen et al., 2009), levels of aggression (Ros, Van der Helm, Wissink, Stams & Schaftenaar; 2013) and levels of symptoms (Oshima et al., 2003 and 2005; Beazley & Gudjonsson, 2011).

As well as appearing to have wide reaching consequences for inpatient care, evidence also indicates that there are numerous ways in which the ward atmosphere can be influenced. Certain staff training has been reported to improve ward atmosphere (Nesset, Røssberg, Almvik & Friis 2009) and various service characteristics have been shown to relate to the perceived ward environment, including the recovery orientation of a ward; Livingston, Nijdam-Jones & Brink, (2012), reported a significant positive correlation between patients' perceptions of how recovery orientated a forensic service was and perceptions of ward atmosphere (the more recovery orientated the service was perceived to be, the more positive the view of the ward atmosphere). Due to the holistic nature of ward atmosphere, there is also evidence to suggest that it may be a valid way of measuring other aspects of interest within mental health. This will be discussed in greater depth in the systematic review in Chapter 2.

1.4 Rationale

Recovery as an orientation within mental health services is high on the agenda for many stakeholders. However, identifying how the recovery process can be supported within services faces a number of obstacles; operational definitions are not agreed upon and the methods of measurement are limited. The importance of services being able to accurately measure and support recovery has been identified (Donnelly *et al.,* 2011), yet the current situation means services are limited in their ability to assess the relevant aspects of the care they provide, regarding the environment and its impact on an individual's recovery journey. Limitations of the measures currently available have been identified.

The recovery literature has consistently highlighted themes from qualitative research which are considered key in supporting the recovery process. These themes, if they could be operationalised, might help services to better support an individual's personal recovery journey. Embedding such concepts within the ethos of a service is recommended by experts in the field (e.g. Donnelly *et al.,* 2011; Slade *et al.*, 2014), as a necessary change to see a recovery orientation best utilised by a system which is currently more comfortable with the medical model (Slade *et al*, 2014).

However, there are difficulties in what practical advice and/or changes can be made within clinical practice to facilitate these concepts being embedding into all aspects of services. One of the first steps in addressing such issues may be to consider other methods of assessing how personal recovery can be supported by inpatient services. One such method would be to consider the environment of inpatient services, by identifying processes which support recovery to be recognised and enabling any areas of weakness to be addressed. Comparing these processes to broad themes which have been identified as supporting people throughout the recovery process, whatever their stage, may offer a way of navigating around the difficulty of patients at different stages of recovery needing different things, as identified by the research into recovery measures.

Bowers *et al.,* (2009) model of inpatient services highlighted the pivotal role staff members' interactions and relationships have on how inpatient services are able to deliver care, what he termed ‘presence +’. Furthermore, as has been discussed, interpersonal relationships, such as those with staff, have been highlighted as meaningful in an individual’s recovery journey, given their ability to either help or hinder it. Therefore, the manner in which staff members interact with patients could be used as a way of assessing whether the recovery processes identified in both the recovery model discussed (Leamy *et al.,* 2011) and reflected in a Bonney and Stickley (2008) review of recovery literature are being enacted by staff members.

The development of the use of the DCM© as an observational measure, suggests that observing interactions is a meaningful way of assessing aspects of inpatient care. As mentioned, the majority of assessments used in inpatient psychiatric settings are self-reported and tend to focus on outcomes, and this does not necessarily sit well within a recovery model (Tew et al. 2012, Social Care Institute for Excellence *et al.,* 2007).

The work done by Bradford university has demonstrated that the use of an observational tool has proved effective and useful within Dementia care settings in capturing and assessing the qualities of an individual’s lived experience. A similar methodology would provide indicators of the quality of care as assessed by observing staff-client interactions (i.e. the types and quality of interactions between staff members and clients), providing an alternative to self-report measures and could arguably provide a more objective measure. Using the recovery literature as a base would also enable interactions to be viewed through a lens focused on what is most helpful in supporting personal recovery.

1.4.1 Aims and Objectives of the current study

To recap, and for clarity, the aims and objectives of this piece of research are as follows:

1.4.1.1 Aims

- To create an evidence based observational tool to measure the recovery orientation of a service.
- To assess the new tool’s reliability.
- To assess the new tool’s validity.

1.4.1.2 Objectives

To meet the aims described above, the following objectives will be met:

- An in-depth synthesis and critique of the current evidence base into the role that ‘ward atmosphere’ plays in inpatient mental health care will be provided in the form of a systematic review (Chapter 2). Research and Service Implications will also be considered, further developing an understanding the processes via which care is delivered in inpatient services.
- To develop the observational tool anchored within a sound empirical base, two relevant reviews (Bonnet & Stickley 2008; Leamy *et al*., 2011) reviewing the evidence base surrounding the ‘recovery orientation’ will be used by the research team to create umbrella descriptions of types of interactions between staff and patient on inpatient mental health wards, which will be termed codes, and examples of how these codes might be enacted, termed examplars (see Chapter 3 for further details).
- The new tool will be further refined via use of a Delphi method, which will be employed to establish which of the codes developed should be included in the new tool (see Chapter 3 for further details).
- To assess the new tool’s validity and reliability, inpatient Rehabilitation wards will be recruited, and a standardised measure of the recovery orientation of the ward will be employed in each ward.
- The new tool will be piloted in the two selected wards. The data gathered from the piloting will be used to assess the reliability via inter-rater agreement (see Chapter 4).
- The hypotheses relating to the recovery orientation of the two wards and pro and hindering- recovery interactions observed, will be statistically tested (see Chapter 4).
- The findings from all stages of the tool’s development and testing will be discussed, drawing on the available evidence to offer possible explanations for the findings (see Chapter 5).

2. Systematic Review

2.1 Aims and scope of the literature review

The aim of this review is to provide an overview of the literature on ‘ward atmosphere’ and to assess the contribution of clinical, patient and service factors in contributing to the quality of inpatient services. The focus of this review is the role of ward atmosphere within inpatient mental health settings. On reviewing the literature, a number of factors were identified which either impact upon ward atmosphere or were impacted upon by ward atmosphere. As such, numerous factors in relation to ward atmosphere have been included to ensure all relevant evidence was reviewed.

2.2 Systematic Review Question

‘What factors within inpatient services influence or are influenced by ward atmosphere?’

2.3 Method

2.3.1 Literature review strategy

The following databases were searched for articles relevant to the systematic review question:

PsycINFO
SCOPUS
Web of Science (all databases).
Social Care Online
PubMED
AMED
EMBASE
Ovid
PsychArticles

2.3.2 Search terms

The following search terms were used in each of the above databases, using the Boolean term AND an asterisk (or database specific equivalent):

in*patient*, mental*health*, ward*

2.3.3 Inclusion criteria

The following inclusion criteria were used to select relevant articles:
The article must relate to mental health.
The article must be published in peer-reviewed journal
The article must include an adult inpatient sample.
The study must include a measure of ward atmosphere.
The study must consider at least one further factor in addition to ward atmosphere.
The article must be in English.
The article must have been published after 2000.

2.3.4 Exclusion criteria

Sources other than articles (books, conferences or webpages).
Papers which solely examined psychometric properties of tools to assess ward atmosphere.

2.3.5 The review process

2,821 articles were identified using the search terms and databases described above; titles of the articles were reviewed with reference to the criteria of the review, in the cases of uncertainty articles were kept.

Via reviewing titles, 318 articles were identified as potential papers and further reviewed. 40 papers were removed due to being duplicates. The abstracts of the 278 papers were then read. This eliminated a further 47 papers; 21 used a community sample only, 17 did not relate to mental health services and 9 related to the psychometric properties of measure of ward atmosphere. The 231 remaining papers were collected and read in full. 209 were rejected for not meeting the inclusion criteria; 98 lacked a measure of ward atmosphere, 84 did not consider the impact of, or on ward atmosphere, 13 were in a physical-health setting, 9 did not sample an adult population and 5 employed a qualitative methodology which did not consistently assess ward atmosphere.

Twenty-two papers remained and the references of each were examined and articles within them reviewed. This process was carried out for each new paper identified, until no new papers were found. An additional 6 papers were located, leaving 28 articles for the review. Details of the selection process is provided in Appendix 1.

2.4 Reviewing the papers

The remaining studies were reviewed with regard to their focus on ward atmosphere, which was considered in relation to a range of inpatient clinical and service factors. To manage this and increase the meaningful comparisons which could be drawn, the papers were read and divided into groups, based upon what aspect of inpatient services and its relationship to ward atmosphere they were considering. From this, 6 groupings were identified. Information on the studies is provided in tabulated form in Table (i).

- Service characteristics (8 papers)
- Satisfaction (7 papers)
- Differences in staff and patient perceptions of ward atmosphere. (4 papers)
- Patient characteristics (4 papers)
- Symptoms (3 papers)
- Other (2 papers; aggression and helping alliance).

2.4.1 Quality review

A quality review of the papers within this systematic review was completed using a framework for critiquing health research (Caldwell, Henshaw & Taylor 2005). This framework consists of 18 items, on which papers can be scored 0-2. Therefore, the possible score range on this framework is 0-36. For the full table of results and the framework, see Appendix 2. To aid the reliability of the quality ratings, the first four of the papers were rated by two people (the supervisor of the project and the author) independently. The scores were the same on each item for one paper, and varied by one point on one item for the other three papers. Where scores differed there was a discussion, in order to reach a consensus on how the framework would be interpreted.

The range of scores within this body of literature was from 21 (Kerfoot, Bamford & Jones, 2012) to 35 (Dickens, Suesse, Snyman & Picchioni 2014; Jörgensen, Römma & Rundmo 2009). Therefore, on the whole, the quality of this body of literature was reasonably high; 8 of the papers scored between 21-29 (Braham, Heasley & Akiens,2013; Kerfoot *et al*., 2012; Livingston, Nijdam-Jones & Brink, 2012; Osborn *et al*., 2010; Long, Langford, Clay, Craig & Hollin, 2011b; Gjerden & Moen, 2001; Oshima, Mino & Inomata, 2005; Oshima, Mino & Inomata, 2003), the remaining 20 scored 30 or above.

The item within the framework which consistently had the lowest scores across the papers was that of ethical considerations; with 12 of the papers scoring 0 (Kerfoot *et al*., 2012; Long *et al*., 2011b; Middelbøe Schjødt, Byrsting & Gjerris, 2001; Gjerden & Moen 2001;

Røssberg, Melle, Opjordsmoen & Friis, 2008; Røssberg & Friis 2004; Jansson & Eklund 2002; Schjødt, Middelbøe, Mortensen & Gjerris 2002; Oshima *et al*., 2003; Oshima *et al*., 2005; Beazley & Gudjonsson 2011; Ros, Van der Helm, Wissink, Stams & Schaftenaar 2013) and only three of the papers scoring 2 (Long *et al*., 2010a; Dickens *et al*., 2014; Jörgensen *et al*., 2008). The item on which this body of research scored the best overall was that of providing a rationale; all but Kerfoot *et al*. (2012) scored 2.

Author, date and country	Title	Focus	Quality Rating	Design. Statistic used	Sample size	Measure of ward environment	Main finding
Livingston, Nijdam-Jones, Lapsley, Calderwood & Brink. 2013 Canada	Supporting Recovery by Improving Patient Engagement in a Forensic Mental Health Hospital: Results From a Demonstration Project.	Service Characteristics	34	Longitudinal. T-test.	Patient: 25 Staff: 125	EssenCES.	No Change reported in EssenCES scores.
Long, Anagnostakis, Fox, Silaule, Somers, West & Webster 2011a UK	Social Climate Along the Pathway of Care in Women's Secure Mental Health Service: Variation With Level of Security, Patient Motivation, Therapeutic Alliance & Level of Disturbance.	Service Characteristics	34	Cross-Sectional. Mann Whitney U. Correlation.	Patient: 65 Staff: 80	EssenCES.	Positive social climate associated with lower levels of security & behavioural disturbances & higher levels of motivation, treatment engagement & therapeutic alliance
Braham, Heasley & Akiens 2013 UK	An Evaluation of Night Confinement in a High Secure Hospital.	Service Characteristics	26	Longitudinal. T-test.	Patient: 31/33 Staff: 84/65	EssenCES.	Patients' & staff perceptions of ward atmosphere did not differ in any domain pre & post night confinement
Kerfoot, Bamford,& Jones 2012 UK	Evaluation of Psychological Provision into an Acute Inpatient Unit	Service Characteristics	21	Longitudinal. None.	Various, unspecified.	EssenCES.	Positive trend in patient experience of the ward atmosphere after the introduction of psychological services.
Livingston, Nijdam-Jones & Brink 2012 UK	A Tale of Two Cultures: Examining Patient-Centred Care in	Service Characteristics	28	Cross-sectional. Correlation.	Patient: 30 Staff: 28	EssenCES.	Positive correlation between perceptions of ward atmosphere and

	Forensic Mental Health Hospital						Recovery orientated care.
Osborn, Lloyd-Evans, Johnson, Gilburt, Byford, Leese and Slade 2010 UK	Residential Alternatives to Acute In-patient Care in England: Satisfaction, Ward Atmosphere & Service User Experience.	Service Characteristics	29	Cross-sectional. Correlation. Regression.	In-patient: 125 Community: 102	WAS.	Alternative units reported greater autonomy, greater support & less anger & aggression on the WAS
Nesset, Røssberg, Almvik & Friis 2009 Norway.	Can a Focused Staff Training Programme Improve the Ward Atmosphere & Patient Satisfaction in a Forensic Psychiatric Hospital? A Pilot Study.	Service Characteristics	33	Longitudinal. Z-Scores.	Patients: total = 22 (T1, 10,T2, 8, T3, 8) Staff: total = unknown (T1 49, T2 48, T3 50)	WAS.	Change in the desired direction from staff & patients across subscales measured.
Long, Langford, Clay, Craig & Hollin 2011b UK	Architectural Change & the Effects on the Perceptions of the Ward atmosphere in a Medium Secure Unit for Women.	Service Characteristics	27	Longitudinal. Wilcoxon.	Patients: 9 Staff: 16	WAS.	No significant difference in WAS ratings in new unit for staff or patients.
Middelbøe, Schjodt Byrsting & Gjerris 2001 Denmark	Ward Atmosphere in Acute Psychiatric In-patient Care: Perceptions, Ideals & Satisfaction.	Satisfaction.	30	Cross-sectional. Correlation.	101 (patients).	WAS (Real and Ideal).	Order, organisation & support were significant predictors of satisfaction. Relationship & system maintenance dimensions of the WAS predictive of satisfaction. Perceived gap between ideal & real perception of WA explained 45% of variance in satisfaction.

Kuosmanen, Hatonen, Jykinen, Katajisto & Valimaki 2006 Finland	Patient Satisfaction with Psychiatric Inpatient Care.	Satisfaction.	32	Longitudinal. ANOVA.	313 (patients)	SPRI	Patients with a hospital stay of over three months were more satisfied with the WA & physical milieu than patients of a stay of 1-3 months.
Bressington, Stewart, Beer, & MacInnes 2011 UK	Levels of Service User Satisfaction in Secure Settings- A survey of the Association Between Perceived Social Climate, Perceived Therapeutic Relationship & Satisfaction with Forensic Services.	Satisfaction.	30	Cross-sectional. Correlation.	44 (patients)	EssenCES.	Patients perceptions of the WA were found to have a significant relationship with patients' satisfaction with care. The variable with the strongest association with satisfaction were patients' perceptions about the nature of the therapeutic relationship with staff.
Gjerden & Moen 2001 Norway	Patient Satisfaction & Ward Atmosphere During a Crisis in Open Psychiatric Ward.	Satisfaction.	28	Longitudinal. Chi-Square.	40 (T1:11, T2:10, T3:9, T4,9) (patient)	WAS	WAS was markedly lower on 5 of the 10 sub-scales during the period of crisis.
Røssberg, Melle, Opjordsmoen & Friis 2008 Norway	The Relationship Between Staff Members' Working Conditions & Patients' Perceptions of the Environment	Satisfaction.	32	Longitudinal. Correlation.	Patient: 129 Staff: 359	WAS.	Staff satisfaction with work was strongly, positively correlated with patients' perceptions of WA on the domain practical orientation. Staff satisfaction with work was strongly negatively correlated with patients' perceptions of WA on the domain staff control.

Røssberg, Melle, Opjordsmoen & Friis 2006 Norway	Patient Satisfaction & Treatment Environment: A 20-year Follow-up Study From an Acute Psychiatric Ward.	Satisfaction.	33	Longitudinal. Correlation.	129 (Patient)	WAS.	Significant relationship between satisfaction & WAS subscales; Involvement, Practical orientation, Anger & aggression, Staff control.
Jörgensen, Römma & Rundmo. 2009 Norway	Associations Between Ward Atmosphere, Patient Satisfaction & Outcome.	Satisfaction (and outcome)	35	Longitudinal. MANOVA.	80	WAS.	Differences in perceived ward atmosphere associated with perceptions of patient satisfaction. Mixed association between ward atmosphere and outcome.
Røssberg & Friis 2004 Norway	Patients' & Staff Perception of the Psychiatric Ward atmosphere.	Patients and staff members' perception of ward atmosphere.	29	Longitudinal. T-Test.	Patient: 424 Staff: 640	WAS.	Staff members had sig higher scores than patients on 9/11 WAS sub-scales. Patients & staff WAS scores were moderately correlated. Patient WAS scores were strongly correlated with patient satisfaction.
Brunt & Rask 2007 Sweden	Patient & Staff Perceptions of the Ward Atmosphere in a Swedish Maximum-security Forensic Psychiatric Hospital.	Patients and staff members' perception of ward atmosphere.	31	Cross-Sectional. Mann Whitney U.	Patient: 35 Staff: 108	WAS.	The patient & staff perceptions of WA differed to a sig level on 8/10 WAS sub-scales.
Jansson & Eklund 2002a Norway	Stability of Perceived Ward Atmosphere Over Time, Diagnosis &	Patients and staff members' perception of	30	Longitudinal. Mann Whitney U.	Patient: 51 (119 questionnaires)	COPES.	Ward atmosphere was fairly stable over time. Few differences between schizophrenia & patients

	Gender for Patients with Psychosis.	ward atmosphere.			Staff: 38 (101 questionnaires)		with other psychosis. No difference in perceived ward atmosphere with respect to gender. Difference in variation between staff and patients' ratings of subscales changed over time
Schjødt, Middelboe, Mortensen & Gjerris. 2003 Denmark	Ward Atmosphere in Acute Psychiatric In-patient Care: Differences & Similarities Between Patient & Staff Perceptions.	Patients and staff members' perception of ward atmosphere.	30	Cross-sectional. T-tests.	Patient: 96 Staff: 66	WAS (Real and Ideal).	WAS R: Staff scored sig higher than patients on 5 of the subscales. The main profiles of patient & staff ratings were similar. WAS I: Staff scored sig higher than patients on 3 of the subscales. Patients scored higher than staff on a further 3 subscales. Main profile was relatively similar.
Jansson & Eklund 2002b Sweden	How the Inner World is Reflected in Relation to Perceived Ward Atmosphere Among Patients with Psychosis	Patient Characteristics	31	Cross-sectional. Regression	37	COPES.	Self-control, paranoid symptoms & social competencies implicated as impacting upon perceptions of ward atmosphere.
Campbell, Allan & Sims 2014 UK	Service Attachment: The Relative Contributions of Ward Climate Perceptions & Attachment Anxiety & Avoidance in Male	Patient Characteristics	30	Cross-sectional. Correlation.	76	EssenCES.	Perceptions of ward atmosphere was more strongly associated with service attachment than personal attachment style. The most important aspect of ward

	Inpatients with Psychosis.						atmosphere for service attachment was perceived Staff support.
Dickens, Suesse, Snyman & Picchioni 2014 UK	Associations Between Ward Climate & Patient Characteristics in a Secure Forensic Mental Health Service.	Patient Characteristics	35	Longitudinal. Regression.	63	EssenCES.	Gender & level of security predicted patient cohesion & experience safety. Level of security predicted therapeutic hold.
Brunt 2008 Sweden	The Ward Atmosphere of Single-Sex Wards in a Maximum-Security Forensic Psychiatric Hospital in Sweden.	Patient Characteristics	30	Cross-sectional. Mann Whitney U.	Patient: 35 Staff: 104	WAS.	No significant differences between wards based on gender. Female wards similar to Relationship Orientated and Insight Orientated programme. Males wards did not resemble any treatment programme.
Oshima, Mino & Inomata 2005 Japan	Effects of Environment Deprivation on Negative Symptoms of Schizophrenia: A Nationwide Survey in Japan's Psychiatric Hospitals.	Symptoms.	28	Cross-sectional. Correlation.	549 (patient)	Personal Possession Scale. Nurses' opinion about Patient Scale. Patient Occupation Scale. Contact with outside world Scale. Ward Restrictiveness Scale.	Significant correlations reported between negative symptoms scale and most of the social environment scales.
Oshima, Mino & Inomata 2003 Japan	Institutionalisation & Schizophrenia in Japan: Social Environments & Negative Symptoms.	Symptoms.	27	Cross-sectional. Correlation.	2758	Personal Possession Scale.	Negative symptoms scales showed significant correlation with under stimulating

						Nurses' opinion about Patient Scale. Patient Occupation Scale. Contact with outside world Scale. Ward Restrictiveness Scale.	social environments in hospital.
Beazley & Gudjonsson 2011 UK	Motivating Inpatients to Engage with Treatment: The Role of Depression & Ward Atmosphere	Symptoms.	31	Cross-sectional. Correlation.	60	WAS.	Depression predicts motivation in a relationship which is mediated by WA. WA predicts motivation in a relationship that is mediated by depression
Johansson & Eklund. 2004 Sweden	Helping Alliance & Ward Atmosphere in Psychiatric In-patient Care	Helping alliance (other)	32	Longitudinal. Correlation.	61	COPES.	9 of the 10 COPES subscales showed significant correlations with helping alliance. Multivariate analysis indicates Support was the most important subscale to helping alliance.
Ros, Van der Helm, Wissink, Stams & Schaftenaar 2013 Holland	Institutional Climate & Aggression in a Secure Psychiatric Setting	Aggression (other)	32	Longitudinal. Correlation.	72	PGCI-SF	Support was negatively associated with aggressive incidents.

Table (i) Details of the papers within the Systematic Review.

Key

Full title of measure of ward atmosphere	**Abbreviation**
Ward Atmosphere Scale	WAS
Essen Climate Evaluation Schema	EssenCES
Community Oriented Programs Environment Scale	COPES
Prison Group Climate Inventory-Short Form	PGCSI-SF
Questionnaire developed by the Swedish Institute for Health Services Development (no further title provided).	SPRI

2.4.2 Settings

This body of literature includes studies from eight countries; nine studies were carried out in the UK (Long *et al*., 2010a; Braham, *et al*., 2013; Kerfoot, *et al*., 2012; Osborn *et al*., 2010; Long *et al*., 2011b; Bressington, Stewart, Beer & MacInnes, 2011; Campbell, Allan & Sims, 2014; Dickens *et al*., 2014; Beazley & Gudjonsson, 2011), six in Norway (Nesset, *et al*., 2008; Gjerden & Moen 2001 ; Røssberg *et al*., 2008; Røssberg Melle, Opjordsmoen & Friis 2006; Røssberg & Friis 2004; Jörgensen *et al*., 2009), five in Sweden (Brunt, 2008; Brunt & Rask 2005; Jansson & Eklund 2002a; Jansson & Eklund 2002b; Johansson & Eklund 2004). Two studies were carried out in Canada (Livingston *et al*., 2013; Livingston *et al*., 2012), Denmark (Middelbøe *et al*., 2001; Schjødt *et al*., 2002) and Japan (Oshima *et al*., 2003; Oshima *et al*., 2005) and one study was carried out in both Holland (Ros *et al*., 2013) and Finland (Kuosmanen, Hätönen, Jyrkinen, Katajisto & Välimäki, 2006).

As per the inclusion criteria, all of the studies involved inpatient mental health services. Eight studies (Livingston, *et al*., 2013; Livingston, *et al*., 2013; Livingston *et al*., 2012; Nesset, *et al*., 2008; Bressington *et al*., 2011; Brunt & Rask 2005; Campbell *et al*., 2014 Ros *et al*., 2013) used a Forensic mental health setting, the remaining 20 papers used a Rehabilitation and/or Recovery setting. Six of the studies compared settings; five of which did so within the same service (Brunt, 2008; Long *et al*., 2010a; Braham, *et al*., 2013; Dickens *et al*., 2014; Jörgensen *et al*., 2009) and one (Osborn *et al*., 2010) traditional in-patient services with alternative community-based services. Osborn *et al*., (2010) was also the only paper to use a comparable community sample.

2.4.3 Design

13 of the studies employed cross-sectional designs (Brunt, 2008; Long *et al*., 2010a; Livingston *et al*., 2012; Osborn *et al*., 2010; Long *et al*., 2011b; Bressington *et al*., 2011; Brunt & Rask 2005; Schjødt *et al*., 2002; Jansson & Eklund 2002b; Campbell *et al*., 2014; Oshima *et al*., 2005; Oshima *et al*., 2003; Beazley & Gudjonsson 2011). The remaining 15 were longitudinal.

2.4.4 Sample

2.4.4.1 Sample Size

The sample size in this body of literature ranged between n = 25 in Long *et al*., (2011b) and n = 2,758 in Oshima *et al*., (2003).

2.4.4.2 Population used

13 of the studies' populations were made up of patients and members of staff (Livingston, *et al*., 2013; Brunt, 2008; Long *et al*., 2010a; Braham, *et al*., 2013; Kerfoot, *et al*., 2012; Livingston *et al*., 2012; Nesset, *et al*., 2008; Long *et al*., 2011b; Røssberg *et al*., 2008; Røssberg & Friis 2004; Brunt & Rask 2005; Jansson & Eklund 2002; Schjødt *et al*., 2002). The remaining 15 of the papers' populations were patients only. In both Oshima *et al*., (2003; 2005) studies, the population identified was the patients; however, the data gathered was third-party reports of the patients.

2.4.4.3 Representativeness of samples

20 of the papers provided information on what percentage of potential participants participated. Kerfoot, *et al*., (2012), Livingston *et al*., (2012), Osborn *et al*., (2010), Gjerden and Moen (2001), Røssberg *et al*., (2006), Røssberg *et al*., (2008), Jansson and Eklund (2002) and Dickens *et al*., (2014) did not. Of the 20 papers which did, the range of potential participants who took part varied from 30% in Livingston, *et al*., (2013) to 100% of potential participants taking part in Long *et al*., (2011b).

2.4.4.4 Inclusion and Exclusion criteria

10 of the studies (Braham, *et al*., 2013; Long *et al*., 2010a; Kerfoot, *et al*., 2012; Nesset, *et al*., 2008; Røssberg *et al*., 2006; Røssberg *et al*., 2008; Røssberg & Friis 2004; Jansson & Eklund 2002a; Jansson & Eklund 2002b; Ros *et al*., 2013) stated their criteria was that an individual should be part of the population they were interested in (most commonly being an inpatient and/or a member of staff in a ward where the research was being carried out) to take part in the study.

Røssberg *et al*., (2008), Røssberg *et al*., (2006) Røssberg & Friis (2004), Jansson and Eklund (2002a), Jansson and Eklund (2002b), Kerfoot, *et al*., (2012), Nesset, *et al*., (2008), Ros *et al*., (2013) did not provide any inclusion/exclusion criteria. The remaining 10 studies had more in-depth criteria, including language skills (Brunt, 2008; Brunt & Rask 2005; Campbell *et al*., 2014; Johansson & Eklund 2004; Livingston, *et al*., 2013; Livingston *et al*., 2012), diagnosis (Campbell *et al*., 2014; Oshima *et al*., 2003; Oshima *et al*., 2005) and stage of treatment (Gjerden & Moen 2001; Kerfoot, *et al*., 2012; Osborn *et al*., 2010).

2.4.5 Measures employed

Across this body of literature, the most common form of data was self-report measures (questionnaires). The exception being Oshima *et al*., (2003:2005) who used third party reports.

2.4.5.1 Measures of ward atmosphere

All of the papers used questionnaires to assess the ward atmosphere. 25 of the studies used a single measure, which was specifically designed to assess ward atmosphere; the most commonly used measure (n = 13) was a version of the Ward Atmosphere Scale (WAS), (Brunt, 2008; Osborn *et al*., 2010; Nesset, *et al*., 2008; Long *et al*., 2011b; Middelbøe *et al*., 2001; Gjerden & Moen 2001; Røssberg *et al*., 2008; Røssberg *et al*., 2006; Røssberg & Friis 2004; Brunt & Rask 2005; Schjødt *et al*., 2002; Beazley & Gudjonsson 2011; Jörgensen *et al*., 2009), 8 of the papers used the Essen Climate Evaluation Schema (EssenCES) (Braham, *et al*., 2013; Bressington *et al*., 2011; Campbell *et al*., 2014; Dickens *et al*., 2014; Livingston *et al*., 2012; Livingston, *et al*., 2013; Long *et al*., 2010a; Kerfoot, *et al*., 2012;), three studies used the Community Oriented Programs Environment Scale (COPES) (Jansson & Eklund 2002a; Jansson & Eklund 2002b; Johansson & Eklund 2004), and one paper used the Prison Group Climate Inventory-Short Form (PGCI-SF) (Ros *et al*., 2013). Two studies used five separate questionnaires to measure various aspects of the ward atmosphere, (Oshima *et al*., 2003; Oshima *et al*., 2005) and one paper assessed ward atmosphere within a standardised patient satisfaction questionnaire, developed by the Swedish Institute for Health Services Development (SPRI) (Kuosmanen *et al*., 2006).

2.4.5.2 Other measures used

Three of the studies only used the measures of ward atmosphere (Kuosmanen *et al*., 2006; Jansson & Eklund 2002a; Schjødt *et al*., 2002). 23 of the studies used an additional questionnaire, the exceptions being Kuosmanen *et al*., (2006), Jansson and Eklund (2002a) Schjødt *et al*., (2002) Dickens *et al*., (2014) and Ros *et al*., (2013). Six of the studies reviewed patient records for information (such as diagnosis and daily dose of medication) (Brunt, 2008; Long *et al*., 2010a; Osborn *et al*., 2010; Middelbøe *et al*., 2001; Dickens *et al*., 2014; Oshima *et al*., 2005). Four used semi-structured interviews (Braham, *et al*., 2013; Livingston *et al*., 2012; Livingston, *et al*., 2013; Long *et al*., 2011b), four recorded service level data (Dickens *et al*., 2014; Kerfoot, *et al*., 2012; Long *et al*., 2011b; Ros *et al*., 2013) and one study used cognitive assessments (Jansson & Eklund, 2002b).

2.5 Critique

This section will offer an overview of issues spanning numerous studies within this review, which may bias or impact upon the validity of the reported findings. These limitations are intended to arm the reader with regards to the weight that is given to the evidence present in findings and synthesis.

2.5.1 Quality

As discussed, as a body of work the quality was reasonably high, and the score range narrow. When comparing the relative focus, *Service Characteristics* and *Symptoms* scored at the lower end; differences between patients and staff members' perceptions of ward atmosphere was in the mid-range; satisfaction, patient characteristics and the papers within the other focus were relatively high. Having said this, the overall difference was not great, the means for each focus ranged from 28.6 for Symptoms, to 32 for Other. Furthermore, there was greater variation between papers within each focus, than between the focuses as a whole. Therefore, to best employ the quality ratings, these will be further explored within the results and synthesis.

2.5.2 Missing Information

Across this body of research, a reoccurring limitation is that certain information is not provided. This was found across a number of areas, including information on the population used and service level factors. This is of particular impact in this body of research, as peripheral factors are likely to have impacted upon ward atmosphere.

2.5.2.1 Environmental Factors

Of the 15 longitudinal studies, only Røssberg *et al*., (2006), provided information on environmental factors, such as crisis in services, in relation to the ward during the times measures were taken. Due to the lack of explanation in the other 14 studies, it is not possible to know whether the findings from these may be confounded by factors outside of the study's remit, in terms of what was causing the change in ward atmosphere.

2.5.2.2 Ward Information

While all of the studies provided basic information about the wards in which they took place (e.g. forensic, low/medium secure, public or private etc.) with the exception of Gjerden and Moen (2001), Jansson and Eklund (2002a;2002b), Jörgensen *et al*., (2009), Osborn *et al*. (2010), Oshima *et al*., (3002;2005) and Schjødt *et al*., (2003), studies did not provide relevant details of the ethos of the wards selected. Arguably many wards may not have a known ethos, but models used or characteristics of the ward would have been helpful in allowing the findings to be more generalisable.

2.5.2.3 Sampling Information

Kerfoot, *et al*., (2012), Røssberg *et al*., (2006), Røssberg *et al*., (2008) and Røssberg & Friis (2004) did not report basic demographic information (age and gender) on the sample used. Therefore, these studies cannot consider demographic factors in relation to ward atmosphere, nor can the representativeness of the sample be evaluated, which makes it difficult to generalise.

The majority of the studies did not provide information on how their participating population related to possible participants, meaning the generalisability was harder to estimate. This is discussed further when sample bias is examined in 2.5.4.2.

2.5.4 Sample Bias

Across this body of research there are a number of methodological issues, which may have biased the sample.

2.5.4.1 Inclusion and Exclusion Criteria

Such criteria are a usual and necessary part of research. However, in this body of research particular criteria may have biased the sample with regards to the type of person who was eligible to take part. This is an important factor when considering the representativeness of the sample used, and how generalisable and robust certain findings are.

The nine studies which used language-exclusion criteria, while pragmatically understandable, may have biased their samples in terms of representativeness of included patients. Kuosmanen *et al*., (2006), required possible participants to be able to complete the questionnaire unaided, which would have biased the sample towards better educated, healthier, native individuals. Middelbøe *et al*., (2001) excluded patients with a dementia or who were too acutely disturbed to be included; they did not state that this was based on the ability to consent, rather having the diagnosis or being deemed ‘too unwell’. This lowers the general age and biases the sample towards better functioning patients. Similarly, Livingston, *et al*., (2013; 2012), Osborn *et al*., (2010), Bressington *et al*., (2011), Beazley & Gudjonsson (2011) and Johansson & Eklund (2004) excluded people, based on staff assessment, however, clear limits of what this entailed were not provided. Ethically, if an individual is unable to consent, this is necessary, however these studies did not state they were asking about individuals’ ability to consent, only if someone was able. The information on exactly what staff were asked is vague and may have meant staff selected people they felt were suitable, or more amenable – the sample may have been biased towards those with a better relationship with staff and therefore may perceive the ward atmosphere more favourably. However, it should be noted that the instructions given to staff were not clearly described, and staff may have been asked about ability to consent only.

2.5.4.2 Representativeness of the Sample

For the findings to be generalisable, the samples used need to reflect the population of interest. Branham *et al.* (2013) reported that 40-45% of possible participants taking part in research which is carried out inpatients services, is average. However, that does not mean any study with such a proportion has a representative sample. As described, all but eight of the studies provided what proportion of possible participants they recruited. As mentioned in section 2.4.4.3, only seven of the studies provided information on how those who did take part compared to those who did not. Jörgensen *et al.,* (2009) completed a comparison analysis to assess if the sample gained was statistically different from the population in question. Both Jörgensen *et al.,* (2009) and Nesset, *et al.,* (2008) stated that the sample did not differ from those who did not take part in terms of age, length of stay and gender. Schjødt *et al.,* (2002) and Middelbøe *et al*., (2001) both reported significant differences in terms of age, with participants being younger and a higher proportion of participants having affective-spectrum disorder and a lower proportion of organic disease, than possible participants who did not take part. The studies of Bressington *et al.,* (2011) and Oshima *et al.*, (2003; 2005), compared their sample to national data on the populations they were using. Bressington *et al.,* (2011) reported their sample was in line with national data other than a higher proportion of woman in the study. Oshima (2003; 2005) reported that in their samples the hospitals had a higher number of beds and a higher number of open wards than national averages. Such differences may impact upon these studies' findings, however on the whole the differences are limited. The other 21 studies did not provide such information.

Livingston *et al.,* (2012), Braham *et al.,* (2013) Bressington *et al.,* (2011), Jansson and Eklund (2002b) and Beazley and Gudjonsson (2011) Nesset *et al*., (2008) all had male dominated samples. None of the studies reported that this was representative of the population the sample was taken from (although it may have been). The relationship between gender and ward atmosphere is unclear, however it may bias the findings in these studies.

Osborn *et al.,* (2010) reported a statistically significant difference in ethnicity of patients between the units (traditional vs alternative) in their research. Furthermore, many more of the sample in alternate units were voluntary compared to those in the traditional units, who were more likely to have been detained under the Mental Health Act. These factors were described, which is of use as it allows the reader to question the relationships between the reported differences in ward atmosphere and service characteristics (as stated by the paper), or the possible influence of patient characteristics.

2.5.5 Uncontrolled Variables

Due to the broad nature of ward atmosphere, factors which studies were not assessing may have impacted upon the findings. These may be important to bear in mind when reviewing what the authors claim have influenced findings. Gjerden and Moen (2001) reported that during crisis periods, patients who would have usually been on another ward due to being more unwell, were on the ward being assessed. This was a clinical necessity and could not have been controlled for, however when reviewing the findings, a possible cause of the change in perceived ward atmosphere could have been characteristics of the patients who were on said ward during the period. The methodology of Osborn *et al*., (2010) altered between the alternative and standard units they were researching; in alternate units questionnaires were completed over 57 days, in traditional units it was done in 10 days. The impact of this was not explored, although it may be of consequence, as evidence from this review supports that perceived ward atmosphere is not consistent over time.

2.5.6 Diversity of Research in the Review

In addition to thinking about the studies at an individual level, considering what this review can offer to the understanding of ward atmosphere in inpatient mental health settings in general, the diversity and how representative this review is also noteworthy. As discussed, the research came from a total of eight countries. However, when considering the focus all of the patient/staff-focused studies were Scandinavian, as were five of the six whose focus was satisfaction. This may make the generalisability of the findings for these focuses questionable.

It was necessary for papers to be published in English for pragmatic reasons. However, it is worth noting this may introduce bias, as statistically significant positive results are more likely to be reported internationally (Egger, 1997).

2.5.6.1 Diversity of Setting in the Review

As per the selection criteria for the review, all of the studies involved some type of inpatient setting. However, within these limits the type of wards selected for the research to be carried out in may have impacted upon how generalisable the findings were. This is of note in relation to Jansson and Eklund's (2002a & 2002b) studies, which were completed in a Psychiatric Rehabilitation Unit, described as distinctly different from traditional rehabilitation units. Two of the studies actively addressed this issue: Jörgensen *et al*., (2009) and Schjødt *et al*., (2003) selected wards without a committed philosophy of milieu therapy, to increases generalisability.

2.5.7 Measures

2.5.7.1 Measures Used

This body of research relies heavily on self-report questionnaires; all 28 studies employed at least one questionnaire. Brunt (2008), Campbell *et al*., (2014), Jorgensen *et al*., (2009), Middelbøe (2001), Osborn *et al*., (2010), Røssberg *et al*., (2005) and Schjødt *et al*., (2003) only used self-report measures to collect all of their data. There are some well-documented limitations to self-reporting: they are relatively subjective measures and rely on participants interpreting questions consistently, honestly and as intended.

2.5.7.2 Accuracy of Construct Measured

In the papers looking at symptoms, in all three cases there is some ambiguity as to what was being measured. In both Oshima *et al*., studies (2003; 2005) all the data was collected about patients from staff members; therefore, it is important that the findings are viewed as staff perceptions of symptoms and ward atmosphere and not as that of the participant (the patient) they relate to.

In the work by Beazley and Gudjonsson (2011) investigating depression, a large percentage of their sample (80%) had a diagnosis of a Schizophrenia-spectrum disorder. Therefore, it is debatable whether Beazley and Gudjonsson measured depression, or undetected negative symptoms of schizophrenia.

2.5.7.3 Choice of Measures

The choices of measures used for three of the studies (Dickens *et al*., 2014; Ros *et al*., 2013; Røssberg *et al*., 2008) may have biased their findings.

Dickens *et al*., (2014), used a retrospective method to collect data, much of which was routinely collected rather than for the study itself. As such, information may have been missed, as the data was not being collected to suit the needs of the study. The definition of treatment engagement was arguably weak – as leave from the ward did not count as 'treatment'. Justification for this is that retrospective records did not provide enough information regarding purpose of leave. However, for patients who were reconnecting with the outside world this may have been an important aspect of their treatment.

In Ros *et al.,* (2013), aggressive incidences were defined by staff. Incidences which patients experienced as aggressive but staff did not, were not considered. The relationship they reported may just be between staff perceptions of aggression and types of ward atmosphere.

Finally, Røssberg *et al.,* (2008), used a modified WAS, which only included six subscales; this could have biased their findings, as the potential role which the other subscales may have played (as in other papers, which used all ten subscales) could not have been supported by this paper, even if the relationship was there to be found.

2.5.7.4 When measures were taken

Braham *et al.,* (2013), Livingston *et al.,* (2013) and Long *et al.,* (2011b) looked at a change and its impact on ward atmosphere. The reported results may have been impacted by how long after the change the measures were taken. This is highlighted by Long *et al.,* (2011b) who reported a relationship between a change in physical environment and levels of satisfaction, symptomatology, risk behaviours, but not ward atmosphere. Yet other studies indicate a relationship with ward atmosphere and satisfaction and symptomatology, which might lead one to expect a change in these facts to predict a change in ward atmosphere. It may be perceptions of ward atmosphere take longer to change in light of improvements than satisfaction and symptomatology, and as such had the measure of ward atmosphere been taken at a later stage, a difference might have been found.

2.5.8 Analysis

Across this body of research there is limited justification for the choice of statistical procedure or evidence of its robustness. While the choices hold face validity, only two of the studies provided a rationale or justification for their statistical choices; Jörgensen *et al.,* (2009) commented that correlations would have been preferable, but due to the number of factors and number of participants they decided against this and performed MANOVA. While information may have been missed, the study appears more robust for it. And Ros *et al.,* (2013), rationalised the use of a high number of factors in a correlation in the first stage of analysis by attempting to minimise the risk of a type I error by only using factors correlating to a degree of $p < .01$ in stage two. The only study which conducted a power analysis was Dickenson *et al.,* (2014).

2.6 Findings and Synthesis

In continuation from the limitations to this body of research, the findings and a synthesis of such will now be presented, in an attempt to consider the weight of the evidence with regards to how it answers the review question. Due to the broad nature of the review question and the subsequent variety in the nature and focus of the papers, the evidence from this body of research will be reviewed under each of the areas outlined above. However, this is a single review which aims to explore the role of ward atmosphere in relation to how it is perceived in inpatient services and how it relates to indices of quality of care in-patient services; therefore, evidence from across the body of research will be used where applicable to help develop a holistic understanding of the evidence available.

2.6.1 Service Characteristics

Eight of the papers, in this review, considered the relationship between ward atmosphere and service characteristics. Long *et al.,* (2011a) and Osborn *et al,* (2010) compared ward atmospheres between services. The remaining six papers considered the relationship between a particular service characteristic and ward atmosphere, these being: the relationship between recovery-orientated care and ward atmosphere (Livingston *et al*., 2012); the impact of a change within a service (Braham *et al.,* 2013; Long *et al.,* 2011b; Kerfoot *et al.,* 2012) and the impact of interventions (aimed at staff members and patients respectively) (Nesset, *et al.,* 2008; Livingston *et al.,* 2013).

The quality of these particular studies is more diverse then the other focuses, ranging from 21 (Kerfoot *et al.,* 2012) which was the lowest score in the review, to 34 (Long *et al.,* 2011a: Nesset *et al.,* 2008) which are among the highest. This is of particular note with regards to the findings reported by Braham *et al.,* (2013), Long *et al.,* (2011b) and Kerfoot *et al.,* (2012), as the weight given to these studies is limited, compared to the remaining five, due to their quality ratings. The remaining five papers scores were narrow and high (29-34), consistent with the general scores in this review.

Five papers indicated some differences in ward atmosphere in relation to service characteristics (Kerfoot *et al.,* 2012; Livingston *et al*., 2012; Long *et al.,* 2011a; Nesset, *et al.*, 2008; Osborn *et al,* 2010), while three reported no relationship.

At the broadest level, all five papers which reported a relationship supported the premise that differences in services result in differences in ward atmosphere. This finding is particularly highlighted by Long *et al* (2011a) and Osborn *et al* (2010) who compared ward atmosphere between services which were designed to differ from one another. Long *et al* (2011a) compared four types of ward within one service (medium-secure (MS) admissions, medium-secure (MS) treatment, low-secure recovery ward and a low-secure and Borderline Personality disorder wards) and reported patient scores of ward atmosphere were significantly lower on all domains on the EssenCES from participants on the MS admission wards, compared to the other wards. This finding may reflect that the wards are designed to support people in different stages of recovery, and that the psychological well-being of participants caused the difference seen, as the paper implies. However, staff members also scored the MS admission ward lower across the three domains and total EssenCES, suggesting differences on the EssenCES was in part due to the ward atmosphere and not only reflective of patients' perception as influenced by their well-being. Osborn *et al* (2010) stated that different ward atmospheres are found depending on the service. They compared traditional inpatient units and community alternatives. Using the WAS, they reported significant differences between wards (based upon the aforementioned criteria) on five subscales; Involvement, Autonomy, Anger and Aggression, Order and Organisation and Programme Clarity, with the alternative units scoring higher on all but Anger and Aggression. A regression model indicated the subscale Anger and Aggression as significantly associated with the type of unit. This relationship was found to hold, when adjusted for patient characteristics. The separation of patient and service characteristic is of interest, as the evidence from this review as a whole indicates that patient characteristics influence perceptions of ward atmosphere. Long *et al* (2011a) reported that Experiences Safety (ES) on the EssenCES was different between wards. This supports the reported significance of the Anger and Aggression subscale to a degree; however, Long did not report ES as having the same pivotal role as Osborn suggested Anger and Aggression did. While the broad findings of these studies are similar, this highlights the differences found in which aspects of ward atmosphere were reported to vary between services, as is true of the other studies in this focus. This is not unexpected given the large differences in what the studies looked at and methodologies and measures used. Furthermore, this difference is consistent with the overall findings from this review; that is, how ward atmosphere impacts upon, or is impacted upon, depends on the nuances which are examined or manipulated.

Livingston *et al*., (2012), Kerfoot *et al*., (2012) and Nesset *et al*., (2008) each looked at a more specific aspect of the service (recovery orientated care, the impact of psychological services and staff training, respectively). While they provided less detail on which subscales or domains of ward atmosphere changed than the previous two studies, focusing on a narrower aspect of service characteristics and their relationship to ward atmosphere allows greater explanation as to how service characteristics might be able to impact upon ward atmosphere.

The reported findings of Livingston *et al*., (2012) indicated both patients and staff who perceived the hospital as being more consistent with recovery principles, had a more positive view of the ward atmosphere; with significant, moderate correlation between perceptions of recovery-orientated care and mean rating of EssenCES scores, for both patients and staff. Kerfoot *et al*., (2012) reported an increase on all three of the dimensions on the EssenCES over three time points, after the introduction of psychology to a service. However, no statistical analysis was performed, and the quality of this study was low. Nesset *et al*., (2008) reported a positive impact on five of the six WAS subscales measured, as reported by patients, and improvement on all six by staff members, after a staff training intervention. In both cases a change in the desired direction in that number of subscales is significant. This indicated that staff training can influence both patients and staff member's perceptions of ward atmosphere.

In contrast to the aforementioned papers, three of the papers found no relationship between service characteristics and ward atmosphere (Braham *et al*., 2013; Livingston, *et al*., 2013; Long *et al*., 2011b). Braham *et al*., (2013) looked at impact of the introduction of night confinement, Long *et al*. (2011b), the impact of a change in physical environment and Livingston *et al*., (2013), the impact of interventions aimed at improving patient quality of life. These studies contradict the findings of the other five studies in this focus. However, how this impacts upon the overall findings from this focus is debatable. Of the three studies, Braham *et al*., (2013) and Long *et al*., (2011b) are at the lower end of the quality ratings (26/27) lending less weight to their findings. As well their quality rating, the findings of both Braham *et al*., (2013) and Long *et al*., (2011b) may be a result of the time at which they collected their data. In both cases data was collected once before and once after the service change had occurred (see section 2.5.7.4 for discussion).

Having said this, Livingston, *et al*., (2013) was a high-quality study, and thus greater importance can be placed upon their findings, which indicate ward atmosphere is not influenced by changes to services. The conclusions drawn by Livingston *et al*., (2013) contradict Nesset *et al*., (2008) despite the studies looking at very similar aspects of service characteristics (i.e. the impact of service level intervention). Given that both are of a high quality (33 and 34 respectively), it seems likely that the difference in findings was due to the differences in interventions provided; either that staff training (Nesset *et al*., 2008) has a greater impact or a stronger link to the ward atmosphere than interventions aimed at patient participation (Livingston *et al*., 2013), or that the Nesset invention was of higher quality.

2.6.2 Satisfaction

Seven of the articles in this review considered satisfaction in relation to ward atmosphere. Five looked at patient satisfaction (Bressington *et al*., 2011; Kuosmanen *et al*., 2006; Røssberg *et al*., 2006; Gjerden, & Moen 2001; Middelbøe et a., 2001), one looked at patient satisfaction and outcome (Jörgensen *et al*., 2009) and one considered patient and staff satisfaction, patient's perception of ward atmosphere and staff perceptions about working conditions (Røssberg *et al*., 2008). Six of the seven studies indicated that there was a relationship between ward atmosphere and satisfaction; Kuosmanen *et al*., (2006), alone did not indicate a relationship between satisfaction and ward atmosphere; the only reported significant finding was between length of stay and satisfaction with ward atmosphere. The study was rated as relatively good quality (32) however, the breadth of interests of the study was wide; only a very limited portion of which was related to the topic of this review. Indeed, a single ward atmosphere scale was not used (rather perceptions of the ward atmosphere were captured as an aspect of satisfaction with the service, in a general satisfaction questionnaire). As such, while the findings do not add to this body of work, due to the nature of the study the lack of findings do not challenge the findings of the other six papers. Furthermore, although not the focus of their studies, Long *et al*., (2011a), Osborn *et al*., (2010) and Røssberg & Friis (2004) also considered satisfaction and ward atmosphere to a degree; all of which also indicated a relationship between ward atmosphere and satisfaction.

The papers within this focus were scored between 28 - 35. When comparing the quality of this focus to the others within the review, it seems to have a wider range; Jörgensen *et al*., (2009) had the highest score in the review, and Gjerden & Moen (2001) scored 28, which was a lower score. Having said this, over half of the papers (57%) in this focus scored over 32, indicating that on the whole the quality of these papers is high.

This focus had the most consistent overall finding: that there is a relationship between perceived ward atmosphere and satisfaction. However, the relationship becomes less clear when this is examined in greater detail. Various subscales and dimensions of ward atmosphere were identified by different papers as having a relationship to satisfaction, and when looking at the subscales no consistent pattern emerged. This may be because satisfaction is multifactorial, as indicated by these papers. As such, broader patterns may be of more help. To do this, and to help impose a structure, the three WAS domains (Relationship, Personal Growth and Organisational {or Systems Maintenance}) will be used to help organise the information. Not all of the studies used the WAS, the evidence provided by these studies will be considered in relation to these WAS dimensions.

Middelbøe *et al*., (2001) reported that the Relationship and Organisational domains of the WAS are predictive of patient satisfaction. Subscales from these domains (or the equivalent from other measures) appear to be most consistently indicated as important.

From the Relationship dimension, all three of the subscales were indicated as having a relationship to satisfaction by at least three sources, furthermore, two papers reported a significant relationship between the Therapeutic Hold (TH) domain from the EssenCES and satisfaction. Support was associated with satisfaction by two of the satisfaction focused papers (Gjerden & Moen 2001: Middelbøe *et al*., 2001), as it was in two of the papers from other focuses (Røssberg & Friis 2004: Osborn *et al*., 2011). These findings were echoed by research using other measures of ward atmosphere; using the EssenCES, Bressington *et al*., (2011) reported a strong positive correlation between overall satisfaction and TH, which supports this claim. Long *et al*., (2011a) found wards which had statistically higher scores on TH, were also the wards with statistically higher levels of satisfaction. Three of the seven satisfaction papers reported a significant relationship between patients' scores on Involvement from the WAS and satisfaction (Gjerden & Moen 2001; Jörgensen *et al*., 2008; Røssberg *et al*., 2006), as did Røssberg and Friis (2004). The subscale Spontaneous Behaviour was indicated as having a significant relationship to support by Gjerden and Moen (2001), Røssberg and Friis (2004) and Jörgensen *et al*., (2009).

These reported findings indicate support, from more than one source, for each of the three subscales in the Relationship domain from the WAS. The relationship was also found to hold when using other measures, as demonstrated by the reported relationships between satisfaction and TH from the EssenCES satisfaction, which related to the Relationship domain.

From the Organisational domain, all three of the subscales are indicated as having a significant relationship to satisfaction, by at least one study. The subscales Order and Organisation and Staff Control were most consistently found to have a relationship across the board. Along with Support, Middelbøe *et al*., (2001), reported Order and Organisation as one of the two subscales which predicted satisfaction. Gjerden and Moen (2001), reported a decline in perceived Order and Organisation during a crisis period, when satisfaction also declined. Osborn *et al*., (2011), and Røssberg and Friis (2004), reported an association between this domain and satisfaction. Jörgensen *et al*., provided some support for this subscale; reporting a significant relationship between Order and Organisation and satisfaction, however of the eight subscales they found to have a significant relationship to satisfaction, this had the weakest. Evidence about Staff Control from this review indicates an inverse relationship, by which lower levels of Staff Control were associated with higher levels of satisfaction. Such a negative correlation with satisfaction was reported between patient score by Røssberg *et al*., (2006) and Røssberg and Friis (2004). Røssberg and Friis also reported a partially negative correlation with staff satisfaction and Staff Control. The relationship between negative satisfaction for both patients and staff members and Staff Control was further supported by the Røssberg *et al*., (2008), who reported patient scores of Staff Control and staff-rated satisfaction strongly negatively correlated. Gjerden & Moen (2001) Røssberg *et al*., (2008) Røssberg *et al*., (2006), Røssberg and Friis, (2004) and Jörgensen *et al*., (2009), all reported a significant relationship between the Practical Orientation subscale and patient satisfaction.

From the Personal Growth dimension, the subscale most frequently shown to have an association with satisfaction was Anger and Aggression, which, similar to Staff Control, was reported to negatively correlate with satisfaction. Røssberg *et al.,* (2006) and Røsseberg and Friis (2004), reported lower levels of Anger and Aggression associated with higher satisfaction. Osborn *et al.,* (2010), reported levels on this subscale as statistically higher in traditional units, where satisfaction was lower. Bressington *et al.,* (2011), reported a moderate correlation between satisfaction and ES. ES reflects aggressive tension or threats of violence, which indicated that patients who felt safer on the unit were more satisfied, also supporting the claim.

Overall this body of research identifies a relationship between perceived ward atmosphere and satisfaction. However, as has been demonstrated, the aspect or subscale of ward atmosphere indicated as relating to satisfaction varied. The variation between studies is also found across this body of work as a whole. Differences may be explained by variations in study design and what the studies were looking at. Due to the high quality of a number of the studies within this focus, the difference may also be explained by the fact that both satisfaction and perceived ward atmosphere are likely to be impacted upon by multiple factors, thus making all the results less generalisable.

2.6.3 Staff and Patients' Perceptions of Ward Atmosphere

Four of the papers primarily compared perceptions of ward atmosphere between patients and staff (Brunt & Rask, 2005; Jansson & Eklund, 2002a; Røssberg & Friis 2004; Schjødt *et al*., 2003). Dickens *et al*., (2014), Long *et al*., (2011a) and Livingston *et al*., (2011) also provided some evidence as to staff and patients' perceptions of ward atmosphere.

The quality of the papers within this focus was in the medium/high range for this review, ranging from 29 (Røssberg & Friis 2004) to 31 (Brunt & Rask, 2005). As with this review as a whole, the quality range was narrow and high.

The findings from this focus were relatively consistent; at the broadest level, all reported that staff and patients' perceptions of ward atmosphere differed to a degree. Furthermore, all four studies reported that staff tended to rate more subscales higher (i.e. indicating staff members rated the ward atmosphere as being more positive) as did the findings from research focused on other areas (Dickens *et al*., 2014; Long *et al*., 2011a; Livingston *et al*., 2011). Brunt & Rask, (2005) Røssberg and Friis (2004) and Schjødt *et al*., (2003) also indicated that there was a relationship between patient and staff perceptions. Schjødt *et al*., (2003) demonstrated this both graphically and statistically; Brunt and Rask also demonstrated this finding graphically and Røssberg and Friis did so statistically.

However, there were also some differences when reviewing the data at a broad level. Notably, there was a difference in how many subscales were reported to differ between staff and patients, this varied from 3/10 on two occasions in Jansson and Eklund, (2002a), to 10/11 in Røssberg and Friis (2004) study. One explanation for the difference in the degree to which patient/staff perceptions vary across studies may be the designs employed. In both the longitudinal studies (Jansson & Eklund 2002a; Røssberg & Friis 2004) staff and patient perceptions of ward atmosphere were reported to vary to some degree over time. Jansson and Eklund (2002a) used Eklund and Hansson (1996), categories of the COPES (low, medium, high) to categorise scores in their study. Using these divisions, they reported of the 10 factors, six remained stable over time on patient scores, and eight on staff scores. They also reported the relationship between patient and staff perceptions of ward atmosphere varied over time. Such alterations over time may explain differences in reported findings between studies, as the cross-sectional studies (Brunt & Rask 2005; Schjødt *et al*., 2003) offer a 'snap shot' only.

At a narrower level, there was further consistency across the papers as to which subscales patient and staff opinion differed significantly on: all reported a statistical difference between staff and patient scores on Involvement, Practical Orientation and Program Clarity, where staff scored higher, and on Staff Control, on which patients scored higher. The Anger and aggression subscale was the subscale which had the least consistent findings reported. Only two of the papers reported staff scored higher levels of Anger and Aggression at some points (Jansson & Eklund 2002a; Røssberg & Friis 2004). There was some further support for this finding from the wider review, Dickens *et al*., (2014), and Livingston *et al*., (2012), reported staff members' scores as significantly lower than patients' on Experienced Safety (ES) in the EssenCES; indicating staff perceived a greater threat of violence/aggression. However, Schjødt *et al*., (2003) findings indicated that staff members' ideal of Anger and Aggression was significantly higher than patients (as measured by the WAS-I).

2.6.4 Patient Characteristics

Brunt (2008), Dickens *et al*., (2014), Campbell *et al*., (2014) and Jansson and Eklund (2002b) focused on the relationship between aspects of patient characteristics and ward atmosphere. Brunt (2008) focused on gender, Campbell *et al*., (2014) considered personal attachment style, Dickens *et al*., (2014) whether and how clinical characteristics were associated with ward climate and Jansson and Eklund (2002b) looked atsocial functioning, diagnosis, personality traits, psychiatric symptoms and cognitive functioning, and how these influence perceptions of ward atmosphere. Of the studies which focused on other areas, Bressington *et al*., (2011), Jansson and Eklund (2002a), Middelboe *et al*., (2010) and Osborn *et al*., (2010), also considered patients' general demographic factors (age and gender) and ward atmosphere.

The quality of the papers within this focus was mid to high in comparison to the scores across the review. Dickens *et al*., (2014), was rated as 35, the highest score received in the review, none of the others scored below 30. This is also true of the additional studies, with the exception of Osborn *et al*., (2010) who scored 29. This indicated the quality of the evidence in this focus is good.

When considering what this focus adds to the understanding of ward atmosphere, it appears that as with service characteristics, which aspects of patients' characteristics is reviewed dictates if there is a relationship to ward atmosphere. As has been indicated, the characteristics of interest vary between studies, the most consistently considered being gender and age, which are discussed first.

Brunt (2008), Bressington *et al*., (2011), Dickens *et al*., (2014), Jansson and Eklund (2002a), Middelbøe *et al*., (2001) and Osborn *et al*., (2010) all considered the role of gender, the evidence for which is mixed. Both Brunt (2008) and Osborn *et al*., (2010) compared services and indicated a relationship between gender and ward atmosphere. Osborn *et al*., reported more positive perceptions of ward atmosphere in alternative vs tradition services, and significantly more females in alternative units, although this is only correlational. Using Moos (1989, as cited Brunt, 2008) conceptualisation of treatment settings, Brunt stated that types of ward atmospheres seen varied between wards which were separated based on gender; however only one subscale from the WAS significantly differed between wards. With both these studies, it is unclear if it is the service characteristics, the patient characteristics, or a combination of the two that have created the ward atmosphere. In contrast, Bressington *et al*., (2011) Jansson and Eklund (2002a) and Middelbøe *et al*., (2001) found no relationship between patient perceptions of ward atmosphere and gender. Dickenson *et al*., reported a significant proportion of variance in the Patient Cohesion (PC) and ES domains of the EssenCES were accounted for by gender, but that there was no relationship between Therapeutic Hold (TH) domain and gender. This indicates the relationship between gender and ward atmosphere as complex. This study was of high quality (35) which is worthy of note when considering the balance of the evidence. That the role of gender is unclear would account for both Dickensen's reported findings and the inconsistency between studies.

The relationship between age and ward atmosphere appears less controversial. All three studies which reported on the relationship (Bressington *et al*. 2011; Campbell *et al*., 2014; Middelbøe *et al*., 2001) stated that age was not associated with perceptions of ward atmosphere. However, when trying to evaluate the evidence across the review, Kuosmanen *et al*., (2006) and Jansson and Eklund (2004) reported a relationship between age and the factors they were considering (level of satisfaction and Helping Alliance respectively), which were also reported to have a relationship with ward atmosphere. This suggests age may have a relationship to ward atmosphere as moderated by other factors. Indeed, the evidence for both gender and age may indicate that the impact of patients' demographic characteristics are mediated or moderated by other factors in how they relate to ward atmosphere. A further consideration is the ethnicity of the studies; social constructs around ideas such as gender roles could also affect the impact of such factors.

As discussed, three of the four papers within this focus consider further aspects of patient characteristics. Campbell *et al*., (2014), using the EssenCES reported a positive correlation between service attachment and patients' perceptions of ward atmosphere; the more positive the perception of ward atmosphere, the stronger the service attachment. When examining the data in greater depth, Campbell *et al*., (2014) reported that Therapeutic Hold was the only domain independently related to service attachment. The authors argued these findings indicated the importance of patient perception of ward atmosphere and the staff/patient relationship, for patients' attachment to services. However, the importance of TH domain is not supported in relation to other aspects of patients' characteristics. Dickensen *et al*., (2014) reported an association with patient characteristics and ward atmosphere. Using regression models to explore the relationship between each of the EssenCES domains, they reported the model fit was significant for all three domains of the EssenCES and patient characteristics; however the model fit was considerably weaker for TH than the other domains. The only patient characteristic reported as constantly having a significant relationship across all domains was Level of Security. This is consistent with Long *et al.,* (2011a) findings, however, whether such reported findings are due to patient or service characteristics is unclear.

Jansson and Eklund (2002b) reported all patients' characteristics they measured related to one of five of the COPES 10 subscale (Support, Practical Orientation, Anger and Aggression, Order and Organisation and Program Clarity). They also reported that the patient characteristics of self-monitoring and restraint, Paranoid ideation and certain cognitive tasks predicted scores on those five subscales. One finding of note for this review was that those who performed better on tasks of abstracted reasoning perceived higher rates of Anger and Aggression, indicating a need for a level of abstract thought to perceive anger and aggression. This links back to the reported findings that members of staff report higher levels of anger and aggression than patients (e.g. Jansson & Eklund 2002a; Livingston *et al*., 2012). If some patient subgroups have lower cognitive functioning at the time of testing, this may illustrate a mechanism to explain the reported differences between staff members and patients' perceptions of levels of Anger and Aggression

2.6.5 Symptoms

Three of the studies explored the relationship between symptoms and ward atmosphere; Beazley and Gudjonsson (2011), Oshima *et al*., (2005) and Oshima *et al*., (2003). In addition to these three studies, further evidence from this review can be found from Dickensen *et al*., (2014), Jansson and Eklund's (2002a;2002b) and Jörgensen *et al*., (2008), who did not consider symptoms per se, but explored diagnosis in relation to ward atmosphere.

Despite reporting on different symptoms, the findings from this focus appear consistent; that there is a relationship between symptoms experienced and perceptions of ward atmosphere. This assertion is further supported by the additional studies. Having said this, the quality of the studies within this focus is at the lower end of those in this review with a range of 27-31, and there are also some methodological aspects to be discussed, which may limit what can be safely concluded from these studies.

In national studies across Japan, Oshima *et al.,* studies (2003; 2005) considered the extent to which under-stimulating social environment in hospitals contributed to negative symptoms of schizophrenia among institutionalised patients. Ward atmosphere was assessed using five questionnaires in both studies, symptoms were assessed using three questionnaires in 2003, and four questionnaires assessed patients' daily dose of anti-psychotic medication in 2005. On both occasions all data was collected from staff about patients.

The 2005 study reported a significant correlation between four of the five negative symptom scales and all but one of the measures of ward atmosphere, and all of the ward atmosphere measures were significantly correlated with the measures of negative symptoms. Similarly, in the 2003 study, the three negative symptom scales and five social environment scales significantly correlated. Weaker, but significant correlations between the five scales used to assess ward atmosphere and positive symptoms were also reported.

Beazley and Gudjonsson (2011) considered the relationship between depression, ward atmosphere and motivation. They investigated if symptoms of depression were associated with motivation to engage with treatment, as mediated by ward atmosphere and secondly, if ward atmosphere was associated with motivation, as mediated by depression. Significant relationships were reported for both models.

There is further evidence of a relationship between symptoms and aspects of ward atmosphere from studies in other areas; Jansson and Eklund's (2002b), found patients grouped together based on high-levels of 'Paranoid Ideation,' had a significantly higher risk of perceiving a low level of Program Clarity, and higher levels of Anger and Aggression. Equally Jansson and Eklund (2002a), reported patients with schizophrenia compared to those with psychosis rated Autonomy and Support significantly lower.

While there appears to be a relatively consistent picture, there are some methodological issues to consider when reflecting on the robustness of these claims, namely the accuracy of the constructs measured (see 2.5.7.2). In Both the Oshima *et al.,*(2003;2005) studies, staff completed the questionnaires relating to ward atmosphere. This links to findings from the wider review, which indicate that members of staff and patients have different perceptions of ward atmosphere. Staff views are still valid, however it is important to hold in mind that the relationship reported is between staff perceptions of patients' symptoms and the ward atmosphere. There is also some contrasting evidence, with two high quality studies, Dickens *et al*., (2014) and Jörgensen *et al*., (2008) whose reported findings both only partial support an association between symptoms and ward atmosphere.

2.6.6 Other

Two of the studies main focus in relation to ward atmosphere did not fit with others in this review; Ros *et al*., (2013) looked at the relationship between perceived ward atmosphere and levels of aggression, and Jansson and Eklund (2004), looked at the helping alliance and ward atmosphere. In addition to these two studies, Jörgensen *et al*., (2009) considered satisfaction and outcome. Satisfaction has been discussed, but this opportunity will be taken to explore the evidence from this study in relation to outcome.

Ros *et al*., (2013) reported a negative relationship between positive ward atmosphere and aggression; specifically, that the subscale Support (from the Prison Group Climate Instrument) was negatively related to the reported number of aggressive incidents. Indirect effects were tested for – results indicating the relation between growth and aggressive incidents was mediated by Support. The authors suggested ward atmosphere impacted upon aggression, however it is unclear what they used to support their directional claims. This was a relatively high quality study (32), and, despite no other study considering aggression and ward atmosphere directly, there is support from other studies; Osborn *et al*., (2011), reported that community services were statically higher on Support and lower on Anger and Aggression, as measured by the WAS, and Long *et al*., (2011a), reported significantly lower numbers of management or prevention of violence and seclusions on the wards where the ward atmosphere was rated as more positive.

Jansson and Eklund (2004), looked at helping alliance and ward atmosphere. They reported significant positive correlation between nine of 10 subscales from the COPES, and strength of patient/therapist alliance (all but the subscale Staff Control). Using Multiple regression, similar to Ros *et al.,* (2013) findings, the Support subscale was reported to be the most important, accounting for 44.9% of variance where helping alliance was the dependent factor. The indication of a relationship between helping alliance and ward atmosphere was supported by Long *et al*., (2011a) and Bressington *et al*., (2011). Long reported that the ward which scored lower than the others on ward atmosphere, also scored significantly lower on therapeutic alliance, and Bressington reported patients' perception of therapeutic relationship with staff was the strongest predictor of satisfaction. The Relationship dimension also appears to be significant in predicting satisfaction, which could indicate a mediating or moderating relationship between therapeutic alliance, ward atmosphere and satisfaction. Indeed, Bressington reported a significant positive correlation between overall satisfaction and perceived therapeutic alliance, and between overall satisfaction and the EssenCES domain ES and TH. No contradictory findings were reported within the studies reviewed.

With regards outcome, Jörgensen *et al*., (2009), reported that the better the ward atmosphere the better the outcome. When comparing three wards, one ward scored significantly lower on 8 and 9 of the 11 WAS subscale than the other two wards respectively. This lower scoring ward reported a significantly smaller reduction in overall difference of patient symptoms and a significantly lower Global Assessment of Functioning scores of patients at discharge. However, no difference between wards was found when comparing other measures used to assess outcome. These findings were partially supported by Jannsoan and Eklund (2004), who reported wards with more positively-rated ward atmosphere also had a significant difference in the reported reduction of overall symptoms. These findings further link to the assertions made in Oshima *et al*., studies (2003;2005) that the ward atmosphere impacted upon symptoms.

2.7 Clinical and Research implications of the findings

Ward atmosphere has been demonstrated to have an impact upon satisfaction (e.g. Bressington *et al*., 2011), motivation (Beazley & Gudjonsson 2011) and outcome (Jörgensen *et al*., 2009), and evidence supports perceived ward atmosphere is altered in relation to service characteristics (e.g. Livingston *et al*., 2012) and certain characteristics of those utilising the services (e.g. Campbell *et al*., 2014; Dickensen *et al*., 2014) and is thus an important aspect of inpatient care. Clinically, the importance of the role of relationships within ward atmosphere is of note; from this review it is the aspect of ward atmosphere which has been most consistently indicated as important (e.g. Campbell *et al*., 201; Jansson & Eklund 2004; Middelbøe *et al.,* 2001). Its implications for clinical practice are very meaningful, as it indicates that interactions between staff and patients, which develop such relationships, are likely to be of importance when considering how to improve ward atmosphere and therefore the multiple factors it is associated with. This is not of course a straight forward endeavour, for example the research which considered the role of patients' characteristics indicates that individuals may respond and perceive ward atmosphere differently. However, there is evidence (Nesset *et al*., 2008) that interventions, such as staff training, can improve ward atmosphere. Further research in this area may be beneficial, as not all interventions have the desired affect (Livingston *et al*., 2013).

Due to ward atmosphere's relationship to other aspects of care, by measuring ward atmosphere it might be possible to assess other hard to quantify factors within inpatient services. This was suggested by Long *et al.,* (2011a) who proposed that assessing an individual's perceptions of ward atmosphere may offer a valid insight into that individual's progress. This could be extended to other variables; for example, ward atmosphere may help services monitor whether recovery principles are being met (Livingston *et al*., 2012).

The range of factors which have a relationship to ward atmosphere is of interest in and of itself, although not unexpected given what ward atmosphere consists of. From a research point of view this presents both opportunities and difficulties. Due to the multiple factors involved, research in clinically relevant settings will always have difficulties managing so many factors. This also links back to the point about mediating and moderating variables; what is seen within this review is the complex nature of ward atmosphere and the difficulty in separating out variables. For example, the differences in ward atmosphere demonstrated by studies such as Long *et al.,* (2011a) who found differences in service characteristics produced differences in ward atmosphere, could in part be explained by the differences in patient characteristic, as demonstrated by research such as Jansson and Eklund (2002b). Yet, due to the broad nature of ward atmosphere, the breadth of areas which can be usefully considered is vast. As ward atmosphere relates to such a range of variables within inpatient settings, it could be a useful starting point to triangulate areas of interest. Indeed, some of the mediating and moderating variables which have been discussed may provide useful areas of research, as they could be used to explore possible mechanisms to either affect, or use, the impact of ward atmosphere.

2.8 Summary and Conclusion

At the broadest level, the balance of evidence from this body of research is of good quality, and indicates that ward atmosphere has a relationship to a range of indices of care within inpatient mental health settings. When looking for a pattern across this body of research, two points emerge; firstly, taken as a whole, this body of research indicates that the relationship between ward atmosphere and other factors is likely to be bi-directional, and secondly, that the aspect of ward atmosphere most consistently indicated as important is that which related to relationships. It has been indicated in satisfaction (e.g. Middelbøe *et al*., 2001), service attachment (Campbell *et al*., 2014), better perceptions of the helping alliance (Jansson & Eklund 2004) and lower levels of aggression (Ros *et al*., 2013). The importance of this aspect of ward atmosphere was not limited depending on what measure was used; it was reported by studies using a range of measures, including the WAS, COPES and EssenCES. This is of interest both in terms of future research and clinical practice, as it would indicate the relationships and interactions between staff members and patients could have widespread impact via its influence on ward atmosphere. Despite the quality of the evidence in this review being high on the whole, some caution is needed when considering the evidence en masse, due to the differences in what the papers were researching.

Due to the diverse nature of the topics investigated within this review, more in-depth findings are perhaps better explored within their focuses, as done above. The breadth of topics covered in relation to ward atmosphere was necessary, given the aim of the review was to consider ward atmosphere in relation to any additional indices of care. There are several consequences of using such broad search terms and having such an inclusive review question, as was done here. Firstly, it presents difficulties in drawing meaningful conclusions which relate to the whole body of research, as like for like is not being compared. To manage this, certain decisions were made as to how the data was handled. The breath of topics within the research led to the decision to review the evidence, both as a whole and by grouping the papers based upon their focus. Using these focuses was beneficial for managing the quantity of evidence and for ensuring the results could be compared meaningfully. When thinking about what the evidence shows from each focus, it is important to bear in mind that the structure placed upon them was at the author's discretion; it is recognised that although the structure used was logical it was not the only way the papers could have been subdivided, and that the choice of focuses imposed has implications for how the information has been drawn together. In order to minimise any possible bias this may have been produced, particular care was given to ensure that the synthesis was not limited to within focuses, but considered across all papers.

This leads on to the third notable point concerning the breath of topic included in this review. Namely, that it has highlighted possible mediating and moderating variables, which may explain some of the reported variations in ward atmosphere seen. This can be illustrated by the evidence from Jansson and Eklund (2002b), which reported the levels of certain cognitive functions (abstract reasoning abilities) impacting upon the levels of Anger and aggression and the individual's perception. This, taken with the findings from the studies looking at differences between staff and patient perceptions of ward atmosphere, indicates cognitive functioning could be a mediating factor, which explains at least some of the differences reported between staff and patient perception of the Anger and aggression subscale in relation to ward atmosphere. Equally, as has been discussed, findings from patient characteristics, such as age may act as a moderating variable between satisfaction and ward atmosphere and perceived helping alliance ward atmosphere (Kuosmanen *et al*., 2006; Jansson and Eklund 2004). Having an understanding of the possible mechanisms is of particular use when thinking about how findings could be applied clinically. Possible explanations of how ward atmosphere may impact upon satisfaction (for example) could also help researchers develop and test further hypothesis, leading to research which could support services in being able to improve the perceived ward atmosphere.

In addition to what this systemic review offers in terms of information about inpatient services and ward atmosphere, the findings can also be considered in relation to the understanding of the recovery orientation, as another aspect of inpatient services. There are several possible benefits of this. Firstly, there is some evidence of a correlational relationship between perceptions of ward atmosphere and perceptions of recovery orientation within services (Livingston *et al.,* 2012) and secondly, both can be viewed as service level concepts which impact upon individuals' care. As such, an understanding of one may offer insights into the other. Conceptually, one of the notable similarities is that the systematic review highlighted the importance of the role relationships play in ward atmosphere; this has also been indicated to be true when considering the strength of the recovery orientation (Mezzina *et al.,* 2006; Tew *et al.,* 2012). At a more practical level, the systematic review demonstrated that the measures used to assess ward atmosphere are all self-report, as was seen with measures of recovery orientation. As such, for both areas there are certain limitations in the measure available (see 1.3 for further details). The evidence from this systematic review with regards how measures are utilised, also provides information when using self-report measures to assess recovery. Namely, the review highlights that there is a difference in perceptions of staff and patients of ward atmosphere; this could indicate that when considering recovery orientation, both researchers and clinicians could need to assess both patient and staff opinions, or develop a method with is not so reliant on opinion; such as the observational tool proposed within this project.

Chapter 3: Methodology

3.1 Introduction

This chapter will describe the method used in this study. For ease of reading, the chapter has been divided into two parts, the first of which explains the development of the tool, including the initial development of the codes and the Delphi study, and the second provides the method for the validation and reliability testing of the tool within a clinical setting.

The inclusion or exclusion of codes from the final tool via the Delphi process in part 1, and the data collected from a standardised measure of recovery, which was used to differentiate wards in part 2, do not related to one another. However, the outcome of both of these processes are related to the presence of recovery supporting and recovery hindering interactions between staff and patients in inpatient mental health settings. This is the link between these two aspect of the project.

For both aspects of the study, this chapter will include details of the design, recruitment and sample, the measures employed, the procedure for gathering the data and details of the research governance.

3.2 Part 1: Development of the tool

3.2.1 Design

The aim of part 1 was to develop a tool to assess the recovery orientation of a service by categorising the types of interactions seen between staff and patients in adult inpatient mental health services. To do this, the study used qualitative data from the recovery literature and the clinical expertise within the research team to create concrete descriptions of types of interactions commonly seen between members of staff and patients, which the evidence would suggest would either hinder or support the recovery process. These were termed ‘codes’. For each code, descriptions of what might be seen if the code was being enacted were also developed; these were termed ‘exemplars’. To establish which codes needed to be included in the tool, a Delphi method was used to generate a consensus from a panel of experts made up of mental health professionals working in Rehabilitation and Recovery inpatient mental health services.

3.2.2 Initial development of the codes

As discussed, recovery is a poorly defined concept, which relies on self-report measures; both issues which were considered when designing the new tool. As a means to counter some of the conceptual uncertainty, evidence gained from Chapter 2 (the Systematic Review) relating to the role of ward atmosphere was utilised when considering the nature of inpatient services, and how and what the tool being developed aimed to measure. Discussion between the research team led to the conclusion that the tool in development was more akin to a measure of recovery than a measure of ward atmosphere, although the design was complementary to ward atmosphere and recovery. This was reflected in the evidence base which was used to create the initial codes, and as such, a review of the recovery literature was used as an evidence base from which the initial codes were developed. Recent qualitative research findings from Bonney and Stickley (2008) and Leamy *et al.* (2011) (see section 1.2.1 for details of the studies) were reviewed by two members of the research team (the author and the academic supervisor, a consultant clinical psychologist in Rehabilitation services). Similarities and differences between the themes from each study were discussed and compared. From this, seven over-arching themes were identified:

- Identity
- Connectedness and the social domain
- Hope and optimism
- Service needs and agendas
- Power and empowerment
- Risk and responsibility
- Meaning in life

Taking this as a framework, the clinical experience of the team was used to create the initial codes (umbrella descriptions of types of interactions) and exemplars (clinical examples of the codes) for each code (see Table (ii) for the codes and exemplars created). After the initial development, a third member of the team, a consultant clinical psychologist working in rehabilitation and low secure services, independently reviewed the codes and exemplars. No further codes were added at this stage.

3.2.3 Delphi method

The Delphi method is a structured communication method, originally used in defence research in the USA in the 1950s (Helmer & Rescher 1959). It has been used in mainstream research across a number of areas since the 1970s (Spivey 1971; Matthews Mahaffey, Lemer & Bunch 1975; Loughlm & Moore 1979) and is commonly used to develop health care indicators (Nieuwenhuijsen, Verbeek, de Boer, Blonk, & van Dijk, 2005; O'Brien, Boddy, Hardy, & O'Brien, 2004; Boulkedid, Abdoul, Loustau, Sibony & Alberti, 2011). The method consists of identifying relevant individuals to create a panel of experts and questioning them on specific items in a series of rounds. This is done individually and remotely, such as by post. Each expert then responds to the researcher. The individual responses of the panel are scrutinised and collated by the researcher, in order to identify a consensus on the items asked about. The researcher then compiles a comprehensive list of items for re-submission to the panel. The experts are then asked to reconsider the list and respond again, indicating their agreement or disagreement with the items. These responses are collated once more and the process repeated until the pre-determined number of rounds is reached, or a consensus for each item is achieved. The procedure is anonymous and confidential. (Boulkedid *et al.,* 2011; Williams &Webb, 1994).

There are a number of benefits of using this method. Firstly, it is argued that the method provides a consensus without the bias created by being in a face-to-face group, as in alternative methods such as focus groups or committee meetings, thus encouraging honest opinions to be provided, free from peer group pressure or being dominated by an individual member or members of the group (Jairath & Weinstein 1994). It is argued that this is of particular note in health care settings where a clear hierarchy exists (Beech, 1991). Secondly, due to the successive rounds used, participants are able to retract, alter or add to the views they give, with the benefit of considered thought (Williams & Webb 1994). Lastly there are a number of practical benefits. Due to the remote nature of the way data is gathered within this method, larger panels from diverse geographical areas can be created with greater ease than when trying to physically assemble the panel in one location. It also allows participants to answer the questions at their own convenience (Boulkedid *et al.,* 2011).

With this in mind, the Delphi method, consisting of two rounds was employed to gather a consensus on the initial codes. As discussed above, there are a number of advantages to using this design, notably it was selected to allow a consensus to be reached across the three Health Boards involved in the study. It also allowed a larger sample to be gathered than if a focus group or committee had been used.

3.2.4 Ethical Considerations

3.2.4.1 Informed consent

All Mental Health Professionals approached to take part in this project were given a full explanation of the project and their role within it. Participant Information Sheets (PIS, Appendix 3) and an opportunity to ask any questions, either to the local collaborator or via email with the main researcher, were provided. The individual's right to withdraw was also discussed and made clear within the information provided. The Mental Health Professionals asked to take part were given a minimum of two weeks to consider their involvement. Written consent forms (see Appendix 4) were provided to all those who indicated their interest in taking part.

3.2.4.2 Protection of individuals

None of the potential participants who were approached for this study were considered vulnerable. It was made explicate that the individual's decision as to whether to become involved in the project or not had no impact on their clinical practice or job role.

3.2.5 Participants

3.2.5.1 Delphi sample

There is currently no definition of what constitutes a sufficient number of participants for a Delphi survey (Reid 1988; Wilhelm, 2001). Akins, Tolson and Cole (2005), conducted a study in which they used 23 professionals trained in the area of interest, as the panel within a Delphi survey, and compared the results from the panel to two computer-generated samples of 1000 and 2000 resampling iterations, which they created using the bootstrapping technique. The reported results indicated that the number of selected experts in the panel was sufficient to ensure reliability. Therefore, it was decided to involve a minimum of 23 participants within the first round, as this usually contains the largest diversity of responses (Akins, *et al.,* 2005).

34 Mental Health Professionals were recruited to create the expert panel for the first round. 11 of the original 34 were further consulted for the second round. A breakdown of the demographics of the sample is provided within the results chapter (section 4.1.1.1).

3.2.5.2 Sample selection

After due ethical approval was received, an expert panel made up of Mental Health Professionals (professional in this case being defined as an individual registered with the appropriate governing body) working with inpatient rehabilitation mental health services in one of three Health Boards involved was generated. A local collaborator in each Health Board met with the managers of the wards to discuss the project. Once ward managers had given their consent for their staff to be approached to take part in the study, the local collaborator approached staff members to introduce them to the parameters of the study, providing PIS (Appendix 3) and gathering written consent (Appendix 4) from those who indicated their willingness to participate.

A smaller expert panel of 11 people made up from the original sample from two of the Health Boards was used for the second round of the Delphi study. This selection was based upon those who could complete the questionnaire within the time frame.

3.2.5.3 Inclusion criteria

Registered Mental Health Professionals, who had direct contact with clients and worked within inpatient Rehabilitation mental health services were included.

3.2.5.4 Exclusion criteria

Mental health professionals who worked in alternative settings were not eligible to take part, nor were professionals who worked within inpatient Rehabilitation or Forensic Mental health services, but had no clinical contact (e.g. managers). Health care assistants or other professionals working within the area whose job did not require professional registration were not eligible to take part.

3.2.6 Materials

All possible participants were provided with an information sheet (Appendix 3). Those who indicated their willingness to take part were provided with a consent form (Appendix 4) and a demographic information sheet (see Appendix 5).

Members of the panels were provided with two Code Lists, at two separate points over the course of the study.

The first list was comprised of four parts (see Appendix 6 for first code list) namely:

1. A set of instructions.
2. A list of the codes with the Relevance and an Importance rating scale.
3. A list of exemplars relating to each of the codes.
4. An additional section for comments, additional codes and exemplars.

The second list of codes was comprised of three parts (see Appendix 7 for the second code list):

1. A set of instructions.
2. A list of the codes to be re-rated and the new additional codes with the relevant rating scales.
3. A list of exemplars relating to each of the codes.

For the first code list, Bonney and Stickley's (2008) systematic review of recovery research and thematic analysis and the conceptual framework proposed by Leamy's *et al*, (2011) systematic review and modified narrative synthesis, were used as the evidence base. From this the codes were developed by the research team. This was done by working through the thematic analysis and modified narrative synthesis from each piece of work and comparing the findings from each paper. The six main themes from Bonney and Stickley's (2008) literature review (identity, the service provision agenda, the social domain, power and control, hope and optimism, risk and responsibility) and the five recovery processes identified in the framework proposed by Leamy *et al.,* (2011) (connectedness, hope and optimism about the future, identity, meaning in life and empowerment) were reviewed. To do this, each of the 11 themes and the elements which were grouped to create the themes were considered with a view to highlighting overlap and areas of differences between these themes from the two sets of research. From this the following seven over-arching categories were created (see section 3.2.2 for the 7 categories).

These overarching themes were then considered specifically in relation to how they might be embodied within interactions occurring between staff members and patients in an inpatient rehabilitation setting. Using the themes as a starting point, the research team developed a number of codes and exemplars of the codes. The codes created were classified as either positive, meaning the literature would suggest they would support a recovery orientation, or negative, meaning the literature would suggest they would hinder a recovery orientation, based upon what the research indicated had had an impact upon the 'recovery process'. The team's clinical experience was used in addition to the evidence base, and to ensure the codes created covered clinically relevant interactions, the team also produced codes which encapsulated their clinical experience. In these cases, the code was included if it was a) related to one of the themes, and b) if it did not duplicate one of the other codes which had already been created. Employing this process, 26 codes were created. These codes were then independently reviewed by another member of the research team. This process did not create any further codes or any alteration in the codes created. Four of the codes related to service level factors, the remaining 22 were at an individual interaction level. The four service level codes (codes 22-25) were not given a positive or negative value, rather they were included to allow the tool to consider service needs and agendas as identified by the qualitative evidence. Using the evidence from the recovery literature, there was a discussion within the research team to determine if the remaining 22 codes were needed to support or hinder the recovery process; 11 were considered to support, and 11 were deemed to hinder recovery. Table (ii) lists the 26 codes developed and their exemplars. Codes which were decided to support the Recovery process are shown in blue, those which were deemed to hinder the process are shown in black, and the service level codes are shown in red.

Codes	Exemplar interactions
1 Encouraging engagement in activities.	Discussing available activities and what the client might like to do. Offering support with carrying out tasks e.g. I noticed you have some washing to do, shall I come and give you a hand?
2 Forcing engagement in activities.	Demanding people take part in certain activities e.g. you need to do your washing now. We are all playing bingo, come and join in. (Demanding tone and gestures may be used).
3 Encouraging the development of independence.	Example of doing *with*, not doing for; "shall I help you with X?" as opposed to "let me do Y for you, it will be quicker."
4 Non-verbal dismissal	Gesturing "one, two, five minutes" with a hand through the window when a client approaches the nurses' station door/window.
5 Discussing client's personal information (in public settings).	Discussing a private aspect of the client's health/care in communal areas. Making reference to personal events in communal areas e.g. Hurry up the taxi is here for your hospital appointment.
6 Offering choice.	Either open choice; what would you like to do today/this week etc? Or selective choice where necessary e.g. "would you like to do X or Y?"
7 Tokenistic involvement	Brief interactions in ward round where clients are involved minimally e.g. "how are you? Everything going ok? Anything you want to ask, no? Good…"
8 Staff initiated non-clinical interactions.	Initiating general conversation or engaging in any activity which is not directly related to clinical care e.g. NOT medication/self-care/food.

Codes	Exemplar interactions
9 Positive responses to client initiated interactions.	Staff responding in an engaging way to client’s interactions.
10 Negative responses to client initiated interactions.	Staff brushing off the advance e.g. “Not now, I am busy…”
11 Joint involvement in task/activities.	Do something with, not to, or for, someone. “Let’s do X together…”
12 Clients not being engaged with.	Staff members walking passed a client sat in the communal area, without making effort to interact.
13 Interactions having an “illness focus”.	Focus of interactions on symptoms/medication etc., treating a person as a condition rather than as an individual.
14 Seeking input from clients.	Either the service or individual staff looking for input from clients on issues, rather than simply passing on information. Also relates to providing choice.
15 Not seeking input from clients.	Telling clients “how things are”.
16 Active listening skills.	Paying attention, giving eye contact, nodding, responding where appropriate.
17 Staff not being in the moment with clients.	Looking at watches, talking to someone else, staring off into space, checking their phone etc.

Codes	**Exemplar interactions**
18 Recognition of communication needs.	E.g. language barriers; does someone need an interpreter? Is it more effective to communicate with someone via pictures? Does someone use sign language? Can that client read the care plans you have given them to read? Individual staff knowing and communicating appropriately AND service provision for things such as interpreters.
19 Ignoring communication needs.	Not recognising the factors listed above.
20 Staff recognising strengths and interests.	Using information of someone's strengths to help support them e.g. "….well you could try making a list of the things you want to get done if you are worried about not having time to do it all, you are good at planning…."
21 Punitive interactions.	"If you don't do X/continue to do Y, you will…. miss your next smoking break….not be allowed section 17 leave etc."
22 Set meal times.	Are there choices with meals given? Do people eat together? Are the clients fed "hotel style"? Are there facilities for clients to prepare their own food/snacks/hot drinks? Do clients have input into their diet?
23 Set bed times.	Are clients sent to bed/bedrooms at certain times? Are TV turned off? Do communal areas become off limits?
24 Set getting up times.	Do clients have to get up at a certain time? Are doors locked during the day? Is this individual (e.g. agreed by client/staff/family etc. in a care plan?) or a whole unit system?
25 Set smoking times.	Can clients control when and how much they smoke? Is this individual (e.g. agreed by client/staff/family etc. in a care plan?) or a whole unit system?

Codes	Exemplar interactions
26 Encouragement/practical support of pro-social activities, including off the unit.	Suggestions of activities which are pro-social. E.g. calling family; is going out into the community encouraged/valued and facilitated or is “we don’t have the staff” more common?

Table (ii) Codes and Exemplar developed by the research team.

3.2.7 Study Procedure

After the initial codes had been developed by the research team, the Code List was piloted to ensure it was understandable. A Clinical Psychologist and an Occupational Therapist, who both attended a relevant professional faculty training day (Faculty of Psychosis and Complex Mental Health, Division of Clinical Psychology) were consulted. They were asked to complete the rating scales and offer feedback on the instructions and process. No changes were made in light of this feedback.

Once this had occurred, the sample selection produced was followed, (see section 3.2.5.2) Following this, the participants were provided with the codes generated by the research team (see Appendix 6).

Each member of the expert panel was asked to rate each code on the two 1 - 7 point Likert scales. The first scale 'Relevance', asked panel to rate the codes as to whether it was likely to occur in the service within which they worked, where 1 = not relevant 7 = very relevant. The second scale, 'Importance', asked the panel to rate the codes with regard to the impact such interactions have upon the quality of care within their setting, where 1 = no importance, 7 = very important. The panel members were not asked their opinion as to whether they felt such interactions should or should not occur, or whether they felt they would hinder or support recovery. The rating scale of 1-7 reflected the scales used in other studies, which had used the Delphi method to create a consensus relating to recovery in mental health (Lakeman, 2010).

There was also an additional space within the form where participants were invited to make any comments on the wording of the codes and exemplars, provide additional codes they felt were missing, and/or to suggest additional exemplars for the codes.

Once the scores had been collected, descriptive statistics were run on them. Boulkedid *et al.,* (2011) systematic review into the methodologies used in Delphi studies investigating health care indicators, was used to inform the choices made for the Delphi process, as to date there is no agreed methodology. Boulkedid *et al.,* (2011) identified five methods of creating a consensus. The two most common methods were as follows:

"In 22/62 (35%) studies, indicators with median scores above a predefined threshold and a high level of agreement among panel members were selected….. (b) In 10/62 (16%)

studies, selection was based only on a median score greater than a predefined threshold" Boulkedid *et al.,* (2011) pp.4

From this, the cut off of 75% of the expert panel was decided upon. For the purpose of this study, this was considered a consensus. The ratings for both the Relevance and Importance scale for each code were placed onto a spread sheet These scores were subjected to descriptive statistics using SPSS (see Appendix 23 the SPSS output for Round 1). Those codes for which 75% or more of the assessors gave a rating of either 1 or 2 were discarded from the process. Those codes which received a rating of either 6 or 7 from 75% of the panel on either scale, were included in the tool. A code only needed to reach a consensus (75% or more of the sampling scoring it 6 or 7) on one of the scales (Relevance or Importance) to be included in the tool. See Appendices 21 and 22, for details of the codes which were included and excluded in Round 1.

Following this process, a second list of codes for re-rating was distributed to the panel (see Appendix 7 for the second code list). A shorter time period was allowed for the collection of the second round of data (three weeks). This led to a reduction in the size of the panel to 11 professionals from two of the original three health boards. Demographic information was not collected from all panel members from this round, and has therefore not been included. In the second round, the panel were asked to rate the codes on the same scales as in the first round. Additional space was given for further exemplars; however additional codes were not asked for. Once the data from the second round had been collected, these too were subjected to descriptive statistics using SPSS (see Appendix 24). As the second round was the last, codes could only be included if they met the criteria, all others were discarded. To reflect this, the procedure was altered; any codes rated within the top half of the scale (5,6 or 7) would be included, all others discarded. However, to maintain the rigour, the percentage of the panel needed for agreement was increased from 75% to 80%. Those codes that enough of the panel rated sufficiently high were added to those codes included after Round 1 to create the new tool, the rest were discarded. See Appendix 25 and 26 for details of the codes which were included and excluded from Round 2.

3.3 Part 2: Validation of the tool and Reliability testing

3.3.1 Design

Part 2 of the study aimed to assess the validity and reliability of the newly developed tool. A sample of eight wards which fitted the inclusion criteria were recruited to the study via the

Principal Investigator in each of three Health Boards in the south Wales area. Within these wards, a standardised self-report measure of the recovery orientation of the service (The Recovery-Self Assessment, Provider Version {RSA-PV}) was completed by participating staff members on the wards. The mean score for the RSA-PV from each ward was calculated, and the ward with highest overall mean and the lowest overall mean identified. These wards were selected, and the tool was piloted in them. A Chi-Square was used to assess if the number of positive and negative interactions as assessed by the new tool, differed to a statistically significant level between the two wards selected.

A time-based sampling method was used to complete the observations of the interactions which were assessed using the tool. This system involved the researcher moving around three different areas of the ward between five-minute time-slots. This was done to minimise bias of recordings in one area of the ward, and to ensure as far as possible, that the interactions seen were not related to one another, as per the assumptions of a Chi-Square. The observation sessions were carried out across the length of a day to ensure observations were gathered at different times of day and over a two-day period.

The reliability of the tool was assessed by two of the researchers (by main research and academic supervisor) coding 20% of the interactions recorded using the tool independently. Whether the interaction was seen as positive or negative, and which codes each researcher assigned was then compared.

3.3.2 Ethical considerations

3.3.2.1 Confidentiality and right to withdraw

To ensure the anonymity of the data, code-numbers were assigned to each health board (1,2 and 3) and a second digit used to identify each ward. This led to the wards being coded as; 11, 12,13, 21,22,31,33 and 34. All information gathered (observation, RSA-PV and demographic information) was identified using the code-number; this allowed for identification of the ward the information came from without any individuals' or services information being evident. When analysing the data, no services were named or identified, e.g., participant 1; ward 12. Participants were made aware that they had the right to withdraw without explanation at any point during the study. It was also made clear that due to the anonymous nature of the data, once collected it could not be withdrawn, but that the participant was not obliged to take any further part in the study.

3.3.2.2 Participant burden

The observational nature of the study ensured that the burden for participants was minimal. They were asked to complete the RSA-PV. This measure was selected in part because it was shorter than some other valid recovery orientation measures (e.g. Recovery Based Program Inventory, which has 148 items vs RSA-PV 36 items). It was deemed to take 10 minutes to complete (Campbell-Orde, *et al.,* 2005) and it was also supported by literature (Burgess *et al.,* 2010; Campbell-Orde, *et al.,* 2005).

3.3.2.3 Further Ethical issues

The participants were given sufficient time to review the study information before deciding whether to participate or not. Time was provided for questions to be answered in person. Contact details of the main researcher and supervisor were provided. Prior to the study taking place, a system of supervision was put in place, so that if any practice was observed which was viewed as unethical or concerns were raised personally for the main researcher by the process, supervision was available.

It was recognised that, despite not being the focus of the study, patients living within the wards sampled may have been concerned or anxious about their involvement in the observations. Information sheets (see Appendix 8) were generated in collaboration with service users, to ensure they were suitable for the intended audience. The observer also ensured time was spent explaining the research in person to clients on the ward who were interested, and that it was only the actions of the consenting members of staff which were of interest. A protocol was also developed for cases where a client asks about the research. A verbal response was provided, the client information sheets were offered, an explanation as to who the local collaborator was reiterated and contact details for members of the research team were provided.

3.3.3 Participants

3.3.3.1 Sample

The sampling operated at two levels:

Site level; this refers to the eight wards which were involved with the study.

Individual level; this refers to the individual members of staff who agreed to take part and created the sample from each wards.

The sample size related to the number of observations which were recorded. The sample size was determined in order to achieve a desired power of 90%. In order to detect a difference in proportions between groups of 0.5 to 0.712 as statistically significant, using a Chi-square test at the 5%, a total sample size of 240 interactions was required (Pearson, 1900). It was therefore decided to collect 120 observed interactions per ward for 90% power. This was calculated on nQuery v7.0.

3.3.3.2 Sample Selection

Members of staff who worked on the wards that had agreed to be involved in the study were made aware that the study was being carried out on their ward by the local collaborator. The main researcher then attended a staff meeting to given further information and Participant Information Sheets (see Appendix 9) to potential participants, and the opportunity was given for individuals to ask questions (either in person or via email). Members of staff who indicated their willingness to take part were provided with a consent form (see Appendix 10) and written consent was obtained.

On the two wards which were identified for observations to take place on, a further step in sample selection occurred. On these two wards, on the day prior to the observations starting, and on each day of observations, the researcher ensured that all members of staff on shift were made aware that the study was being carried out. Anyone who indicated that they were not aware of the study already was provided with information and the opportunity to take part if they wished, by giving written consent. All members of staff working were made aware that they did not need to take part in the study, and if they did not, any interactions in which they were involved would not be recorded. However, all members of staff working over the two days on the two wards consented to take part.

3.3.3.3 Inclusion criteria

Site level: Inpatient rehabilitation mental health services in 3 Health Boards in south Wales were eligible to take part.

Individual level: Members of staff who worked directly with clients who indicated their willingness for the research to take place were eligible.

3.3.3.4 Exclusion criteria

Site level: No other services (i.e. those which serve different populations or which are community based) were eligible to take part.

Individual level: Members of staff who do not have direct contact with clients (e.g. maintenance or domestic staff) were not eligible to take part.

3.3.4 Materials

The following materials were used;
Participant Information sheets (see Appendix 9).
Consent forms (see Appendix 10).
Demographic information sheet (see Appendix 11)
The RSA-PV questionnaire (see Appendix 12).
Information sheets for service users living within the wards where the research was conducted (see Appendix 8).
Observation Matrix (see Appendix 13)
The newly developed tool (see Appendix 14)

3.3.4.1 Measures used

Two measures were used; the standardised measure of the recovery orientation of a service, and the newly developed tool which was being piloted.

The standardised measure used was the Recovery Self-Assessment Provider Version (RSA-PV) (see Appendix 12). The RSA-PV has versions for four stakeholders (person in recovery version; family/significant others/advocates version; provider version; and CEO/Agency director version). The RSA-PV is designed to measure the extent to which recovery-supporting practices are evident in mental health services. It consists of 36 items which assess five subscales: life goals; involvement; diversity of treatment options; choice; and individually-tailored services. Participants are asked to rate each item on a 5-point Likert scale. The RSA-PV is reported to have undergone an appropriate process of item development, including drawing on stakeholders' input as well as testing, including factor analysis (O'Connell *et al.,* 2005). It was one of four instruments that met the criteria at each level of the hierarchy used within Burgess *et al.,* (2011) review of recovery measures in Australia, as well as being one of the measures recommended by Campbell-Orde *et al.,* (2005) in the second volume of compendia of measures of recovery in the United States.

The second measure used was the tool developed in the first stage of the project, consisting of codes and exemplars (see Appendix 14), which was used to code staff-patient interactions.

3.3.5 Study procedure

3.3.5.1 Identification of wards for observation

Prior to the study gaining the necessary ethical approval and sponsorship, service leads and ward managers within eight wards which met the inclusion criteria in three Health Boards in south Wales were approached and asked if, in principal, they would be willing for the second part of study to take place within their wards. Sponsorship was obtained though Cardiff University. Following this, ethical approval was gained from the National Health Service, via Integrated Research Application System (IRAS). Finally, permission was gained from each of the three health boards, via each of their Research and Development departments (see Appendices 15, 16 and 17 for letters of access). All eight sites indicated their willingness to take part. Wards were then assigned a number to ensure anonymity. These were used throughout the course of the study and the relationship between the coded number and the ward was known only to the researcher. After having initially agreeing to take part, it became apparent that one of the eight wards was unable to participate, because of difficulty with staffing levels at the time. Consequently, seven of the eight wards took part in the study.

Members of staff were then recruited into the study (see section 3.3.3.2 for further details). Those who indicated their willingness to take part were provided with a consent form (see Appendix 10) and written consent was obtained. Participants were asked to complete a demographic information sheet (see Appendix 11) and the Recovery-orientation questionnaires (the RSA-PV, Appendix 12). All data from each ward was marked with the number the ward had been assigned.

The scores from the RSA-PV were entered into a spread sheet by administrative support, thus ensuring that no one on the research team knew which wards were likely to have scored highest or lowest mean score on the RSA-PV. The mean for the RSA-PV total score for each ward was calculated by an individual not involved with the research, to further ensure blind assessment. The two wards with the greatest disparity between the overall RSA-PV scores were identified, and the code-number of these wards was given to the researcher; however, no information on the score levels of the wards on the RSA-PV was provided to the researcher. At this stage, the other wards were informed that no further participation was required from them. See section 4.2.1.3., Table (vii) for the range, mean Standard Deviation and Coefficient of Variation of the data from the RSA-PV for the seven wards.

3.3.5.2 Observations

Using the method described above, two of the initial seven units were selected for the observations to take place upon. The wards selected had been given the code-numbers 12 and 31. Ward 12 is a 10-bed, secure Rehabilitation and Recovery inpatient ward. The staffing levels are five members of staff on the early shift, four on the late shift and one member of staff working on a mid-shift, which spans half the early and half of the late shift. During the night, the staffing levels are two qualified members of staff and two Health Care Assistances (HCAs). Ward 31 is a low secure ward within a Rehabilitation and Recovery service, it has 13 standard beds and one emergency bed. The staffing levels are five members of staff during the day, and three at night.

3.3.5.2.1 Observation protocol

3.3.5.2.2 Timing and quantity of Observations

Once the two wards had been identified, they were both contacted and a week for data collection arranged. On both wards, the data collection was arranged to be carried out on three consecutive days; on the first day, the main researcher went into the wards and explained the procedure to the patients, opportunity was given for people to ask questions and information sheets (see Appendix 8) were distributed to the patients. On the following two days, the observations were carried out. The possibility of the researcher returning if not enough data had been gathered was discussed and agreed upon. However, this was not necessary.

To ensure face validity and reduce bias of the observations, it was decided that the observations should come from across the three periods within a day. To achieve this, the day was divided into three periods; ‘morning’ 0930hrs-1230hrs, ‘afternoon’ 1230hrs – 1530hrs and ‘evening’ 1530hrs -1830hrs. These time frames were used as guides. On the first observation session of the day, the earliest observations started was 09:30, the latest was 10:00. The greatest disparity in start time between the wards occurred on the afternoon session of the second day of observations; on ward 12 this started at 12:25, on ward 31 this started at 13:00. To ensure data was drawn from each time period, it was decided that the researcher would stop an observation period once 20 interactions had been observed. Had insufficient interactions been observed during the allotted time period, the researcher would have returned to the ward on a third day. Between each observation session, the researcher left the ward. The actual starting and stopping times for each observation are presented in Table (iii). All observation sessions fell within the time periods specified.

Ward	Session 1	Session 2	Session 3	Session 4	Session 5	Session 6
12	09:30-1200	12:45-15:10	15:30-17:20	9:50-11:40	12:25-14:00	14:35-17:00
31	09:45-11:30	13:30-15:10	15:35-17:45	10:00-12:15	13:00-15:10	15:30-17:45

Table (iii) Timings of observation periods.

Ward 12 was sampled first. On the first day of the agreed week, the main researcher took part in the morning meeting, during which the study was explained to patients. Two researchers then spent time on the ward discussing how and where the observations should take place. This was piloted with both researchers for 20 minutes and then discussed. In light of this, a slight alteration was made to the method; instead of simply recording every time a new staff member or patient was involved in an interaction, the letters and numbers would be used to identify staff members and patients respectively (e.g. first staff member observed = SA, first patient observed = P1).

Within each unit, a number of communal areas were identified; three in all for both ward 12 and 31. Prior to the observations taking place, it was agreed in which order the observer would move between the communal areas in each ward. To ensure that the interactions were independent of one another, after every observation the researcher would move to the next area. The observer spent up to five minutes observing any one area, if no interactions occurred within that five minutes, 'no interaction' was recorded, and the reason why. The reasons were recorded as one of three possibilities:

- due to no staff being present.
- no patients being present.
- no one at all being present.

A recording of 'no interaction' did not count towards the 120 interactions per observation session.

When an interaction occurred between a participating member of staff and patient, the following steps were carried out:
The interaction was recorded verbatim, including not only what was said and done, but also available information on body language and tone.
The interaction was coded using the tool.
Depending on which codes were assigned to the interaction, it was classed as positive or negative. Positive interactions were those that received codes which are considered to

support recovery, and those that received codes which are considered to hinder recovery were classed as negative.
After an interaction had been observed, the researcher moved onto the next area and the timer was re-started.

This dichotomous coding permitted the Chi-Square analysis to be used. If the researcher felt unsure about which code should be used, the interaction was left un-coded and this was done after the observation period was finished. The interactions were recorded on a matrix with five minute slots, and a space for the codes to be written and whether the code was positive or negative (see Appendix 13 for a blank version). If two dyads were interacting at the same time, the interaction between those physically closest to the researcher were recorded. The second interaction was not recorded. No personal data as to who had been involved in any interactions was kept, rather letters of the alphabet were used to denote different staff members and numbers used to denote different patients. This allowed the calculation of the number of different staff members and patients who had been involved in observed interactions. See Appendix 18 for examples of the interactions observed and recorded.

During the observations, the researcher aimed not to engage with either staff or patients. When the researcher was engaged, information was politely given that they were carrying out an observation for a piece of research which was aiming to develop a new tool. The researcher offered the PIS, or patient information sheet and explained it was perfectly acceptable for the researcher to be ignored. Any specific questions (such as who was being observed etc.) were answered. On some occasions, engagement was necessary. When this occurred, the researcher stopped timing and stopped observing for the duration of the conversation.

After the observation session had been completed, the researcher left the ward and coded any interactions left un-coded during the observation. After all of the observations and coding had been completed, the interactions were re-read and the codes and the exemplars re-checked to ensure that all had been recorded and coded satisfactorily

3.3.6 Reliability testing

To consider the reliability of the tool, it was decided that 20% of the interactions (48 in total) should be independently coded by two of the researchers. Two methods were used to achieve this:

- Both the researcher and academic supervisor recorded the same interactions and then coded them independently.
- The description of the interactions taken during the observation were typed up, with no code included. The interactions were then coded by the academic supervisor.

During the first 13 interactions observed on session 1 on ward 12, two members of the research team (the researcher and academic supervisor) carried out the observations and the subsequent coding. Two of the researchers also carried out joint observations during session 4 on ward 12, on observations 11, 14, 15, 16 and 17. This was done independently and findings then compared. When a disagreement occurred, this was discussed and an agreement reached; the differing codes were not changed. The same interactions from ward 31 were typed up and sent to the second researcher, who then coded them independently.

This lead to 39 of the interactions having been independently coded by two people using the new tool. For the final nine interactions, an online random number generator was used to generate three numbers to determine which of the interactions would be coded by a second researcher; the first number was set between 1-2, 1 referring to Ward 12 and 2 referring to ward 31. The second number was set between 1-6 which denoted which observation session the interaction should come from. The final number was between 1-20, and dictated which interaction should be used. If a combination came up which had already been rated by both researchers, it was discarded and the process repeated. The list of which interactions were used is provided in Table (iv). These interactions were typed up and given to the second researcher. All interactions which were typed up are given in Appendix 18.

Ward	Observation session	Interaction
31	6	20
12	2	18
31	5	7
31	4	13
12	2	7
12	6	13
31	3	11
31	6	10
31	6	16
31	5	4
12	4	8
31	3	6

Table (iv) Interactions which were randomly selected for rating by a second researcher.

3.3.7 Analysis

To assess the reliability of the tool, the percentage agreement for a number of circumstances which occurred over the 20% of the interactions which were coded by two members of the research team, was calculated. These were:

- Whether both researchers coded the interaction as being positive or negative.
- Whether all the same codes, and only the same codes, had been applied to the interaction by both researchers.
- Whether at least one of same codes had been applied to the interaction by both researchers.
- Whether the same code had been used by the researchers, but one researcher had also used additional codes.
- Whether none of the codes used were the same.

The Kappa value for the level of agreement on whether the interaction was positive or negative was also calculated, using Statistical Package for Social Sciences, version 23 (SPSS).

A Chi-square was used to assess if there was a statistically significant difference between the two units on the proportion of positive and negative interactions seen. 240 interactions in total were observed; this gave the study 90% power.

The interactions which had been coded as positive were transformed into ‘0’ and interactions which had been coded as negative were transformed into ‘1’. The analysis was then carried out on SPSS.

To meet the assumptions of a Chi-square, specifically that observed events are independent of each other, the research moved between communal areas of the ward as described (3.3.5.1).

Specific hypotheses (see 3.3.8) were made about the relationship between the ward having the highest or the lowest RSA-PV mean score and the proportion of positive and negative interactions seen on each ward.

3.3.8 Hypotheses

In light of the evidence base for this study, the following hypothesis about the findings of the new observational tool was made:

H1: A greater number of positive interactions will be observed on the ward with the higher staff rated Recovery-orientation than the ward with the lower staff rated Recovery-orientation.

H0: There will be no difference in the number of positive interactions observed between the two wards.

Chapter 4: Results

This chapter will report the findings from this study. For clarity and ease of reading it will be presented in two parts; part 1, the development of the tool, and part 2, the validation and reliability evaluation of the tool. Part 1 includes the Delphi method used to produce ratings of the codes by an expert panel. This was used to determine which codes were included in the tool.

Part 2 included the recruitment of eight wards, and sampling of seven which were able to take part. From the seven wards, two were selected based upon which had the highest and the lowest recovery orientation scores. Observations were carried out on the two wards, which were coded using the new tool. The reliability of the tool was examined using percentage agreements. A Kappa test was performed on the main reliability finding. The frequency with which each code was used on each ward was compared using a Wilcoxon signed rank test to assess each code's discriminant validity. The validity of the new tool was assessed by comparing the number of positive and negative interactions observed within each ward, using a Chi-square.

Throughout this chapter, all statistical and mathematical calculations were performed using the Statistical Package for the Social Sciences (SPSS) version 23.

4.1 Part 1, Development of the tool

4.1.1 Delphi method: Round 1

4.1.1.1 Sample

Within the first round of the Delphi method, 34 members of staff from across the three health boards consented to take part. Information on the following five demographic areas was collected:

- Profession
- Clinical setting worked in.
- Gender.
- Length of service in current clinical setting.
- Age range.

There was some missing data on the demographic information. Of the 34 participants, 33 completed at least three aspects of the demographic information sheet, one participant declined to answer any aspect. The demographic characteristics are presented in Table (v).

This Table provides the number and the percentage of the sample who answered each of the five areas asked about. As can be seen from Table (v) the sample was predominantly female (81.82%) from a nursing background (75.76%), aged between 18 – 45 (85.47%) who had been working in their service from between 0 -7yrs 11months (84.38%).

	Number of responses to Phase 1 of the Delphi demographic characteristics.	% of respondents
Profession	(33)	
Nursing	25	75.76
Clinical Psychology	4	12.12
Occupational Therapist	2	6.06
Doctor (medical)	1	3.03
Clinical Lead	1	3.03
		Total % (100)
Clinical Setting	(33)	
Mental health (rehab/recovery)	26	78.79
Forensic	7	21.21
		Total % (100)
Gender	(22)	
Female	18	81.82
Male	4	18.18
		Total % (100)
Length of Service	(32)	
0 – 3 years 12 months	15	46.88
4 years – 7 years 12 months	12	37.5
8 years – 11 years 12 months	3	9.38
12 years – 15 years 12 months	1	3.13
30 years – 33 years 12 months	1	3.13
		Total % (100)
Age Range	(33)	
18-30	6	18.18
31-45	22	66.67
45-60	5	15.15
60+	0	0
		Total % (100)

Table (v) Demographic characteristics of panel for Round 1 of Delphi.

4.1.1.2 Rating of the Codes on Relevance and Importance: Round 1 of the Delphi study.

4.1.1.2.1 Relevance

For the Relevance scale, the number of codes that participants rated ranged between 22–26 (from a total of 26). The number of codes that each participant rated on the Relevance scale is given in Appendix 19. The majority of the sample (85.29%) rated all 26 codes. When considering each code on the scale, the range of responses for each code was between 31 and 34 responses, with code 7 having the fewest responses. The number of participants rating each code is given in Figure (i).

4.1.1.2.2 Importance

For the Importance scale, the number of codes that the participants rated ranged between 10–26 (from a total of 26). The number of codes that they rated on the Importance scale is given in Appendix 20. Over half of the sample (67.65%) rated all 26 codes; fewer than rated all codes on the Relevance scale. When considering each code on the scale, the number of responses for each code ranged between 27 and 34 responses, with the fewest participants rating code 10. The number of participants who rated each code is shown in Figure (i).

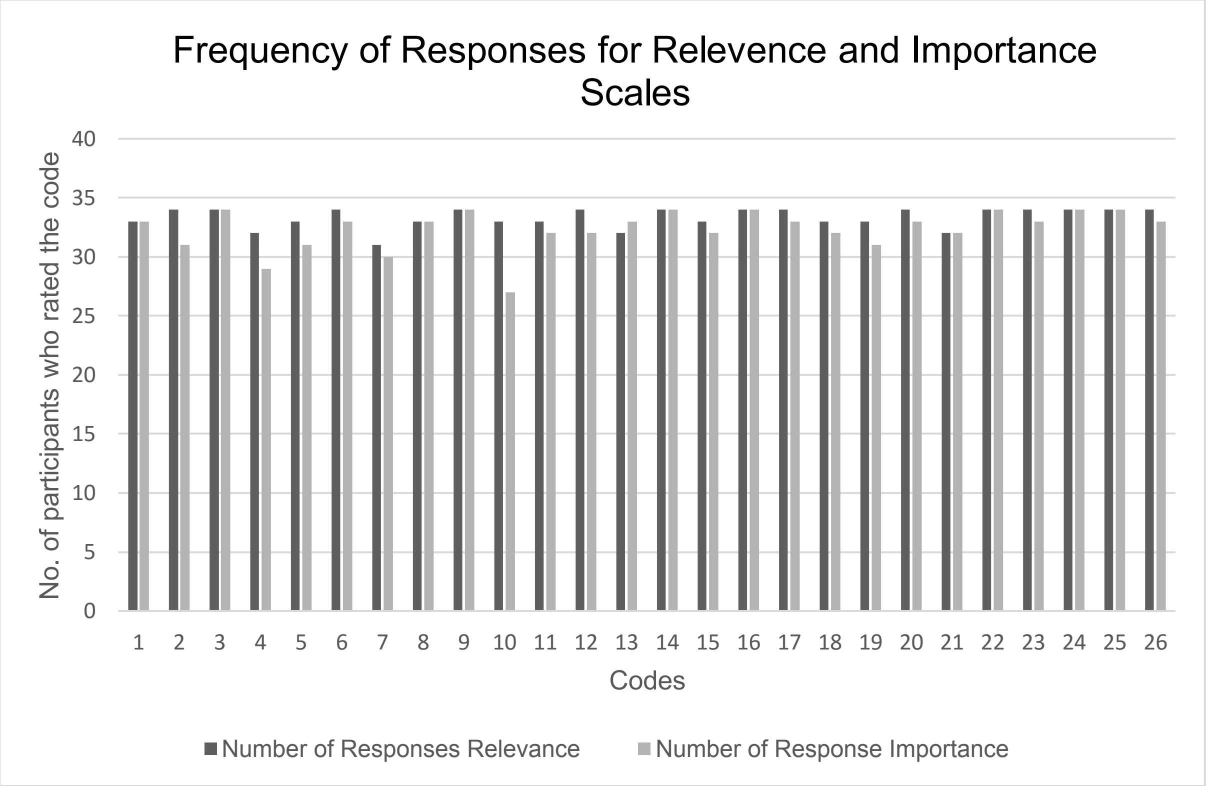


Figure (i). Frequency of Responses for each code on Relevance and Importance Scales.

4.1.1.3 Item Analysis for Relevance and Importance scales: Round 1

It was decided that those codes for which 75% of the panel scored 6 or 7on, either the Relevance or the Importance scale, would be included in the tool, while those codes which 75% of the panel scored either 1 or 2 on either scale would be discarded from the tool. All other codes would be re-rated in the second round. These decisions were based upon a literature review which covered the use of the Delphi method (see section 3.2.7). The percentage of the sample who scored each code 1 and 2, and 6 and 7, and the results for each code is presented in Figure (ii) for Relevance and Figure (iii) for Importance. Appendix 21 provides the scores for each code, and whether this led to them being included, excluded or re-coded in Round 1 on the Relevance scale. Appendix 22 provides the same information for the Importance scale. For full SPSS output for each code, please see Appendix 23.

On the Relevance scale, 11 of the codes were included in the tool, and 10 of the codes were included on the Importance scale. The same codes were included on both scales, apart from code 11, which was included based upon the Relevance scale, and would have been re-rated based upon the Importance score. All other codes on both scales were rated so as to be re-rated in Round 2. Neither scale produced any code which the panel determined should be omitted from the tool.

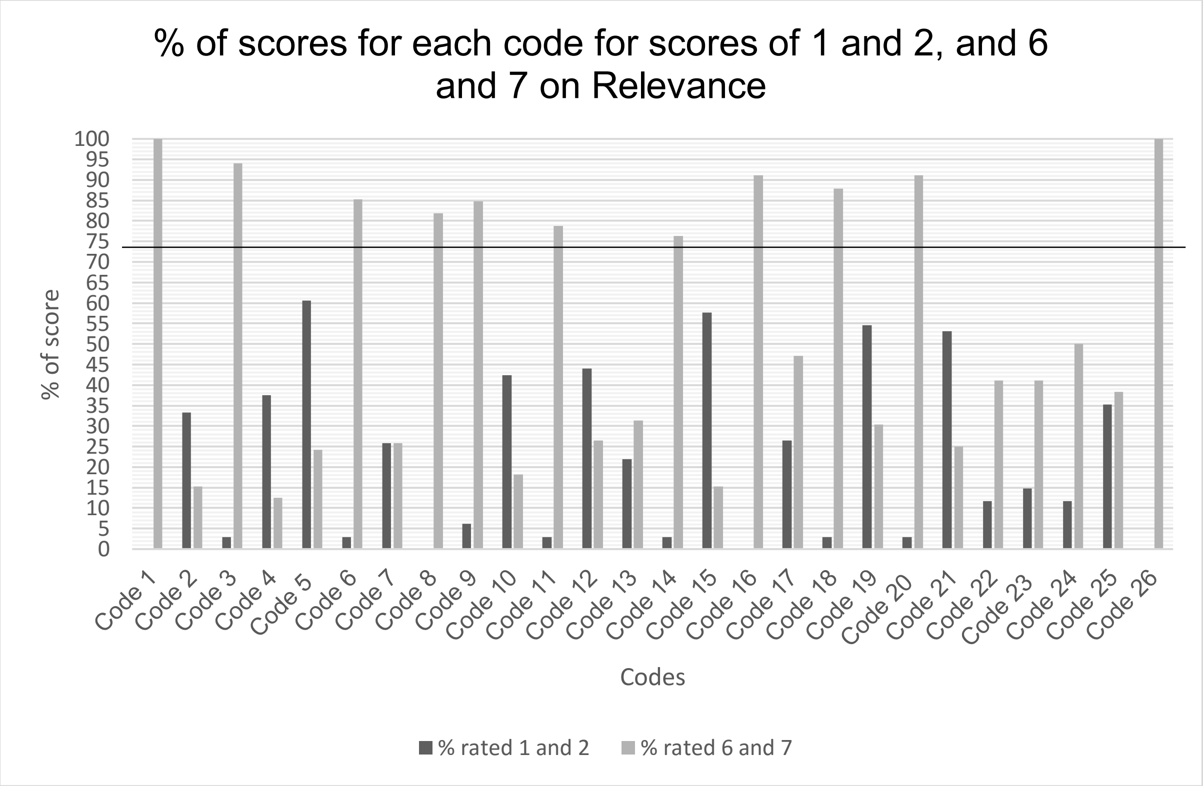


Figure (ii) Percentage of the sample who rated each code 1&2, and 6&7 on the Relevance scale.

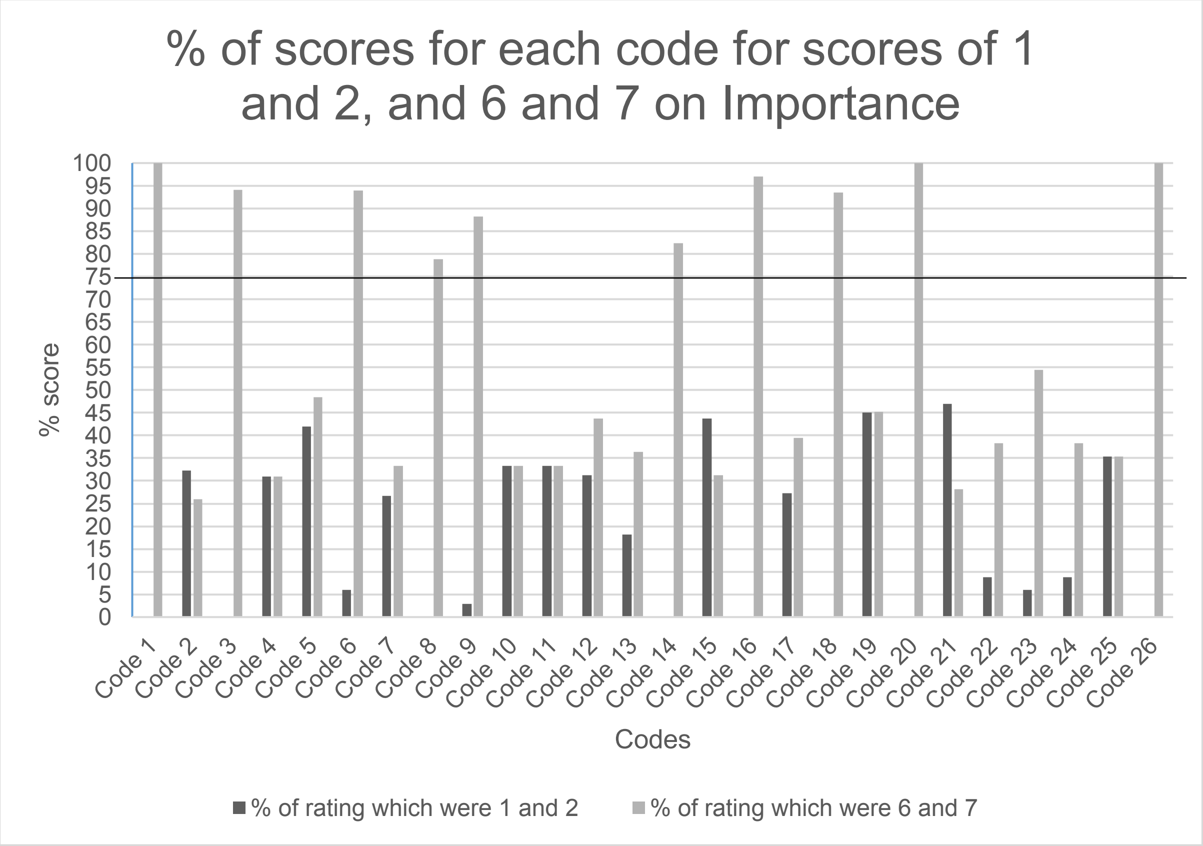


Figure (iii). Percentage of the sample who rated each code 1&2, and 6&7 on the Importance scale.

4.1.2 Delphi method, Round 2

4.1.2.1 Sample

Eleven participants from the original panel took part in the second round of the Delphi survey. This was due to pragmatic reasons: a shorter period (three weeks) was given for data collection. This panel was made up from two of the three health boards in the original panel. Demographic information was collected on less than half of the second panel, and consequently has not been included here.

4.1.2.2 Additional Codes

In Round 1, in addition to rating the codes the research team had created, the panel was asked to add any codes they considered should be included from their clinical experience. They were also asked to generate any further exemplars of the codes from clinical experience.

No additional exemplars were generated. However, four additional codes were generated, by a single participant. These were:

Code 27: PRN medication offered as a ‘quick fix’/alternative to talking.
Code 28: Empathy shown in response to distress.
Code 29: Doing with, not doing for.
Code 30: Asking how a client would like to deal with a problem; exploring options.

4.1.2.3 Rating of the Codes on Relevance and Importance, Round 2

The codes which had been identified to be re-rated in the first round of the Delphi survey and the newly generated codes, were provided to the 11 participants in Round 2. Once again, the panel was asked to rate each code on the two scales; Relevance to the service i.e. whether there were interactions which were likely to occur, and Importance, i.e., the impact such interactions could have upon the quality of care within these settings. After three weeks, all responses which had been gained were analysed. All of the 11 participants who took part in the second round rated every code on both scales.

4.1.2.3 Code Analysis for Relevance and Importance, Round 2

Descriptive statistics for the ratings for both scales which had been provided by the participants were calculated. In this round, those codes which were scored 5,6 or 7 by 80% of the panel on either the Relevance or the Importance scale were included. Any code which was not rated as 5, 6 or 7 by 80% of the panel on either scale was excluded. A full SPSS break down of the codes’ scoring is provided in Appendix 24. The percentage of each code rated by the sample was computed to determine if the code should be included in the tool. In the second round, those codes which were not rated sufficiently high enough by the panel were excluded.

The percentage of participants who scored each code as 5, 6 or 7 is presented in Figure (iv) for Relevance and Figure (v) for Importance. Appendix 25 provides the percentage of participants who rated each code 5,6 or 7, and the decision made in light of this on the Relevance scale. Appendix 26 provides the same information for the Importance scale.

On the Relevance scale, only one code was rated so as to be included, code 28, the other 18 codes received ratings which indicated they should be excluded. However, on the Importance scale, 12 of the 19 codes were rated so as to be included (code 28 was included on the Importance scale as well). Therefore, all 12 codes were included in the final tool.

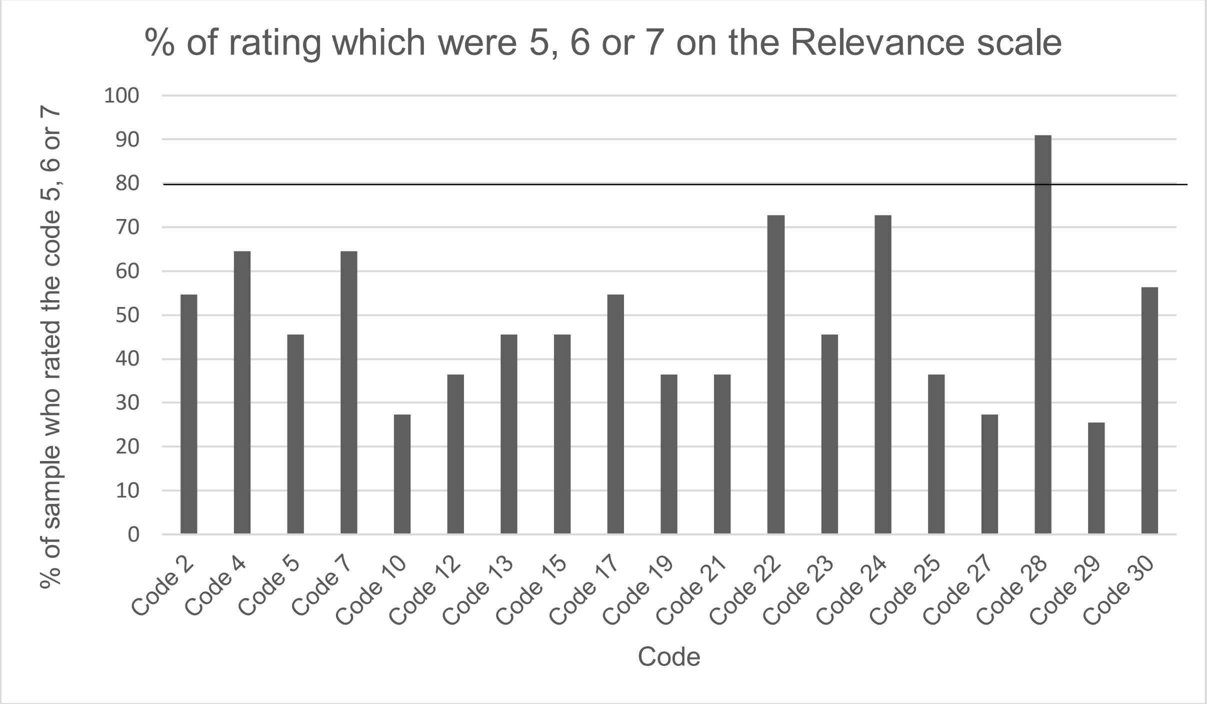


Figure (iv) Percentage of sample who rated each code 5,6 or 7 on the Relevance scale.

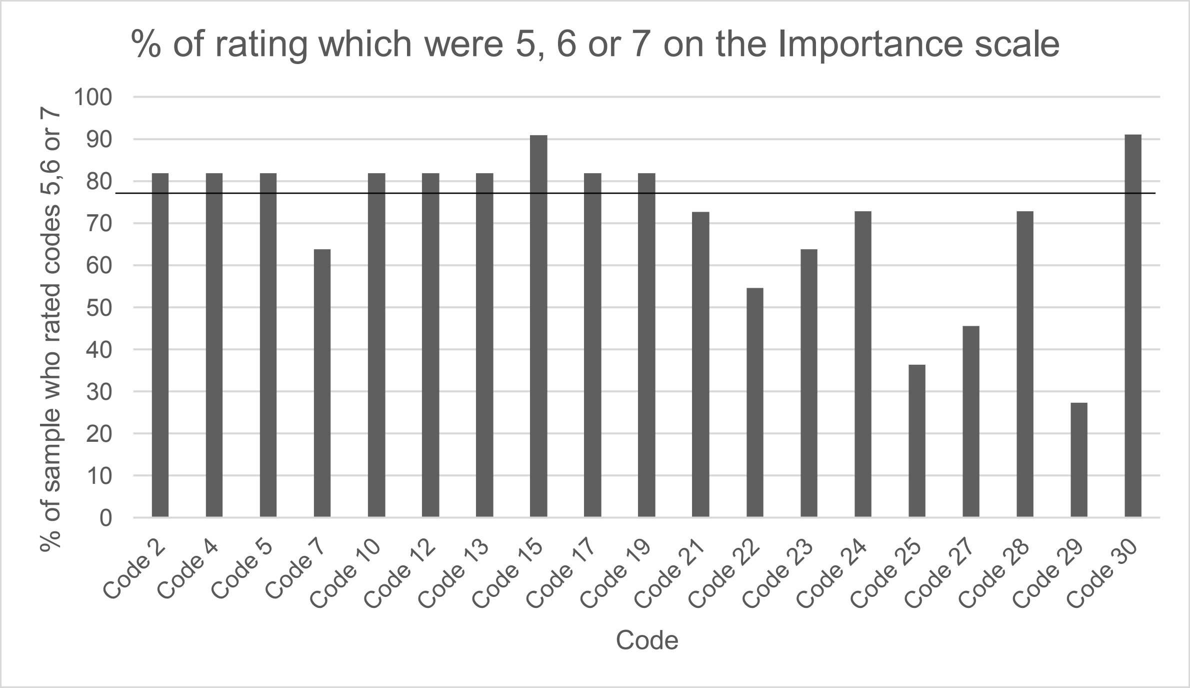


Figure (v) Percentage of sample which rated each code 5, 6 or 7 on the Importance scale.

Using the rating from both the first and the second round, the codes were divided into those which would be used to construct the final tool, and those which were to be discarded. The codes which were included, (and in which of the rounds of the Delphi method that inclusion took place) are given in Appendix 27; those which were excluded (and in which round that occurred) are given in Appendix 28.

26 codes were initially developed by the team, and an additional four codes were added by the panel, resulting in a total of 30 codes being rated by the panel. Of those 30 codes, 23 were selected by the panel and one was removed by the research team, leaving 22. Of the final 22 codes, 12 were considered to ‘support’ the recovery process, and 10 to ‘hinder’ this process in accordance with the initial development based upon the recovery literature (Bonney & Stickley, 2008; Leamy *et al.,* 2011). These were the codes by which the observed interactions were rated in Part 2 of the study.

Code 24 (Set getting up times) was the only service level code which was included in the final tool, based upon the panel’s judgement. The service level codes had been designed to be used together. As a single code, it was decided code 24 could not add any additional information to the tool and was therefore removed.

4.2 Part 2. Validity and Reliability testing of the Tool

4.2.1 Identification of wards for Observations

A sample of eight wards which fitted the inclusion criteria were recruited to the study. However, having agreed to take part, it became apparent one of the wards was not able to participate. Therefore, the sample was made up of seven wards. From these eight wards, two were identified: the ward with the highest and the ward with the lowest scores on the standardised measure of recovery. They were identified using a standardised self-report measure of the recovery orientation of the service (Recovery Self-Assessment, Providers Version [RSA-PV]). The mean score for the RSA-PV from each ward was calculated and the ward with highest overall mean and the lowest overall mean, identified. These wards were selected for the tool to be piloted in.

4.2.1.1 Sample

Seven rehabilitation and recovery inpatient services from across three health boards took part in the second stage of the study; participants were recruited from members of staff on these wards. To ensure anonymity, each Health board were given a number, 1, 2 or 3. Each of the wards were then numbered, by employing the first digit in the number relating to the Health board they came from, and the second identifying each ward. Thus the wards were; 12 and 13; 21 and 22; 31, 32, and 33.

The total number of participants across the seven wards involved at this stage of the research was n = 57.

The number of participants per ward ranged between n= 6 (ward 31) to n = 10 (wards 12 and 33). The number of participants from each ward is given in Appendix 29.

Demographic information was collected from participants at this stage, asking:

- Profession
- Gender
- Length of service in current clinical setting
- Age range

The demographic information from each ward is presented in Appendix 29. Demographic information from across the seven wards is given in Table (vi). There was a total of 15 items of data missing from the demographic information provided; profession (3 times), gender (5 times), length of service (3 times), and age range (4 times).

	Number of participants	% of the whole sample	% of the sample who completed the demographic information
Profession			**54**
Nursing	22	38.59%	40.74%
Clinical Psychologist	3	5.26%	5.55%
Ward Manager	2	3.5%	3.7%
Health Care Assistant	17	29.82%	31.48%
Occupational Therapist	4	7.01%	7.4%
Activity Co-ordinator/Therapies Co-ordinator	2	3.5%	3.7%
Clinical Lead	2	3.5%	3.7%
Psychiatrist	2	3.5%	3.7%
Prefer not to say/did not complete	3	5.26%	N/A
Total	57	100%	100%
Gender			**52**
Female	33	57.89%	63.46%
Male	19	33.33%	36.53%
Prefer not to say/not completed.	5	8.77%	N/A
Total	57	100%	100%
Length of time at service			**53**
0 – 3 years 12 months	26	45.61%	49.05%
4 years – 7years 12 months	15	26.31%	28.30%
8 years – 11 years 12 months	6	10.52%	11.32%
12 years – 15 years 12 months	3	5.26%	5.66%
16 years -19 years 12 months	0	0%	0%
20 years – 23 years 12 months	0	0%	0%
24 years 27 years 12 month	1	1.175%	1.88%
28 years – 31 years 12 month	1	1.175%	1.88%
32 years – 35 years 12 months	1	1.175%	1.88%

Prefer not to say/not completed	4	7.01%	N/A
Total	57	100%	100%
Age			**53**
18-30	11	19.28%	20.75%
31-45	24	42.10%	45.28%
46-60	16	28.07%	30.18%
60+	2	3.5%	3.77%
Prefer not to say/not completed	4	7.01%	N/A
Total	57	100%	100%

Table (vi) Demographic information from part 2 of the study, ward selection phase.

As can be seen from Table (vi), of the sample who did provide the demographic information, it was a predominately female sample (63% female). The most common profession was nursing, representing 40% of the sample. The most common age range was between 31-45 years old (42%), and the most common length of service was between 0 and 3years 11 months (49%).

Comparing the wards, some differences are apparent: ward 31 had fewer nurses, four of the six participants were Health Care Assistance (HCAs) (66% of the wards sample) compared to one nurse (16% of the wards sample). The ward with the greatest difference in gender was ward 12, where eight of the 10 participants were female (80%), whereas ward 22 had more males than females (57% male). In terms of length of service, ward 21 differed from the sample as a whole, as none of the sample had worked for less than four years. The majority having worked for between 4years and 7years 11 months (55%).

4.2.1.2 Missing data

Two participants had a group of scores missing on the RSA-PV, and were therefore discarded from the analysis. Therefore, the overall number of participants whose data was used decreased to n=55, and the sample size for ward 12 and ward 33 both decreased to n=9 from n=10.

As the data was anonymous, it was not possible to remove these participants' information from the demographic information collected.

4.2.1.3 Analysis of differences between wards

The data from the RSA-PV were input blind in order for the wards with the highest and lowest scores to be concealed from the researcher. For each ward, the mean was calculated for the total RSA-PV score.

The mean scores for each ward ranged from 128.22 (ward 12) to 142.67 (ward 31) and as such these wards were selected for the observation stage to be completed in. The range of possible scores from the RSA-PV is 36 -180. The Standard Deviations ranged from 11.74 on ward 32, to 23.06 on ward 12. The Coefficient of Variation on all wards was under 1, which is considered low-variance (Brown, 1998). The mean total scores, score range, standard deviation and coefficient of variation for each ward are present in Table (vii).

Ward	Mean RSA	Range	Standard Deviation (SD)	Coefficient of Variation (CV)
12	128.22	77-157	23.06	0.18
13	136.25	113-153	12.43	0.09
21	136.78	107-159	15.97	0.12
22	130.86	95-156	21.73	0.16
31	142.67	110-174	22.66	0.16
32	135.63	119-152	11.74	0.08
33	138.22	112-154	12.62	0.09

Table (vii) Mean, range, SD and CV of the RSA-PV scores for each ward.

4.2.2 Observations

4.2.2.1 Sample

The sample size for this phase of the study was n= 240. The sample size was dictated by the number of interactions observed. A Chi-square power calculation was performed. This indicated that using a total sample size of 240 gave 90% power.

Those involved in the interactions observed were staff members on shift during the observations periods in the two wards selected. Data collection followed the procedure given in section 3.3.5.2.1

The number of different members of staff and patients who were involved in a recorded interaction was noted at each observation session. On ward 12, the number of different staff members involved in an interaction within an observation session ranged from 6 – 8, and the number of different patients with whom staff interacted with ranged from 5 – 10. On ward 31, the number of staff members involved in interactions within an observation session range from 6 – 11, and the number of different patients with whom staff interacted ranged from 6 – 10. This is outlined in Table (viii).

	Ward 12		Ward 31	
	Number of Staff	Number of Patients	Number of Staff	Number of Patients
Session 1	7	9	9	9
Session 2	8	8	8	8
Session 3	6	10	9	10
Session 4	7	9	7	9
Session 5	8	5	11	6
Session 6	8	9	6	10

Table (viii) Number of different Staff members and Patients involved in at least one interaction in each observation session.

4.2.3 Analysis of the tool

4.2.3.1 Reliability analysis

The tool's reliability was measured; see method section 3.3.6. for details. 20% (48) of the interactions were rated by two researchers. Of the 48 interactions rated, whether the interaction was positive or negative, there was agreement on 45 of those interactions. In terms of which codes were assigned by each researcher, in 43 of the cases, at least one of the codes assigned was the same for both researchers. Of these 43 cases, there was total agreement (both parties using all the same codes) on 25 of the cases. In a further 18 of the interactions, although the researchers put at least one of the same codes, one researcher also put an additional code or codes. These numbers and the percentage they represent are given in Table (ix). The Cohen's kappa (k) was calculated for the percentage agreement of the positive and negative ratings of the interactions. (k) = .827, $p<.000$. Based on the guidelines from Altman (1999), a kappa agreement of .81-1 indicated a strength of agreement of 'Very Good.' As the (k) found here is within this range, the strength of the agreement is very good.

	Number	Percentage (of 48)
No. of cases where positive/negative rating was in agreement.	45	93.8%
No. of cases where all codes were in agreement.	25	52%
No. of cases where at least one code was in agreement.	43	89.6%
No. of cases where initial code was in agreement, but one researcher had an additional code.	18	37.5%
No. of cases where no codes were in agreement.	5	10%

Table (ix) Percentage agreement between testers of observed interactions.

4.2.3.2 Analysis of the validity of the tool

The data was analysed to assess if the tool in its current state was able to differentiate between the two wards that had scored the highest and the lowest on a standardised measure of recovery orientation of a service. This was done using a Chi-square.

The number of codes assigned to each interaction was investigated. This was investigated to establish if the ’quality’of the interactions differing between wards, based upon how many codes were needed to capture the interaction seen. Which codes were used to describe the interactions on each ward was also analysed. This was done to assess the discriminatory validity of the codes and to consider if there was a relationship between which codes where used and their Importance scores from the Delphi survey.

4.2.3.2.1 Descriptive data

Each interaction was coded, using as many of the codes as appropriate. The total numbers of positive and negative codes used for each ward is displayed in Figure (vi). This was done to assess if there was a marked difference in the number of either positive or negative codes seen between each ward, as the main analysis (a Chi-square) compared the number of positive and negative interactions seen, without regards to the number of positive or negative codes used.

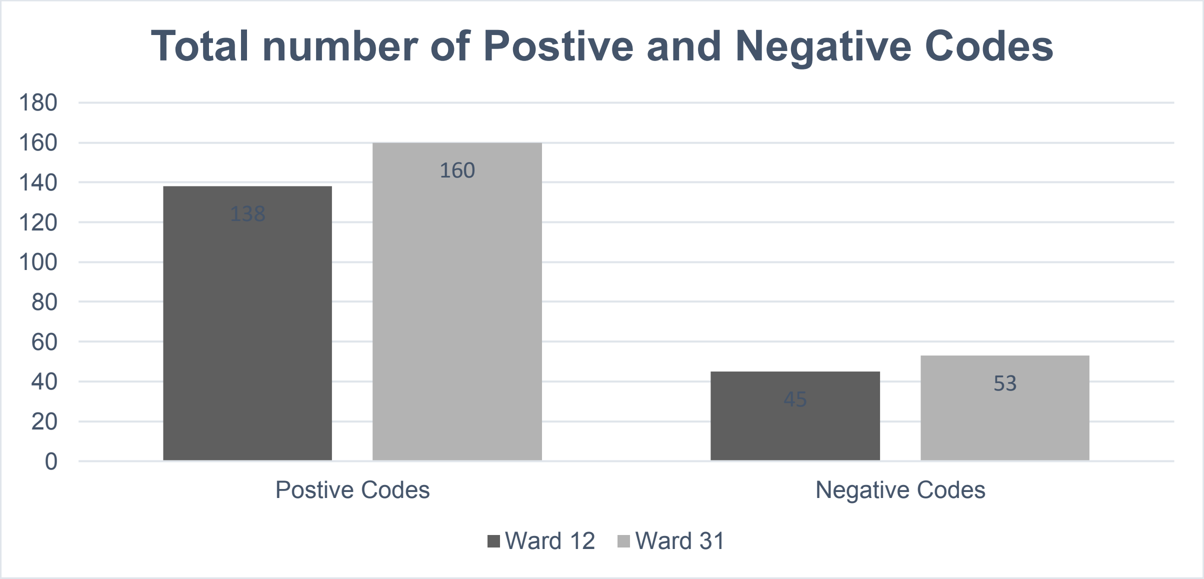


Figure (vi) The total number of positive and the total number of negative codes used for each ward.

The number of codes assigned to each interaction observed varied between 0 – 6 for positive codes, and 0-3 for negative codes. To assess if there was a significant difference in the number of codes assigned to a single interaction between the wards, the frequency with which the number of codes were assigned to an interaction was calculated. This is displayed in Figure (vii) for positive codes, and Figure (viii) for negative codes. In every interaction observed, the most frequent outcome was that only one code was applied, irrespective of whether it was a positive or a negative.

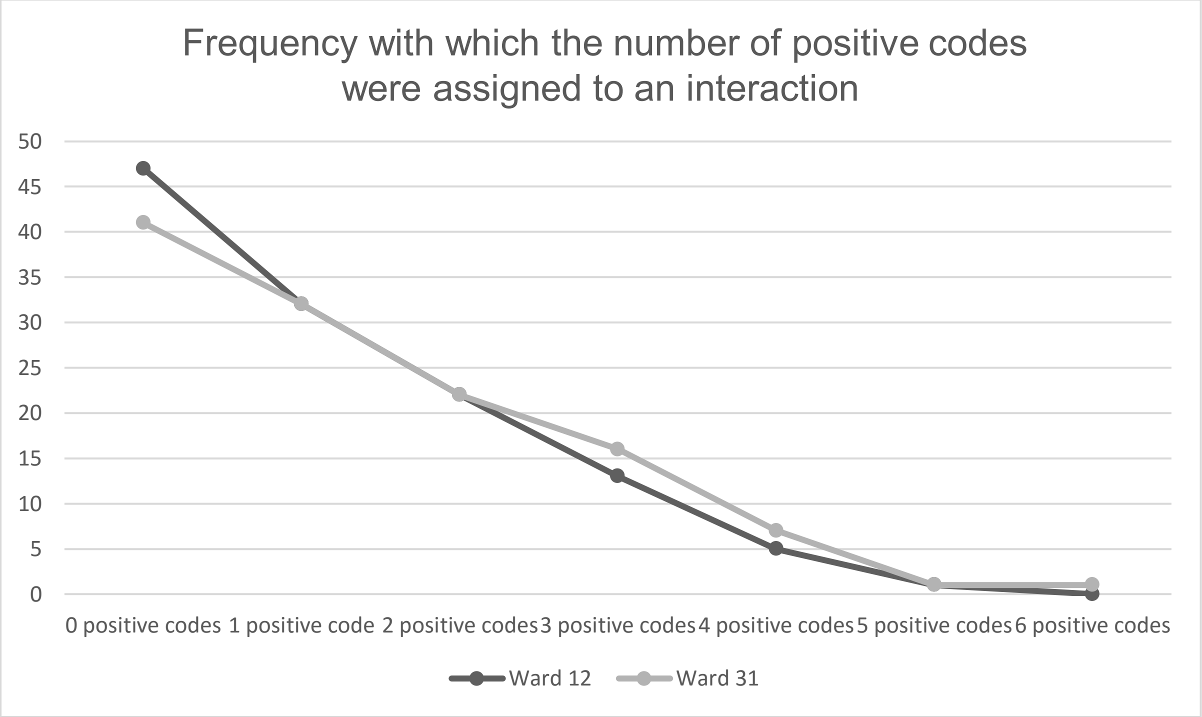


Figure (vii) the number of times 0,1,2,3,4,5 or 6 positive codes were used to describe one interaction.

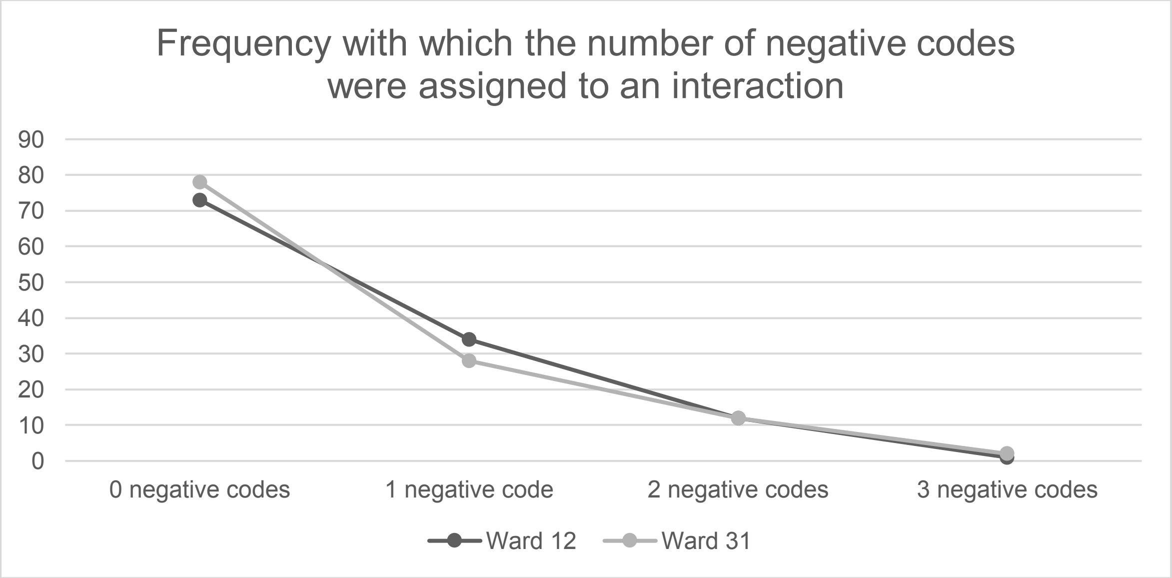


Figure (viii) the number of times 0,1,2 or 3 negative codes were used to describe one interaction.

As can be seen from the two graphs, the frequency with which the number of positive and negative codes were applied to an interaction was very similar between both wards.

4.2.3.2.2 Discriminant validity

To examine the discriminatory validity of the codes within the tool, the Importance data from the Delphi survey in part 1 was used, and the frequency with which the codes were used in each ward was examined.

The number of times each code, which was used across the 120 observations, was calculated and presented in Figure (ix). The most frequently used code 8 (Staff initiated non-clinical interactions) which was used 68 times across both wards. The only code which was not used on either ward on any occasion was code 30 (Asking how a client would like to deal with a problem; exploring options).

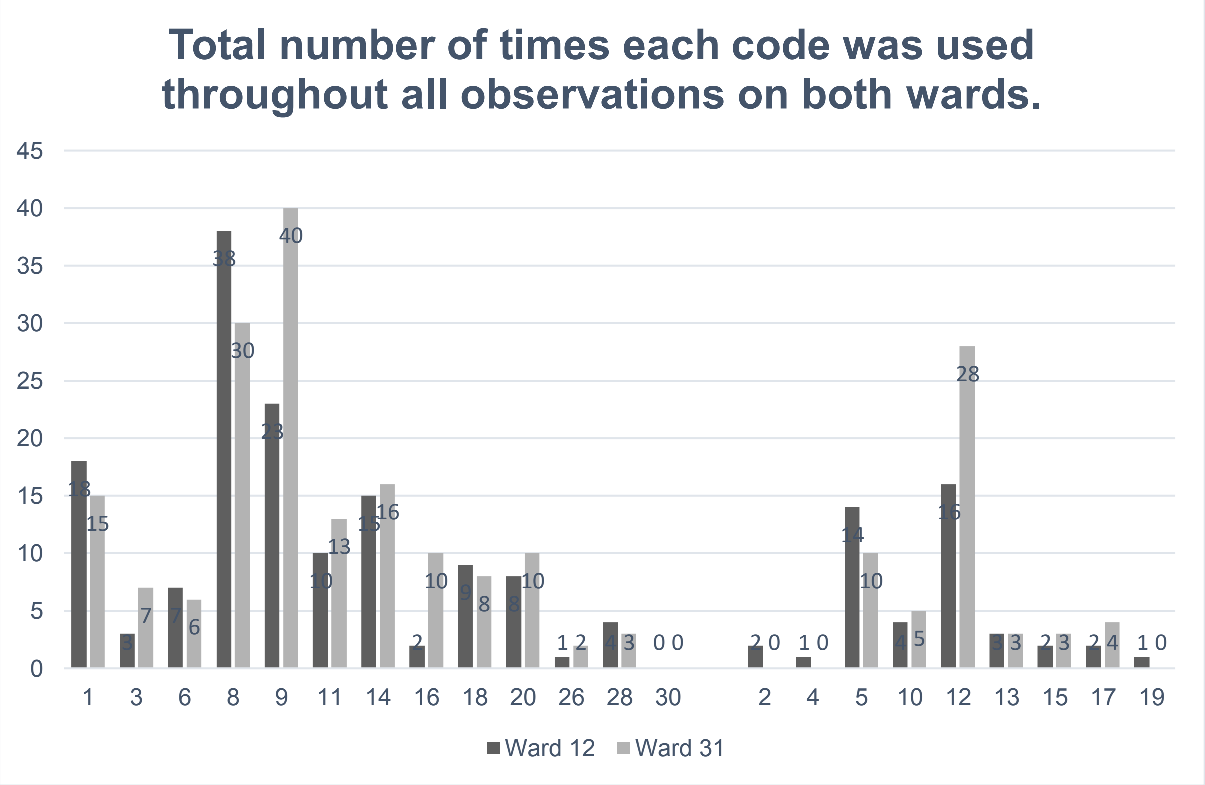


Figure (ix) Number of times each code was used to describe an interaction across the 120 observed per ward.

The mean Importance score from the Delphi survey was calculated for each code within the tool. The codes were then ranked, based upon their importance. For the codes which support recovery, the Importance rating ranged from 7 (code 12) to 4.1 (code 11), the mean score and the ranking are provided in Table (x). For the codes which hinder recovery, the mean Importance score ranged from 6.5 (code 12) to 5.6 (code 2). the mean score and the ranking are provided in Table (xi).

Code	Mean score	Rank	Frequency of use on Ward 12	Frequency of use on Ward 31
26	7	1	1	2
1	6.9	2	18	15
16	6.9	2	2	10
20	6.9	2	8	10
3	6.8	5	3	7
18	6.8	5	9	8
6	6.7	7	7	6
14	6.6	8	13	16
9	6.5	9	23	40
8	6.3	10	38	30
30	5.9	11	0	0
28	5.4	12	4	3
11	4.1	13	10	13

Table (x) Mean Importance ratings and rank of codes which support recovery.

Code	Mean score	Rank	Frequency of use on Ward 12	Frequency of use on Ward 31
12	6.5	1	16	28
19	6.3	2	1	0
15	6.3	2	2	3
13	6.2	4	3	3
4	6	5	1	0
17	6	5	2	4
5	5.9	7	14	10
10	5.9	7	4	5
2	5.6	9	2	0

Table (xi) Mean Importance ratings and rank of codes which hinder recovery.

As can be seen from the tables, the mean Importance scores are generally higher on the codes which support recovery. The range of mean scores is also greater for the codes which are supportive of recovery. When reviewing the frequency with which the codes were used for each ward, there is not any obvious difference. This has also been assessed statistically.

It would have been preferable to examine the frequency of each code between wards individually. This would have allowed greater scrutiny of the sensitivity and specificity of each code. However, this would have involved 22 independent variables (as there are 22 codes), which was inappropriate given the quantity of data. Therefore, a Wilcoxon signed rank test was performed on the codes which support recovery, and those which hinder recovery to assess if there was any statistically difference between wards.

For the codes which support recovery, a Wilcoxon signed rank test indicated that there was not a statistically significant difference between codes which were used to describe the interactions between the two wards $Z = -1.105$, $p < .269$. For the full SPSS output, please see Appendix 30.

For the codes which hinder recovery, a Wilcoxon signed rank test indicated that there was not a statistically significant difference between codes which were used to describe the interactions between the two wards $Z = -.071$, $p < .943$. For the full SPSS output, please see Appendix 31.

4.2.3.2.3 Statistical testing of the Validity

A Chi-square test of independence was performed to examine the relationship between the number of positive interactions seen between the two wards which had been differentiated, based upon mean RSA-PV scores.

The relationship between these variables was not significant;
X2 (1, N = 240) = .646, p= .422

Therefore, the null hypothesis, is accepted; There was no difference in the number of positive interactions observed between the two wards.
For the full SPSS output please see Appendix 32.

Chapter 5: Discussion

5.1 Introduction

This chapter will provide a summary of the results obtained, followed by an interpretation of such in the context of the methodological choices, as discussed in Chapter 3 (Method) and the literature on recovery, measures of such and inpatient care, as discussed in the Chapter 1 (Introduction). This chapter will then present the strengths and limitations of the current study, followed by the clinical implications of the work done to date and recommendations for future research, involving the tool designed. Finally, overall conclusions from the current study will be presented.

5.2 Summary of Results

The aim of the study was to carry out the initial development, reliability and validity testing of a tool for assessing the recovery orientation of Rehabilitation and Recovery inpatient mental health services. The importance of recovery orientation within mental health services is recognised by stake holders and by policy (see section 1.2); however its implementation faces several barriers, one of which is being able to meaningfully assess how services support personal recovery (CSIP, 2007; Donnelly *et al.,* 2011). A number of measures exist (see section1.3), however these have received criticism (e.g. Burgess *et al.,* 2011; Shanks *et al.,* 2013). Reviewing the measures captured within available compendia, three areas of weakness appeared:

- The conceptual understanding of recovery is diverse, and therefore what domains various tools measure is inconstant.
- What is measured to assess recovery varies, processes which occur that support or hinder the recovery journey are not closely scrutinised.
- All tools are self-report measures.

There appears to be a gap for a measure which assesses the processes or modes by which support of the recovery orientation might be delivered within services, which is not reliant upon self-report techniques. With this in mind, a four-stage process was employed to design and test a new tool to assess the recovery orientation of services, using an observational method, focusing upon the manner in which members of staff interacted with patients.

5.2.1 Development of the initial codes

The first stage involved the research team examining the themes, which have been captured by research and identified as proximal to the personal recovery process (Bonney & Stickley, 2008: Leamy *et al.,* 2011), see section 1.2.1 for further details. Using the themes as a basis and the team's clinical expertise, the research team developed a number of codes and exemplars of the codes. This produced 26 codes. Four of these codes related to service characteristics and were not classed as supporting or hindering recovery. Of the remaining 22 codes, using the evidence from the recovery literature, 11 were considered to support the recovery process and were termed positive, and 11 were deemed to hinder recovery and termed negative.

5.2.2 The Delphi method

The second stage of the project involved gathering a consensus as to which codes should be used within the new tool. This was achieved by using the Delphi method over two rounds. A panel of 34 experts was created from professionals working in inpatient Rehabilitation and Recovery services across three health boards. Within the first round, the panel was asked to rate each of the codes on its Relevance to the services within which they worked and on its Importance to the recovery process, relating to how much of an impact they felt these had. They were also asked to provide any codes they believed had not been developed but should be included, and further examples of the codes provided from their experience. From this, four new codes were generated (from one participant). After the first round, the panel had agreed on 11 codes (all positive) to be included in the tool. The remaining 11 had been rated to be reconsidered in the second round. Within the second round, due to time limitations, the panel was made up of 11 from the original 34. The panel was asked to reconsider the codes they had already reviewed and consider the four new codes added after Round 1. From Round 2, 12 of the remaining 19 codes were included in the tool, the other codes were discarded. Throughout both rounds of the Delphi method, 23 of the 30 codes presented to the panel were included in the final tool.

5.2.3 Selection of wards for piloting of the new tool

In order to assess the tool's validity, the tool was used in two wards with differing recovery orientations, as assessed by a standardised measure of the recovery orientation of a service (the RSA-PV). Eight wards were initially recruited into the study; however, it became evident one ward was unable to take part. Therefore, staff members from seven wards were recruited to take part in the study. Participants on each ward completed the RSA-PV and the mean score was calculated for each ward. This produced scores ranging from 128.22 to 142.67 (see Table (vii) for details). The two wards with the highest (ward 31) and the lowest (ward 12) mean scores were used to pilot the tool in.

5.2.4 Observations: Reliability and Validity testing of the new tool

To assess the tools reliability and validity, the tool was piloted in the two wards identified. The researcher carried out observations spanning the working day, over a two-day period. One-hundred and sixty interactions per ward were observed. To assess the reliability of the tool, 20% of the interactions were coded by two members of the team, and the level of agreement compared. This was done via two methods. Eighteen of the interactions were observed by both the researcher and the academic supervisor, these were then coded independently and the results compared. The remaining 30 interactions which were used to check reliability, were observed by the researcher alone and then typed up from the observation matrix without including any codes. The academic supervisor then coded these interactions. The coding performed by the researcher and the academic supervisor were compared to consider:

- Whether both coded the interaction as being positive or negative.
- Whether all the same codes, and only the same codes, had been applied to the interaction by both researchers.
- Whether at least one of same codes had been applied to the interaction by both researchers.
- Whether the same code had been used by the researchers, but one researcher had also used additional codes.
- Whether none of the codes used were the same.

In 93.8% of the interactions there was agreement as to the positive/negative rating. This gave a Kappa rating of .827, which indicates a strength of 'Very Good' (Altman, 1999). In 52% of interactions, there was agreement on all of the codes that had been used. In 89.6% of interactions there was agreement on at least one of the codes that had been applied. In 37.5% of interactions, both researchers agreed on one code, but one researcher had applied an/some additional codes, which the other researcher had not. In only 10% of the interactions used for reliability testing was there no agreement between researchers, as to which codes should be applied. This indicated the reliability of the newly developed tool as being good.

To assess the validly of the tool, a Chi-square was used to compare the number of positive and negative interactions observed between the two wards. The hypothesis made was that there would be a greater number of positive interactions seen on the ward which had the higher RSA-PV mean score. The Chi-square did not find a significant result; X^2 (1, N = 240)

= .646, *p*= .422. This indicated that the current tool was not able to differentiate between the two wards identified.

5.3 Discussion of findings

5.3.1 Initial development of the codes

5.3.1.1 Literature base used

The empirical basis of the study relied heavily upon the Model of Recovery presented by Leamy *et al.,* (2011)**,** in addition to the themes identified by Bonney and Stickley (2008) in their review. This was most directly employed during the initial development of the codes.

The model of recovery presented by Leamy *et al.,* (2011) was part of a wider National Institute for Health Research funded study. They used an extensive search to identify papers that described or developed a conceptualisation of personal recovery from mental illness. Electronic databases, hand-searching and a web search were all completed. A robust quality rating process was used, completed by three members of the team, then moderated by a further two, using different quality frameworks for the qualitative and quantitative papers (RATs and Effective Public Health Practice Project (EPHPP, 2009) respectively).

The data was analysed using a modified narrative synthesis. This process developed an initial conceptual framework, which was more heavily based on the papers with the highest quality ratings than the others in the review. Two methods were employed to assess the robustness of the synthesis. Firstly, qualitative studies which were rated as moderate quality were thematically analysed until category saturation was achieved, and the resulting themes were then compared with the preliminary conceptual framework developed. The second method used an expert consultation panel to whom the preliminary conceptual framework was sent. The panel comprised 54 advisory committee members of the REFOCUS Programme, including those with academic, clinical or personal expertise of recovery. The preliminary conceptual framework was modified in response to these comments, producing the final conceptual framework (Leamy *et al.,* 2011).

There are several particular strengths to this piece of research. The use of a quality frame work, which was modified by more than one researcher, indicated that the quality of the papers was well reviewed. This piece of work then utilised the information by placing more emphasis on the information gained from the higher quality studies. The processes of validation for the conceptual framework were also thorough, using both the empirical base and exploring the face validity with experts in the field who represented a range of stakeholders. This adds validity to the framework developed.

However, there were limitations. Conceptually the model presented is limited because, although it synthesises the current literature, it cannot be seen as definitive, due to the nature of recovery and a disconnect been recovery and research principals. One of the fundamental scientific challenges of research and development within recovery is that the philosophy gives primacy to individual experience and meaning, whereas mental health systems and the current dominant scientific mode, give prominence to group-level aggregated data. In terms of the technical limitations this imposes on the work by both Leamy *et al.,* (2011) and recovery research more generally, research is primarily focused at the bottom of the hierarchy of evidence (Onken, Craig, Ridgway, Ralph, & Cook, 2007).

Methodologically, the work also presents some limitations. Firstly, the narrative synthesis approach was modified, which the authors agree could have been widened. Secondly, the emergent categories were only one way of grouping the findings. That opinions are likely to differ is highlighted by the categories being changed as a result of expert consultation. The categorical separation provides structure, and the logic of the choices is evident, however studies which follow may not create the same overall thematic structure. When reviewing the analysed data, the analysis synthesised the interpretation of the primary data provided in each paper, rather than utilising the primary data directly. Analysis of primary data may have allowed a stronger voice for certain stakeholders, such as consumer researchers, or for different stakeholders to be directly compared. And finally, the systematic review highlighted the dominance of recovery literature originating from the USA. This is valuable data, and a result of where the research is being done, however, the cultural relevance may be diminished for use within UK services, which is due consideration, given the role of cultural relevance within work on recovery (Donnelly *et al.,* 2011).

The review by Bonney and Stickley (2008) aimed to review the various ways the respective stakeholders conceptualised the notion of recovery in mental health, reviewing over 170 papers from the previous six years, which were selected upon the basis of being recovery literate and generated within the United Kingdom, unless it was a paper in English which was considered to be seminal. Unlike Leamy *et al.,* (2011) whether this included recovery models was not made explicit, nor was any use (or not) of a quality rating system. The Bonney and Stickley review did however, explicitly explore concepts of recovery from different authors: service users, health care providers and policy makers, offering a view of recovery which explicitly involved three of the main stakeholders. Having said this, it is reported that there was an expectation that the findings would directly relate to the respective agendas of the three groups, e.g. the service users demanding autonomy, the health care providers focusing upon methods and models and the policy makers,' control (LoBiondo-Wood & Haber 2002), which was not reported to be altogether fulfilled.

The quality of the Bonney and Stickley (2008) review was not as high as that presented by Leamy *et al.,* (2011), although it was still a robust piece of work. Both pieces of work had limitations, as have been discussed. However, Leamy's work incorporated other models of recovery, which is of relevance to this study, as developing the codes for the new tool from a conceptual understanding of recovery which is meaningful, was key to the codes accurately relating to the processes involved in a recovery orientation. The review by Bonney and Stickley (2008) offered some areas of strength, where Leamy *et al.,* (2011) study was limited, such as in cultural relevance. That both studies identified very similar key themes also adds weight to the validity of using such themes for the basis of the codes, as it demonstrates consistent results across two independent, reliable sources.

5.3.1.2 Individuals involved in initial development

The individuals involved in the initial development of the codes used to create the tool were not a recruited sample, rather the research team whose aim it was to create the tool. However, it is worth noting that all three individuals involved were psychologists; two were Consultant Clinical Psychologists and one a Trainee Clinical Psychologist. This may have led to a bias about the codes created, due to professional beliefs shared by the research team. There is some evidence that beliefs about mental health conditions may vary between professional groups. For example, using vignettes to describe individuals with either schizophrenia or depression, work by Jorm, *et al.,* (1997) and Jorm, Korten, Jacomb, Christensen, and Henderson, (1999) indicated clinical psychologists were generally found to be more optimistic than psychiatrists and GPs about the prognosis of people with schizophrenia or depression (Jorm *et al.*, 1997; 1999). Work following on from this, reported that mental health nurses were significantly more optimistic than the other professional groups concerning the prognosis for the person described with schizophrenia, after receiving professional help (Caldwell & Jorm, 2001). While not directly related to views on recovery, this does support that views about concepts within mental health may be influenced by profession.

5.3.1.3 Alternative methods

Using a strong empirical base for the codes is important, however there are other methods which could have been employed to generate those codes and other models of recovery or evidence, which could have been used.

The research could have started backwards from the current method, and taken observations and then generated codes to describe the interactions, relating them to the evidence base until all of the themes and interactions which were being observed could be coded with the codes already in existence. Equally, the same methodology could have been employed, but with a wider and more diverse selection of people; focus groups of professionals and service users could have been employed to discuss the themes and generate examples from their own experiences.

As has been discussed, the evidence base used was robust and both pieces of worked included systematic reviews of the literature. As such, the evidence base was able to offer a review of many sources and use these to establish the themes involved in the recovery process**.** These were then utilised within this project. It is recognised, however that other models of recovery exist and other themes have been identified.

5.3.2 The Delphi method

The use and reporting of Delphi studies for health indices needs improving, there are few recommendations and little agreement across sources for researchers who use it (Boulkedid *et al*., 2011). Therefore, while the methodological choices made were based upon research into the Delphi method within indices of quality of health care (Boulkedid *et al*., 2011) and other surveys which have used the Delphi method to explore aspects of recovery in mental health (e.g. Lakeman, 2010; Law & Morrison, 2014) it is recognised that the evidence base for the use of these is inconsistent.

5.3.2.1 Influences in Panels' rating of the codes

Of the 11 codes included in Round 1, 10 of them were rated so as to be included in the measure on both the Relevance and the Importance scale. There was only one code (code 11) which was included, based on the Relevance score alone. During the second round, there was much less consistency between the two scales; only one of the codes was rated to be included on the Relevance scale, however 12 codes were rated so as to be included from the Importance scale. There are several possible explanations for this. It may reflect the difference in the panel used, which was considerably smaller in Round 2 than Round 1 (11 compared to 34 respectively). Secondly, it may reflect that the panel was uncomfortable rating the negative codes highly on the Relevance scale, for fear of a negative reflection on their service. This links to the findings within the first round, where all 11 codes which were included were supportive of recovery, or positive codes. This too may reflect the panels' unease with rating negative codes as being relevant to their service, or indeed important. Rating negative codes highly may have been seen to imply that the individual agreed with such behaviour, rather than the acknowledgement of such interactions' impact, or there could have been a misunderstanding with some of the panel, who might have felt they were being asked about what should occur in services.

5.3.2.2 Sample

The significant characteristics of demographics of the panel used was it was female dominated (81.82%) and the majority were nurses (75.76%), (see Table v for further details of the demographics of the sample). The available data from the NHS across Psychiatry Doctors and Psychiatry staff for May 2106 in England, reported that: 56,662 people were employed in total. 8,626 of these were Medical staff (NB this is all Medical staff), 2,105 were Clinical Psychologists and 13,398 were Staff Nurses, with an additional 3,409 Community Nurses, (NB this does not represent all Nurses position described) (NHS Digital, May 2016). As such, having a sample which included more Nursing professionals than any other kind may not be unreasonable given the makeup of the work force.

The Delphi method involves creating a panel made up of experts in the area of interest (Boulkedid *et al*., 2011); in this case recovery within Recovery and Rehabilitation inpatient mental health services. The members of the panel were deemed to be experts in the area due to having registered with a professional body and working within the aforementioned setting; what is referred to in the Delphi literature as 'expert by experience' (e.g. Lakeman, 2010; Law & Morrsion,2014). This is a recognised practice within the Delphi method (Boulkedid*et al*., 2011; Lakeman, 2010; Law & Morrsion, 2014). However, no other measures of expertise were taken, nor were attitudes or beliefs towards the recovery orientation assessed.

While all members having membership to a professional body guaranteed a level of training, it excluded Health Care Assistants (HCAs). Under national profiles for nursing services available from National Health Service Employers (2008), the formal job statement states that this staff group: 'Undertakes personal care duties for patients in hospital or similar settings and reports patient condition to qualified staff'. There is no formal training required for the role, and therefore the same level of training and knowledge cannot be assumed of

HCAs compared to registered professionals. However, it is recognised that this role produces the most chances for patient interaction, as highlighted by the job description. Excluding this group may have biased the panel towards being top heavy in terms of levels of seniority and decreased the amount of time spent with patients on a daily basis within the panel, compared to one which included HCAs. This could have impacted upon the codes selected.

The second round of the Delphi method had 11 out of the 34 original panel members, who came from two of the three health boards. There were pragmatic reasons for this, as time constraints meant there was a three-week period in which to collect data, and due to the local collaborator in one health board changing roles, it was not possible during that time for the data to be collected in the third health board. This resulted in a smaller sample; demographic information was not collected from all panel members and was therefore not included. There is very limited data on the size of a panel needed. The work by Akins, Tolson and Cole (2005), indicates that a panel of 23 is sufficient to insure reliability. However, the reliability of a smaller panel, such as the one used in Round 2, is unclear. Those involved in the second round of the Delphi method were arguably more involved in the process, as they were more willing to complete the second round, despite the tighter time scale. It is possible that the 11 involved in Round 2 were also more comfortable with the process, as they rated every code on both scales, unlike some members of the panel in Round 1.

5.3.2.3 Clarity of instructions

In endeavouring to ensure the clarity of the task asked of the panel, the instructions and the rating scales were given to a Clinical Psychologist and an Occupational Therapist who worked outside of the Trusts involved; the feedback was that the instructions were clear and the scales easy to use. However, due to the remote nature of the Delphi method, it was not possible to ensure the panel were clear on what they were being asked to rate. On some of the panels' feedback, codes were left unrated. This may have indicated uncertainty on how to rate the code to reflect their view. Indeed, on two occasions it was indicated that the panel member felt that a particular code was negative (as it was designed to be) and as such, they did not want to rate it as being important, although feeling it had a significant impact, as it was important not to act in such a way. This was despite the instructions asking them to. This uncertainly about rating negative codes highly is also a possible explanation for only positive codes being included after the first round, as discussed. It might have been beneficial for the instructions to be reviewed by a wider selection of professionals to ensure their clarity. Contact details were provided, but no one from the panel made contact to ask for further instruction. This may have been because the onus was on the panel member to make contact, rather than the researcher to ensure their understanding.

5.3.2.4 Alternative methods

An alternative methodology to create a consensus would have been to use focus group(s). However, this methodology would have been likely to produce a much smaller sample of professionals. A possible compromise would be using a modified Delphi method, which is identified by the literature (Boulkedid *et al.,* 2011). In the studies reviewed in the aforementioned study, the modified method involved having a focus group of some or all of the panel between rounds, in which the findings are discussed and the focus group is used to make alterations to items to be rated by the panel in the next round. This may have opened up conversations about further missing codes, or difficulties found in rating the codes.

5.3.3 Selection of wards for piloting of the new tool

5.3.3.1 Choice of measure

The Recovery Self-Assessment, Provider Version (RSA-PV) was used to differentiate between the wards. This is a standardised measure of recovery in the form of a self-report questionnaire. It is reported to have undergone an appropriate process of item development (drawing on stakeholders' input) and testing (factor analysis), and has been published in peer-reviewed journal (O'Connell, et al., 2005). The face validity is supported as items were derived from extensive literature reviews and discussions with: persons in recovery, mental health service providers, family members and administrators. However, quantitative indicators of validity have yet to be established (Campbell-Orde *et al.,* 2005).

Within this project the RSA-PV total scores were used, which provides information on staff members' perspectives on the recovery orientation of the ward, by asking them to rate service level factors which are believed to influence how recovery orientated the service is. The RSA-PV can also be used to calculate scores for the five subscales it is comprised of: life goals, involvement, diversity of treatment options, choice and individually-tailored services. However, these were not utilised within this project, which may have decreased the information which could have been extracted from the RSA-PV.

Donnelly *et al*., (2011) report that recovery measures should be culturally sensitive to be relevant. The RSA-PV was designed for use within the United States, while the project in question was carried out in the United Kingdom, therefore, it may not have been a culturally appropriate measure.

The newly developed measure aimed to assess whether the interactions of staff embody recovery themes, as identified in Leamy *et al.,* (2011) and Bonney and Stickley (2008) review of the recovery literature. Despite both measures aiming to assess the same concept, the five domains which comprise the RSA-PV differ to some extent from the seven overarching themes which were used to create the tool in this project. This reflects the fact that the recovery concept has been operationalised differently by the two tools. No standardised measures currently available assess all of the concepts covered within Leamy *et al.,* (2011) model of recovery (Shanks *et al.,* 2013), which was part of the rationale for the development of the tool. Therefore, it could be argued that the two measures used within this project are not designed to measure the same domains. Having said this, they do both aim to measure the same concept and therefore it was deemed reasonable to use the standardised measure with which to assess the new tool's validity.

5.3.3.2 Sample

Demographic information was collected on participants who completed the RSA-PV on the seven wards involved. Comparing the two wards in which the observations were carried out on the sample there was some notable difference; ward 12 had a higher proportion of females than the sample as a whole, 80% vs 63%, whereas ward 31 had an even split on gender, therefore having fewer females than the sample as a whole. Ward 31 had a lower percentage of nurses making up their sample (16%) than across the sample, where 39% were nurses, compared to ward 12, which had a higher percentage of nurses (60%) than the whole sample. A considerably higher percentage of the sample from ward 31 were HCAs (66%) than compared to the sample as a whole, (31%) or from ward 12, who had 20% HCAs, which was less than the sample as a whole. As has been commented upon, there is some evidence that views relating to mental health may differ between professions (Caldwell & Jorm, 2001; Jorm *et al.*, 1997; 1999), which may mean the differences in the professional make-up from each ward could have impacted upon the scores generated.

When looking at the samples from each of the seven wards, the greatest difference was the number of participants. This ranged from nine on wards 12, 21 and 33 and six participants on ward 31. This is of note, as ward 31 was involved in the observations as the ward with the highest RSA-PV score; however it was the opinion of a smaller sample which may have impacted upon this. The Coefficient of Variation for each ward was less than 1 and therefore classed as low variance, meaning the opinions of the sample on each ward did not vary greatly. There is no guidance on how many of a staff team is ideal to complete this measure for an accurate representation of the wards recovery orientation. Furthermore, accurate data from the wards on how many permanent and bank members of staff were on their teams was not collected, therefore it is not possible to determine the proportionate sample size for each ward. The face validity might have been improved if a percentage of the permanent staff team had been agreed upon, and only wards in which respondents from that number of staff members had been achieved were then included in the study.

5.3.3.3 Difference between wards

Using the standardised measure to differentiate between wards offered a number of pragmatic benefits to this study; it allowed seven wards to be involved and ensured the participant burden was kept to a minimum. However, this choice of methodology may have influenced the outcome of the validity testing, as the difference between the two wards in the RSA-PV total mean score was not that great. The difference between the two wards used was only 14.45, where the range of scores of the measure was 36 – 180. Therefore, it is arguable that the lack of difference found by the new tool was due to a lack of difference between the two wards. This poses a considerable difficulty to the study, as the validity of the tool can be said conclusively to be poor, as it may be there was no significant difference to find.

5.3.3.4 Alternative methods

An alternative method in which the information from the RSA-PV could be utilised could be to select particular subscales from the RSA-PV which are considered to most closely match with the conceptual basis used to create the new tool, and select wards which differ most of all on these subscales, rather than using the RSA total.

An additional alternative which would have utilised the information from the RSA-PV to a greater extent, would be to use the scores from each of the subscales on each of the wards, as dependent variables. Observations could then be carried out on each of the wards involved. Which codes are used most frequently could be assessed, and differences between which codes were used on which wards could be related back to which subscales the wards score highest and lowest on, on the RSA-PV.

The study could have been improved by having two wards which differed more significantly from one another on their recovery orientation. One method of achieving this would be to use the same methodology, but sampling a larger number of wards. However, this would present pragmatic difficulties (such as being able to sample the wards in observations afterwards if they are geographically very diverse, having local collaborators and managing an increased number of Research and Development departments etc.), so this would have been beyond the scope of this project, and would still not have guaranteed a larger difference.

An alternative would have been to use a repeat measures design. Within this the same ward would have been sampled on two occasions, before and after an intervention which would have been expected to change the recovery orientation. The tool would be utilised to establish a base line and then an intervention, which was likely to change the recovery orientation would be carried out, and the tool used again to assess if it could establish a difference. A standardised measure could be used alongside the new tool to triangulate the results between the expected change (due to the intervention), any difference (or not) reported by the new tool and any difference found by the standardised measure. Staff training is a possible example of such an intervention. Specific staff training has been found to impact upon factors influencing quality of care in inpatient mental health services. Nesset *et al.,* (2009) demonstrated that staff training on ward atmosphere could lead to a statistically significant improvement in the reported ward atmosphere. One option for such might be using the REFOCUS intervention, which in addition is based upon the themes that are included in the new tool.

5.3.4 Observations: Reliability and Validity testing of the new tool

5.3.4.1 Characteristics of the codes and the data generated by the tool

All but one of the codes which were included in the tool were utilised during the observations across the two days on each ward. The code not used was code 30, ‘Asking how a client would like to deal with a problem; exploring options.’ This indicated that the Delphi method did produce a tool in which the codes were relevant to the services it was designed for. All of the other codes within the measure were used at least once and there was a dominance of use of code 8: ‘Staff initiated non-clinical interactions.’ Code 9, ‘Positive response to client initiated interactions.’ and code 12, ‘Clients not being engaged with.’ No interactions occurred that the researcher was unable to code using the tool; indicating again the relevance of the codes included, and that the tool included sufficient codes to cover the interactions seen.

5.3.4.2 Characteristics of the data collected

The design of the measure meant any interaction, however small, which occurred in the designated area during the five-minute time slot was recorded and then coded. The current tool was not based upon DCM©. However, using the DCM© as a comparison for other observational tools used within health care setting, it is notable that DCM© only records 'significant" interactions'. As such, the result may have been produced because insignificant interactions, such as an exchange of greeting, were recorded and counted. On the other hand, such interactions may also illustrate the ethos of the ward.

The participant information and information given by the researcher on the two wards on which the observations took place, made it explicit that the aim of the research was to examine the validity and reliability of the new tool, not to assess the quality of the ward. However, the act of having an observer on the ward, who all members of staff knew was involved in measuring recovery principals, may have produced demand characteristics in the staff team, which in turn may have influenced the findings. Demand characteristics are a well-known and widely used concept within psychology (Dallos, 2012). Originally proposed in the work of Orne 50 years ago, they refer to participants being aware of what the researcher is trying to investigate, or anticipating findings, and what this implies for how participants are expected to behave (Orne, 1959; Orne 1962). However, caution needs to be applied when considering their impact in this study. It is not possible to quantify such an effect, so its impact on the results is speculation. Secondly, a recent systematic review into the effects of demand characteristics on participant behaviours in non-laboratory settings indicated that there were few dedicated studies to such and further research was necessary (McCambridge, de Bruin & Witton, 2012).

The design of the tool produced categorical data; an interaction was either positive or negative. Using these dichotomous categories may have missed important information or forced interactions to be coded as one or the other, which may not have been wholly accurate. All codes which were included in the measure were viewed as equally relevant and important. It may be that particularly types of interactions have a greater impact than others. It could have been that although there was little difference in the number of positive and negative interactions seen, the quality of those interactions was different and the current measure is unable to capture this.

5.3.4.3 Diversity of people and locations sampled

Due to the wish to keep participants anonymous, demographic information on those involved in the observations was not taken, therefore it is not possible to comment on the demographic information of those involved in the interactions seen. The number of different members of staff involved in interactions ranged from 6-8 with a mean of 7.3 across the six sessions on ward 21, and 6-11 with a mean of 8.3 on ward 31. With regard to the patients involved, this ranges from 5-10 with a mean of 8.3 per session on ward 21, and 6-10 with a mean of 8.6 different participants involved per session. These numbers are reasonable given the size of the wards in terms of patients and how many staff are on shift at a given time. However, what is not captured by the data is that certain patients and members of staff did interact more often with one another. This highlights what the design of this tool misses, which is patients who are not engaging with services.

The diversity of the staff and patients involved in the observations was reasonable, however there were certain staff and patients who were observed more frequently, due to who was on the ward, this biases the interactions to staff and patients who spend time in communal areas opposed to those in offices, their own private space (such as bedrooms), or who were off the ward. The tool has no way of accounting for patients who are not interacting due to being in their bedroom, or avoiding social contact. Similarly, the level of staff members who avoid interactions cannot be accounted for within this tool. There is not a way of quantifying what this missed information may have showed, however it is a notable shortcoming in the measure's current design, and worthy of note when considering possible reasons for a lack of difference between wards. This tool focused only upon communal areas within the ward. This was done, because following staff and patients off the ward was considered to be invasive by the research team, and similarly observing personal consultations, such as ward rounds would have been inappropriate. However, missing these types of activities is likely to miss key interaction, which shows how recovery orientated a service is. For example, connectedness is a theme identified by both Leamy *et al.,* (2011) and Bonney and Stickley's (2008) reviews of the recovery literature. Therefore, how patients and staff interact in public, and how patients are encouraged to engage social (or not), is of relevance. Equally, events such as ward-round run the risk of being dominated by professionals, where patients are invited in at others' convenience, may highlight areas where services would benefit from thinking about the power balance; but the observational nature of the measure developed may be unsuitable to assess it.

5.3.4.4 Timing of observations

The feedback from consultations with the wards was that the least distress would be caused if an explanation of the study was given and the observations then carried out quickly afterwards. Therefore, the observations were completed over two days. The face validity was reasonable, as the observations were carried out evenly over the course of the day. However, a two-day period is a relatively short snapshot, which may have misrepresented the wards.

All the observations were carried out during the conventional working week, despite the two services running constantly. Therefore, the data gathered can only be said to be representative of the wards within those time constraints.

5.3.4.5 Alternative methods

A way of countering the possible observation bias could be to use video cameras or CCTV from wards, and then randomly select clips to review and rate. This would avoid the presence of a researcher, which could alter the behaviours seen. However, there are pragmatic difficulties with this method, and more importantly it would be ethically challenging, especially if some individuals do not want to be involved.

To increase the face validity and take longer snapshots of the wards, the observations could have been done over a longer period, perhaps selecting three days of observation over a three-week period. Taking the samples from a more diverse time period may have produced a more accurate representations of the wards. It may have also captured a more diverse range members of staff. However, this would be inappropriate if it could not be managed in a manner which would avoid patients' distress. Alternatively, the two wards could have been sampled on the same day; for example, sampling ward 12 in the morning on the first day and ward 31 in the afternoon and then swapping over for the second day.
Methods of addressing the possible problem, with all codes being as important as one another, could be to alter the tool so the codes are weighted. Using the current evidence base, greater weight could be given to codes which relate to connectedness, hope and optimism about the future, identity, meaning in life and empowerment, as the themes stated to have the most proximal relevance to clinical practice (Lemey *et al.,* 2011). Secondly, the Importance ratings provided by the panel within the Delphi method could be consulted. Looking at the frequency with which members of the panel rated codes highly could be used to establish a hierarchy of importance.

5.4 Strengths and limitations of the current study

There were several strengths and limitations identified in the current study. These were generally in relation to the empirical base from which the study was derived, the power within the statistical procedure offering strengths, whereas the wards selected for observation and the samples used may have presented some limitations. These are outlined below.

5.4.1 Strengths

5.4.1.1 Empirical base

One of the strengths of this study was the empirical base from which the initial codes used to create the tool were developed from. There are several strengths with the choice of literature used as the empirical base. The recovery processes identified in Leamy *et* al., (2011) was a very robust piece of work; the literature review included other models of recovery and the development of the conceptual framework included two checks, both evidence based and from a panel of experts. The model postulated has also undergone further validity testing (Bird *et al.,* 2014; Slade *et al.,* 2012). The consistency between the themes identified by Leamy *et* al., (2011) and Bonney and Stickley (2008) adds face validity. That Bonney and Stickley (2008) focused on evidence from the United Kingdom, ensures that the tool developed is culturally relevant.

By basing the tool on broader themes, as opposed to narrower concepts it was hoped to manage the difficulties of implementing and measuring recovery posed by the lack of a universal definition of recovery, and the issue of it being a contested concept (Bonney & Stickley, 2008). Incorporating a strong thematic base into the tool also avoided the criticisms which have been levied at current measures of recovery that have not (Shanks *et al.,* 2013).

However, it is recognised that the model postulated by Leamy *et al.,* (2011) is in its relative infancy, and that while these two systematic reviews are consistent, other literature into the recovery orientation may be more diverse. While the choices made were sound, other evidence could have been employed.

5.4.1.2 Power

The observational aspect of the study was adequately powered (90%). This was based upon the number of interactions seen on each ward. This enables the research to be confident that any difference would have been detected. However, the positive impact of this is somewhat limited, given that whether there was enough of a difference between wards was not established clearly enough, (see section 5.4.2.2).

5.4.2 Limitations

5.4.2.1 Sample

Throughout the study, the samples used consisted of members of staff; patients were not included at any stage. There were several reasons for this. This project was the very first stages of the development of the measure, and it was always presumed that further work will be needed to be completed before the tool can be used effectively clinically. As such, patients' involvement was visualised for later stages of the development. The literature which was used as a base from which the initial codes were developed did include patients' first-hand experience, which was felt to anchor the codes in an evidence base which reflected patients' thoughts and experiences. Pragmatically, the Ethics committee was more supportive of a project which focused upon members of staff. Having said this, patients or service users not being part of the sample could have impacted upon the measure developed. Evidence from the systematic review presented in Chapter 2 indicates that staff members and patients perceive ward atmosphere differently from one another (Brunt & Rask, 2005; Jansson & Eklund, 2002a; Røssberg & Friis 2004; Schjødt *et al.*, 2003). This could indicate that one might expect to find differences between patient and staff members' perceptions on other experiences within an inpatient setting. Differences in which codes were of importance and relevance to people's first-hand experience may have led to different codes being developed. Furthermore, the design of the measure only considered staff behaviour, the patients' behaviour is used as context with which to code the staff within the interaction.

5.4.2.2 Wards selected

The greatest difficulty in assessing the results from this project related to the validly of the tool. The result was not statistically significant; however, this may have been because the recovery orientation of the two wards did not differ enough for a significant difference to be found. As such, it is not possible to comment with any certainty on this aspect of the study. The methodology chosen created a not insignificant amount of data which was unused (the data from the remaining five wards, such as the RSA-PV and demographic information). However, in principle the design was reasonable in that it differentiated the wards using a standardised measure. The limitation occurred, because the results on the RSA-PV from all wards were relatively similar. Alternative methods piloting the tool (see section 5.3.3.4), such as a repeat measure design have been discussed. When considering the design that was used, a safeguard which could have been employed, would be for a pre-determined size in the difference between wards being determined prior to the calculation, which if not reached, could entail a change in the methodology. In such a case, a possible option could be for all the wards to be sampled and a correlation performed, to assess if the number of positive interactions increased as the mean of the RSA-PV increased.

5.5 Clinical Implications and Future Research

In its current format, the measure developed does not offer anything to clinical services, as its validity cannot be established. However, as has been discussed, the current self-report measures available to assess the recovery orientation of services include a number of pitfalls, which using an observational measure could counter. The proposed measure attempts to consider how individual recovery is supported at a service level via assessment of how staff members who work most closely with patients interact with them. If the measure could be developed to a point where it is valid, the information yielded would be able to support services, to highlight periods where enacting a recovery orientation is more difficult, in which service level changes may be able to support them. It also has the potential to be used as a training aid, highlighting to staff the times when a more recovery orientated manner would be possible.

However, the measure is looking at the recovery orientation of a service by only looking at one level of how members of staff interact with patients, whereas the reality of current inpatient services is that there are aspects of the system which are at odds with a recovery orientation to such an extent, a true recovery orientated service may not be possible to achieve by working from the bottom up only. Slade *et al.,* (2014) offers seven misuses or ‘abuses’ of the recovery orientation, which are: Recovery is the latest model: Recovery does not apply to ‘my’ patients: Service can make people recover through effective treatment: Compulsory detention and treatment aid recovery: Recovery is about making people independent and normal: Contributing to society happens only after the person is recovered. These ‘abuses’ highlight changes which need to be made at every level of services.

While meaningful change may need to permeate the system at many more levels than directly between patients and members of staff, working from the bottom up may enable wards to highlight points at which they are less able to act in a recovery supporting manner due to local policy and procedure that may not support such. Demonstrating aspects of the system, which are inconsistent with recovery could help change practices.

When thinking about future research, the methodology offers a potentially meaningful way of assessing how recovery supporting principles are enacted in inpatient care settings, and while the tool appears reliable, the codes are not yet of use. It may be that while the rationale for the use of such a measure may still be valid, the development of the codes may need to be re-visited.

At the broadest level, it may have been that the wards did not present different enough recovery orientations for the new measure to be able to tell the difference between. The same measure could be tested using different methodology. Establishing the validity (or not), of the current tool more conclusively would make it clear what the next logical step in development should be.

5.6 Summary and conclusion

Measuring recovery poses multiple problems, not least because what the concept entails is diverse. However, reviewing the literature indicates that a number of common themes have been established (Bonney & Stickley, 2008; Leamy *et al.,* 2011), which if utilised could enable services to better support personal recovery. Current measures utilise self-report measures, which is a further limitation. There is evidence which suggests the relationship between staff members and patients has an impact on various facets of inpatient mental health care, such as the domain of ward atmosphere, which is predictive of satisfaction (Bressington, *et al.,* 2011; Middelbøe, *et al.,* 2001) and has been identified as a mode via which inpatient mental health services operate (Bowers *et al.,* 2009). As such, developing a measure which used an alternative method of data collection to self-report, was firmly based upon consistent themes of recovery, and focused upon the relationship between staff members and patients, offereing a new method of assessing the recovery orientation of a ward.

Based upon this, a four-stage project was designed, which aimed to develop and carry out the initial testing on an observational tool, designed to assess the recovery orientation of a service (in terms of how it supported personal recovery) via assessing the manner in which staff members interacted with patients. This involved; using the available evidence base and clinical expertise to develop descriptions of types of interactions which embodied the themes of recovery. These were termed codes. A Delphi method was then employed, via which a consensus from a panel of experts by experience was gathered as to which of the codes should be included in the final tool, based on each code's perceived relevance and importance to inpatient Rehabilitation and Recovery services. Wards were then recruited to the study, within which staff members were asked to complete a standardised measure of the recovery orientation of the service (RSA-PV). The wards with the highest and the lowest mean RSA-PV score were selected for the tool to be piloted in. Finally, observations of interactions were carried out in the two wards, which were coded using the tool. Twenty percent of the interactions were coded by two researchers to assess the reliability, and the validity was tested using a Chi-square to compare the number of positive and negative codes between wards.

This produced an observational tool with good reliability, but seemingly poor validity. The codes based upon the recovery literature and selected via the Delphi method, were found to be relevant to the services they were used in, as demonstrated by all but one of the codes being utilised during the observations. The greatest difficulty with the work presented is that of validity. A significant difference was not found however, due to similarities within the wards which were sampled; this could have been due to a lack of difference between wards, rather than the tool's lack of validity.

More work needs to be carried out on the validity of the tool, before its clinical implications can be explored. This work could be further improved by involving other stakeholders, not least patients from inpatient services and service users, which was not within the scope of this piece of work; whilst beyond this project, the lack of a role for patients and service users is of note.

The assertion that a recovery orientation should be adopted by services, has been voiced by a range of stakeholders including policy, professionals and service users (e.g. DoH, UK 2012; Tew *et al.,* 2012; Slade, 2009). The author's own perspective on the recovery orientation is that for personal recovery, as well as clinical recovery to be enacted within services, a fundamental change within the power dynamics needs to be seen, whereby the tone and direction of care and research is not dictated by services or professionals with input from patients and service users, but rather service users and professionals utilise one another's skill set to reach a shared goal. With this in mind, a shortcoming of this project is that the methodologies used to create the measure, while scientifically valid, do not reflect the core values of a recovery orientation: the project was directed by professionals. The choices made within this project have reflected the evidence available in terms of both theory and methodology, and worked within the constraints of the scope of this project. However, it could be argued that by prioritising the scientific rigour, changes in how mental health research is carried out which might be more recovery orientated, were not supported. It was recognised from the outset of this project that within the scope of the work it would not be possible to produce a tool, which was fit for purpose and that further work would be needed. This opportunity is being taken to express the author's position: that the next stages of development of this tool must include the voices of patients and service users if the tool in development is going to be able to reflect personal recovery at a services level, as is the intention.

Despite the non-significant result, there is still some merit in the design of the proposed tool. As has been discussed, whether the codes which have been produced so far are of any value is difficult to positively state, given the relatively small difference between the two wards on the standardised measure of recovery in which the tool was piloted. The difficulties for inpatients in managing their role in a system which forcibly hospitalises individuals, while imposing a recovery orientation is an issue. Current rhetoric indicates that changes need to occur at a number of levels for a recovery orientation to be full adopted, and providing services with a valid measure of recovery, such as the tool in development, would support services in this endeavour.

Reference

Akins R.B.Tolson, H. & Cole, B.R. (2005). Stability of response characteristics of a Delphi panel: application of bootstrap data expansion. *BMC Medical Research Methodology*. 1: 5:37.

Andresen R, Oades L, Caputi P. Psychological recovery: beyond mental illness. London: Wiley, 2011.

Anthony WA. (1993) Recovery from mental illness: the guiding vision of the mental health system in the 1990s. Psychosocial Rehabilitation Journal. 16:11-23.

Austin, E. J., Gibson, G. J., Deary, I. J., McGregor, M. J., & Dent, J. B. (1998). Individual response spread in self-report scales: personality correlations and consequences. Personality and Individual Differences. 24: 421–438.

Barker P., & Buchanan-Barker, P. (2005). The Tidal Model. Hove. *Brunner-Routledge.*

Beazley P. & Gudjonsson, G. (2011). Motivating Inpatients to Engage with Treatment: The Role of Depression & Ward Atmosphere. *Nordic Journal of Psychiatry*. 65; 95-100.

Beech B F (1991). Changes the Delphi technique adapted for classroom evaluation of clinical placements. *Nurse Education Today.* 11: 207-212

Bellack A. S. (2006). Scientific and consumer models of recovery in schizophrenia: Concordance, contrasts, and implications. Schizophrenia Bulletin, 32, 432−442.

Bellack A. S. & Drapalski A. (2012). Issues and developments on the consumer recovery construct. *World Psychiatry*. 11:156-60.

Bhandari A. & Wagner T, (2006). Self-reported utilization of health care services: improving measurement and accuracy. Med Care Res Rev. 63(2):217-35.

Bird V. Leamy M, Tew J, Le Boutillier C, Williams J, Slade M (2014) Fit for purpose? Validation of a conceptual framework for personal recovery with current mental health consumers, Australian and New Zealand Journal of Psychiatry, 48, 644-653.

Bonney S. & Stickley, T. (2008). Recovery and mental health: a review of the British literature. *Journal of Psychiatric Mental Health Nursing*. Mar;15(2):140-53. doi: 10.1111/j.1365-2850.2007.01185.x.

Bowers (2005). Reasons for admission and their implications for the nature of acute inpatient psychiatric nursing. *Journal of Psychiatric and mental health nursing*. 12: 231-236.

Bowers,L., Simpson,A., Alexander,J.; Hackney,D., Nijman,H.; Grange,A., Warren,J. (2005). The nature and purpose of acute psychiatric wards: the Tompkins acute wards study. *Journal of Mental Health*.14(6): 625-635.

Bowers,L., Chaplin,R., Quirk,A., Lelliott,P. (2009). A conceptual model of the aims and functions of acute inpatient psychiatry. *Journal of Mental Health.* 18(4): 316-325.

Boulkedid R., Abdoul H., Loustau M., Sibon,O. & Alberti C. (2010). Using and Reporting the Delphi Method for Selecting Healthcare Quality Indicators: A Systematic Review. *PLoS ONE* 6(6): e20476. doi:10.1371/journal.pone.0020476

Bradford Dementia Group. (2005). Dementia Care Mapping. Bradford: Bradford Dementia Group, University of Bradford.

Braham L.G., Heasley, J.F. & Akiens S. (2013). An evaluation of night confinement in a high secure hospital. *Mental Health Review Journal.* 18(1): 21 – 31.

Bressington D., Stewart B., Beer D. & MacInnes D. (2011). Levels of Service User Satisfaction in Secure Settings- A survey of the Association Between Perceived Social Climate, Perceived Therapeutic Relationship & Satisfaction with Forensic Services. *International Journal of Nursing Studies*. 48: 1349-1356.

British Psychological Society. Code of Ethics and Conduct. BPS, 2006 (http://www.bps.org.uk/downloadfile.cfm?file_uuid=5084A882-1143-DFD0-7E6C-F1938A65C242&ext=pdf).

Brooker D., & Surr, C. (2005). Dementia Care Mapping: Principles and Practice. Bradford: Bradford Dementia Group, University of Bradford.

Brunt D. (2008) The Ward Atmosphere of Single-Sex Wards in a Maximum-Security Forensic Psychiatric Hospital in Sweden. *Mental health Nursing*. 29:221-241.

Brunt D. & Rask M. (2005). Patient and staff perceptions of the ward atmosphere in a Swedish maximum-security forensic psychiatric hospital. *The Journal of Forensic Psychiatry & Psychology.* 16(2): 263-276.

Burgess P, Pirkis J, Coombs T, Rosen A. (2011). Assessing the value of existing recovery measures for routine use in Australian mental health services. *Australian and New Zealand Journal of Psychiatry*. 45: 267–80.

Caldwell K., Henshaw L. & Taylor G. (2005) Developing a framework for critiquing health research. *Journal of Health, Social and Environmental*. 6 (1): 45-54.

Caldwell T.M. & Jorm A.F. (2001). Mental health nurses' beliefs about likely outcomes for people with schizophrenia or depression: a comparison with the public and other healthcare professionals. *Australian and New Zealand Journal of Psychiatry*. 10(1):42-54.

Campbell R., Allan S. & Sims P. (2014). Service Attachment: The Relative Contributions of Ward Climate Perceptions & Attachment Anxiety & Avoidance in Male Inpatients with Psychosis. *Criminal Behaviour and Mental Health*. 24; 49-59.

Campbell-Orde T, Chamberlin J Carpenter J & Leff SH. (2005). Measuring the Promise: A Compendium of Recovery Measures. Vol II. *Human Services Research Institute*.

Care Services Improvement Partnership, Royal College of Psychiatrists, Social Care Institute for Excellence. *A Common Purpose: Recovery in Future Mental Health Services.* CSIP, 2007.

College of Occupational Therapists. Code of Ethics and Professional Conduct (Revised). *College of Occupational Therapists*, 2005.

Coolican H. (2009). *Research Methods and Statistics in Psychology*. 5th Edition. Routledge.

Davidge M. Elias, S. & Hayes, B. (1993) *Survey of English mental illness hospitals.* Birmingham, Inter-Authority Comparisons and Consultancy, University of Birmingham.

Davidson L., Lawless M. S. & Leary F. (2006) Concepts of recovery: competing or complementary. *Current Opinion in Psychiatry,* 19 (6); 619–624.

Davidson L Schmutte T., Dinzeo T. & Andres-Hyman R. (2008). Remission and recovery in schizophrenia: practitioner and patient perspectives. *Schizophrenia Bulletin*. 34: 5–8

Davidson L, Tondora J, Lawless MS, O'Connell M, Rowe M. (2009). *A Practical Guide to Recovery-Oriented Practice Tools for Transforming Mental Health Care*. Oxford University Press.

Department of Health. (1999) *National Service Framework for Mental Health.* London. Department of Health.

Department of Health (2005) *Mental Health Policy Implementation Guide: Adult Acute Inpatient Care Provision.* London. Department of Health

Department of Health (June, 2008). *High Quality Care for All: NHS next stage review final report.* London: Department of Health. Retrieved from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/228836/7432.pdf

Department of Health (2009). *Performance data and statistics: beds.* (web publication)http://www.dh.gov.uk/en/Publicationsandstatistics/Statistics/Performancedataandstatistics/Beds/index.htm

Department of Health. (2011). *No health without mental health. Delivering better mental health outcomes for people of all ages*. London: Department of Health.

Department of Health Social Services and Public Safety (Northern Ireland) (2010). *Service framework for mental health and wellbeing.* Belfast: Department of Health Social Services and Public Safety.

Department of Health and Ageing. (2009). *Fourth national mental health plan: an agenda for collaborative government action in mental health 2009–2014*. Canberra: Commonwealth of Australia.

Dickens G.L., Suesse M., Snyman P. & Picchioni M. (2014). Associations Between Ward Climate & Patient Characteristics in a Secure Forensic Mental Health Service. *The Journal of Forensic Psychiatry & Psychology.* 25(2); 195-211.

Donnelly M., Scott D., McGilloway, S. O'Neill, T. Williams, J. & Slade M. (2011). Patient outcomes: what are the best methods for measuring recovery from mental illness and capturing feedback from patients in order to inform service improvement? A report commissioned by the Bamford Implementation Rapid Review Scheme Reference: COM/4409/10. Retrieved from: http://www.publichealth.hscni.net/sites/default/files/Patient%20Outcomes.pdf

Effective Public Health Practice Project. (2009). *Quality Assessment Tool for Quantitative Studies*. Canada. EPHPP(http://www.ephpp.ca/tools.html).

Egger M., Zellweger-Zähner, T Schneider M., Junker C., Lengeler C. & Antes G. (1997). Language bias in randomised controlled trials published in English and German. *The Lancet.* 350: 326-329.

FakhouryW. & PriebeS. (2000) The process of deinstitutionalisation: an international overview. *Current Opinion in Psychiatry*. 15: 187-192

Fakhour,W. & Priebe,S. (2007) Deinstitutionalisation and reinstitutionalisation: major changes in the provision of mental health care. *Psychiatry.* 6(8): 313- 316

Farkas M, Gagne C, Anthony W. & Chamberlin, J. (2005). Implementing recovery oriented evidence based programs: identifying the critical dimensions. *Community Mental Health Journal.* 41:141-58.

Fossey J., Lee, L. and Ballard, C. (2002). Dementia Care Mapping as a research tool for measuring quality of life in care settings: Psychometric properties. *International Journal of Geriatric Psychiatry,* 17: 1064–1070.

Gagné C. & Gaston, G. (2005). Improving self-report measures of non-adherence to HIV medications. *Psychology and Health.* 20(6):803–816.

Gilburt H., Peck, E., Ashton R., Edwards N. & Naylor C. (2014). Service transformation: Lessons from mental health. The King's Fund.

Gjerden. P. & Moen, H. (2001). Patient Satisfaction & Ward Atmosphere During a Crisis in Open Psychiatric Ward. *Social Psychiatry and Psychiatric Epidemiology.*

Godlee F, Jefferson T, Clark J. *How to peer review a qualitative manuscript. In Peer Review in Health Sciences* (2nd edn). (eds Godlee F, Jefferson T): 219–35. BMJ Books, 2003.

Hawkshead J, Krousel-Wood MA. (2007) Techniques for measuring medication adherence in hypertensive patients in outpatient settings: advantages and limitations. *Disease Management & Health Out.* 15(2):109–118.

Helmer O., & Rescher N. (1959). On the epistemology of the inexact science. *Management Science*. 6: 25-53.

Hughes R., Hayward M., & Finlay, W. M. L. (2009). Patients' perceptions of the impact of involuntary inpatient care on self, relationships and recovery. *Journal of Mental Health.* 18(2), 152-160

Jairath N. & Weinstein J. (1994) The Delphi methodology: a useful administrative approach. *Canadian Journal of Nursing Administration*. 7, 29–42.

Jansson J.A. & Eklund, M. (2002a). Stability of Perceived Ward Atmosphere Over Time, Diagnosis & Gender for Patients with Psychosis. *Nordic Journal of Psychiatry*. 56; 407-412.

Jansson J.A. & Eklund, M. (2002b). How the Inner World is Reflected in Relation to Perceived Ward Atmosphere Among Patients with Psychosis. *Social Psychiatry and Psychiatric Epidemiology*.

Johansson H. & Eklund, M. (2004). Helping Alliance & Ward Atmosphere in Psychiatric In-patient Care. *Psychology and Psychotherapy: Theory, Research and Practice*. 77; 511-523.

The Joint Commissioning Panel for Mental Health, June 2013. Guidance for commissioners of acute care – inpatient and crisis home treatment. London. Commissioning Panel for Mental Health.

The Joint Commissioning Panel for Mental Health, November, 2012. Guidance for commissioners of rehabilitation services for people with complex mental health needs. London. Commissioning Panel for Mental Health.

Jörgensen K.N., Römma, V. & Rundmo, T. (2009). Associations Between Ward Atmosphere, Patient Satisfaction & Outcome. *Journal of Psychiatric and Mental Health Nursing*. 16; 113-120.

Jorm A. F., Korten, A. E., Jacomb, P. A., Rodgers, B., Pollitt, P., Christensen, H. & Henderson, S. (1997). Helpfulness of interventions for mental disorders: Beliefs of health professionals compared with the general public. *British Journal of Psychiatry,* 171, 233–237.

Jorm A. F., Korten, A. E., Jacomb, P. A., Christensen, H. & Henderson, S. (1999). Attitudes towards people with a mental disorder: A survey of the Australian public and health professionals. *Australian and New Zealand Journal of Psychiatry*, 33: 77–83.

Kerfoot G., Bamford, Z. & Jones, S.A. (2012) Evaluation of Psychological Provision into an Acute Inpatient Unit. *Mental health Review Journal.* 17(1): 26-38.

Kitwood T. (1997), *Evaluating Dementia Care: The DCM method*. 7th ed., University of Bradford.

Kuosmanen L., Hatonen H., Jykinen A.R., Katajisto J. & Valimaki M. (2006). Patient Satisfaction with Psychiatric Inpatient Care. *Journal of Advanced Nursing*. 55(6): 655-663.

Lakeman R. (2010). Mental health recovery competencies for mental health workers: A Delphi study. *Journal of Mental Health.* 19(1): 55–67.

Law H. & Morrison, A.P. (2014). Recovery in psychosis: a Delphi study with experts by experience. *Schizophrenia Bulletin*. 40(6):1347-55.

Leamy M, Bird V, Le Boutillier C, Williams J, Slade M (2011) A conceptual framework for personal recovery in mental health: systematic review and narrative synthesis, *British Journal of Psychiatry.* 199: 445-452.

Leigh A.J., O'Hanlon K., Sheldrick R., Surr. C. & Hare. D.J. (2014). Care mapping in clinical neuroscience settings: Cognitive impairment and dependency. Neuropsychological Rehabilitation: *An International Journal.* DOI: 10.1080/09602011.2014.951366

Livingston J.D., Nijdam-Jones A. & Brink J. (2012). A Tale of Two Cultures: Examining Patient-Centred Care in Forensic Mental Health Hospital. *The Journal of Forensic Psychiatry and Psychology.* 23(3): 345-360.

Livingston J.D., Nijdam-Jones A., Lapsley S., Calderwood C. & Brink J. (2013). Supporting Recovery by Improving Patient Engagement in a Forensic Mental Health Hospital: Results From a Demonstration Project. Journal of the American Psychiatric Nurses' Association. 19(3) 132-145

Long C.G., AnagnostakisK., Fox E., Silaule P., Somers J., West R. & Webster, A. (2011a) Social Climate Along the Pathway of Care in Women's Secure Mental Health Service: Variation With Level of Security, Patient Motivation, Therapeutic Alliance & Level of Disturbance. Criminal Behaviour and Mental Health. 21: 202-214.

Long C.G., Langford V., Clay R., Craig. L. & Hollin. C.R. (2011b). Architectural Change & the Effects on the Perceptions of the Ward atmosphere in a Medium Secure Unit for Women. The British Journal of Forensic Practice. 13(3). 205-212.

Loughlm K. G. & Moore L. F. (1979) Using Delphi to achieve congruent objectives and activities in a pediatric department. *Journal of Medical Education.* 54(2**),** 101-10.

Matthews M.E., Mahaffey M.J., Lemer R.N. & Bunch, W,L. *(1975)* Profiles of future for administrative dieticians via Delphi technique. *Journal of the American Dietetic Association* 66(5), 494-499.

McCambridge J, de Bruin M, Witton J (2012) The Effects of Demand Characteristics on Research Participant Behaviours in Non-Laboratory Settings: A Systematic Review. *PLoS ONE* 7(6): e39116. doi:10.1371/journal.pone.0039116.

McIntosh C.J., Westbrook J., Sheldrick R., Surr C. & Hare D.J. (2012). The feasibility of Dementia Care Mapping on a neurorehabilitation ward. *Neuropsychological Rehabilitation*. 22(6): 920-941.

Mental Health Act 2007, ch.43. Retrieved from http://www.opsi.gov.uk/acts/acts2007/ukpga_20070012_en_1

Mental Health Advocacy Coalition. (2008). *Destination: recovery: Te ūnga ki uta: te oranga.* Auckland: Mental Health Foundation of New Zealand.

Mental Health Commission of Canada. (2012) *Changing directions, changing lives: the mental health strategy for Canada*. Calgary: Mental Health Commission of Canada.

Mezzina R., Davidson L., Borg M., Marin I., Topor A., & Sells, D. (2006). The social nature of recovery: discussion and implications for practice. *American Journal of Psychiatric Rehabilitation*, 9(1), 63-80.

MiddelbøeT., Schjodt,T., Byrsting, K. & Gjerris. A. (2001) Ward Atmosphere in Acute Psychiatric In-patient Care: Perceptions, Ideals & Satisfaction. *Acta Pschiatrica Scandinavica.* 103. 212-219.

Mitchell N. (2010) *The researcher participant relationship in educational research*. MPhil thesis, University of Nottingham.

Moos R.H., & Houts, P.S. (1968). Assessment of the social atmospheres of psychiatric wards. *Journal of Abnormal Psychology,* 73, 595–604.

Muijen M. (1999). Acute hospital care: ineffective, inefficient and poorly organised. *Psychiatric Bulletin*. 23; 257-259

National Clinical Practice Guideline (2007). *Number 42 National Collaborating Centre for Mental Health & Social Care Institute for Excellence National Institute for Health and Clinical Excellence*. A NICE–SCIE *Guideline on supporting people with dementia and their carers in health and social care.* London. The British Psychological Society & Gaskell.

National Health Service Employers (2008). *National Profiles for Nursing Services*. Retrieved from: http://www.nhsemployers.org/~/media/Employers/Documents/Pay%20and%20reward/Nursing_Services.pdf

National Health Service Digital: Psychiatry Doctors and Psychiatry Staff by Job, Role, Speciality and Area of Work. May 2116.

National Health Service Digital: 18 August, 2016, KH03 Report. NHS Engand.

National Institute for Mental Health in England, N. (2004). *Emerging Best Practice in Mental Health Recovery*. Leeds. National Health Service.

Nesset M. B., Røssberg, J., I. Almvik . R. & Friis S. (2009). Can a Focused Staff Training Programme Improve the Ward Atmosphere & Patient Satisfaction in a Forensic Psychiatric Hospital? A Pilot Study. *Scandinavian Journal of Caring Sciences*. 23: 117-124.

New Freedom Commission on Mental Health. (2003). *Achieving the promise: transforming mental health care in America. Final report.* Rockville: U.S. Department of Health and Human Services.

Nieuwenhuijsen K., Verbeek J.H.A., de Boer A.G.E., Blonk R.W.B., & van Dijk F.J.H. (2005). Validation of performance indicators for rehabilitation of workers with mental health problems. *Medical Care*, 43(10), 1034– 1042.

Nind M. (2008). Conducting qualitative research with people with learning, communication and other disabilities: Methodological challenges. *National Centre for Research Methods* NCRM/012.

Noordsy D, Torrey W, Mueser K, Mead S, O’Keefe C, Fox L (2002) Recovery from severe mental illness: an intrapersonal and functional outcome definition. *International Review of Psychiatry.* 14:318–326

O’Brien A.P., Boddy J.M., Hardy D.J., & O’Brien, A.J. (2004). Clinical indicators as measures of mental health nursing standards of practice in New Zealand. *International Journal of Mental Health Nursing,* 13(2), 78–88.

O'Connell M, Tondora J, Croog G, Evans AL, Davidson L. (2005). From rhetoric to routine: Assessing perceptions of recovery-oriented practices in a state mental health and addiction system. *Psychiatric Rehabilitation Journal*. 28(4):378-386.

O'HanlonK., Leigh A., Sheldrick R., Surr, C. & Hare D.J. (2014): The feasibility and validity of care mapping in the clinical neurosciences. Neuropsychological Rehabilitation: *An International Journal*, DOI: 10.1080/09602011.2014.978873

Onken S.J, Craig C.M., Ridgway P., Ralph R.O. & Cook J.A. (2007) An analysis of the definitions and elements of recovery: a review of the literature. *Psychiatry Rehabilitation Journal*; 31: 9–22

Osborn, Lloyd-Evans, Johnson, Gilbert, Byford, Leese and Slade (2010) Residential Alternatives to Acute In-patient Care in England: Satisfaction, Ward Atmosphere & Service User Experience. *The British Journal of Psychiatry.* 197. 41-45.

Oshima I., Mino Y. & Inomata Y. (2005). Effects of Environment Deprivation on Negative Symptoms of Schizophrenia: A Nationwide Survey in Japan's Psychiatric Hospitals. *Psychiatry Research.* 136; 163-171.

Oshima I., Mino Y. & Inomata Y. (2003). Institutionalisation & Schizophrenia in Japan: Social Environments & Negative Symptoms. *British Journal of Psychiatry*. 183; 50-56.

Orne M.T. (1959). The nature of hypnosis: Artifact and essence. Journal of Abnormal and Social Psychology: 58, 277-299.

Orne M.T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*. 17:776–783.

Paterson D.L., Potoski B. & Capitano B. (2002). Measurement of adherence to antiretroviral medications. *Journal of Acquired Immune Deficiency Syndromes*. 31(3): 103–S106.

Pearson K. (1900). *Philosophical Magazine Series 5*. 50 (302): 157–175. doi:10.1080/14786440009463897.

Persaud M. & Jaycock S. (2001). Evaluating Care Delivery: The Application of Dementia Care Mapping in Learning Disability Residential Services. *Journal of Learning Disabilities*. 5 (4): 345–52.

Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. (2006). Guidance on the conduct of narrative synthesis in systematic reviews. Results of an ESRC funded research project (unpublished report). University of Lancaster.

Prochaska J.O. & DiClemente C. C. (1982) Toward a more integrative model of change. Psychotherapy: *Theory, Research & Practice.* 19(3): 1982, 276-288.

Quirk A. & LelliotP. (2001.) What do we know about life on acute psychiatric wards in the UK? A review of the research evidence. *Social Science and Medicine*. 53(12); 1565-1574.

Quirk A. Lelliott P.& Seale C. (2006). The permeable institution: an ethnographic study of three acute psychiatric wards in London. *Social Science and Medicine*. 63; 2105-2117.

Reid N. G. (1988). The Delphi technique: its contribution to the evaluation of professional practice. In Professional Competence and Quality Assurance in the Caring Professions (Ellis R. ed.). Chapman & Hall, London.

Røssberg J.I. & Friis S. (2004) Patients' & Staff's Perception of the Psychiatric Ward atmosphere. *Psychiatric Services*. 55; 798-803.

Røssberg J.I., Melle,I., Opjordsmoen S. & Friis, S (2006). Patient Satisfaction & Treatment Environment: A 20-year Follow-up Study from an Acute Psychiatric Ward. *Nordic Journal of Psychiatry.* 60; 176-180.

Røssberg J.I., Melle,I., Opjordsmoen S. & Friis S. (2008). The Relationship Between Staff Members' Working Conditions & Patients' Perceptions of the Environment. *International Journal of Social Psychiatry*. 54(5); 437-446

Ros N. Van der Helm P., Wissink I., Stams G.J. & Schaftenaar P. (2013). Institutional Climate & Aggression in a Secure Psychiatric Setting. *The Journal of Forensic Psychiatry & Psychology*. 24(6); 713-727.

Royal College of Psychiatrists, Royal College of Physicians of London, British Medical Association. (2001). *Mental Illness: Stigmatisation and Discrimination within the Medical Profession (CR91*). London. Royal College of Psychiatrists.

Schalast N, Redies M, Collins M, Stacey J, & Howells K. (2008). EssenCES, a short questionnaire for assessing the social climate of forensic psychiatric wards. *Criminal Behaviour Mental Health.* 18(1):49-58. doi: 10.1002/cbm.677.

Schjødt T., Middelboe, T., Mortensen E.L. & Gjerris A. (2003). Ward Atmosphere in Acute Psychiatric In-patient Care: Differences & Similarities Between Patient & Staff Perceptions. *Nordic Journal of Psychiatry*. 57; 215-220.

Schranks B. & Slade M. (2007). Recovery in psychiatry. *Psychiatric Bulletin*: 31:321-325. doi: 10.1192/pb.bp.106.013425

Shanks, Williams, Leamy, Bird, Le Boutillier & Slade (2013). Measures of Personal Recovery: A Systematic Review. *Psychiatric Services*. 64(10): 974-980.

Silverstein S.M. & Bellack, A. (2008). A scientific agenda for the concept of recovery as it applies to schizophrenia. *Clinical Psychology Review.* 28: 1108–24.

Simmons P., Hawley C.J., Gale, T.M. & Sivakumaran T. (2010). Service user, patient, client, user or survivor: describing recipients of mental health services. *The Psychiatrist*. 34; 20-23.

Slade M. (2002). What outcomes to measure in routine mental health services, and how to assess them: a systematic review. *Australian & New Zealand Journal of Psychiatry*. 36: 743–53.

Slade M. (2009) *Personal recovery and mental illness: a guide for mental health professionals.* Cambridge: Cambridge University Press.

Slade M. (2010). Mental illness and well-being: the central importance of positive psychology and recovery approaches. *BMC Health Services Research*. 10:26.

Slade M., Amering M., Farkas M., Hamilton B., O'Hagan M., Panther G., Perkins R., Shepherd G., Tse S. & Whitley R., (2014). Uses and abuses of recovery: implementing recovery-oriented practices in mental health systems. *World Psychiatry*. 13(2); 12–20.

Slade M & Davidson L: *Recovery as an integrative paradigm in mental health*. In Graham Thornicroft (Ed). Oxford Textbook of Community Mental Health (2011). OUP Oxford.

Slade M, Hayward M (2007) Recovery, psychosis and psychiatry: research is better than rhetoric. Acta Psychiatr Scand 116:81–83

Slade M, Leamy M, Bacon F, Janosik M, Le Boutillier C, Williams J, Bird V (2012) International differences in understanding recovery: systematic review. *Epidemiology and Psychiatric Sciences*. 21, 353-364.

Spaniol, L., Wewiorski, N., Gagne, C., & Anthony, W. (2002). The process of recovery from schizophrenia. *International Review of Psychiatry*. 14(4), 327-236.

Spivey B.E. (1971). A technique to determine curriculum content. *Journal of Medical Education*. 46, 269-271.

Tew J., Ramon S., Slade M., Bird V., Melton J., & Le Boutillier, C. (2012). Social Factors and Recovery from Mental Health Difficulties: A Review of the Evidence. *British Journal of Social Work.* 42(3), 443-460. doi: 10.1093/bjsw/bcr076

ThornicroftG. & Bebbington P. (1989) Deinstitutionalisation – from hospital closure to service development. *British Journal of Psychiatry*. 155; 739-753.

Turner D. (2002). Mapping the routes to recovery. *Mental Health Today*. July: 29-30.

Warner R. (2009) Recovery from schizophrenia and the recovery model. *Current Opinions in Psychiatry*. 22:374–80.

Wellesley Institute (2009). *Mental Health "Recovery" Study Working Group. Mental health "recovery": users and refusers*. Toronto: Wellesley Institute, 2009.

Westbrook J.L., McIntosh C.J., Sheldrick R., Surr C. & Hare D.J. (2013). Validity of Dementia Care Mapping on a neuro-rehabilitation ward: Q-methodology with staff and patients. *Disability and Rehabilitation. An international, multidisciplinary journal*. 1-8.

Wilhelm W.J. (2001). Alchemy of the Oracle: the Delphi technique. *The Delta Pi Epsilon Journal.* 43:6–26.

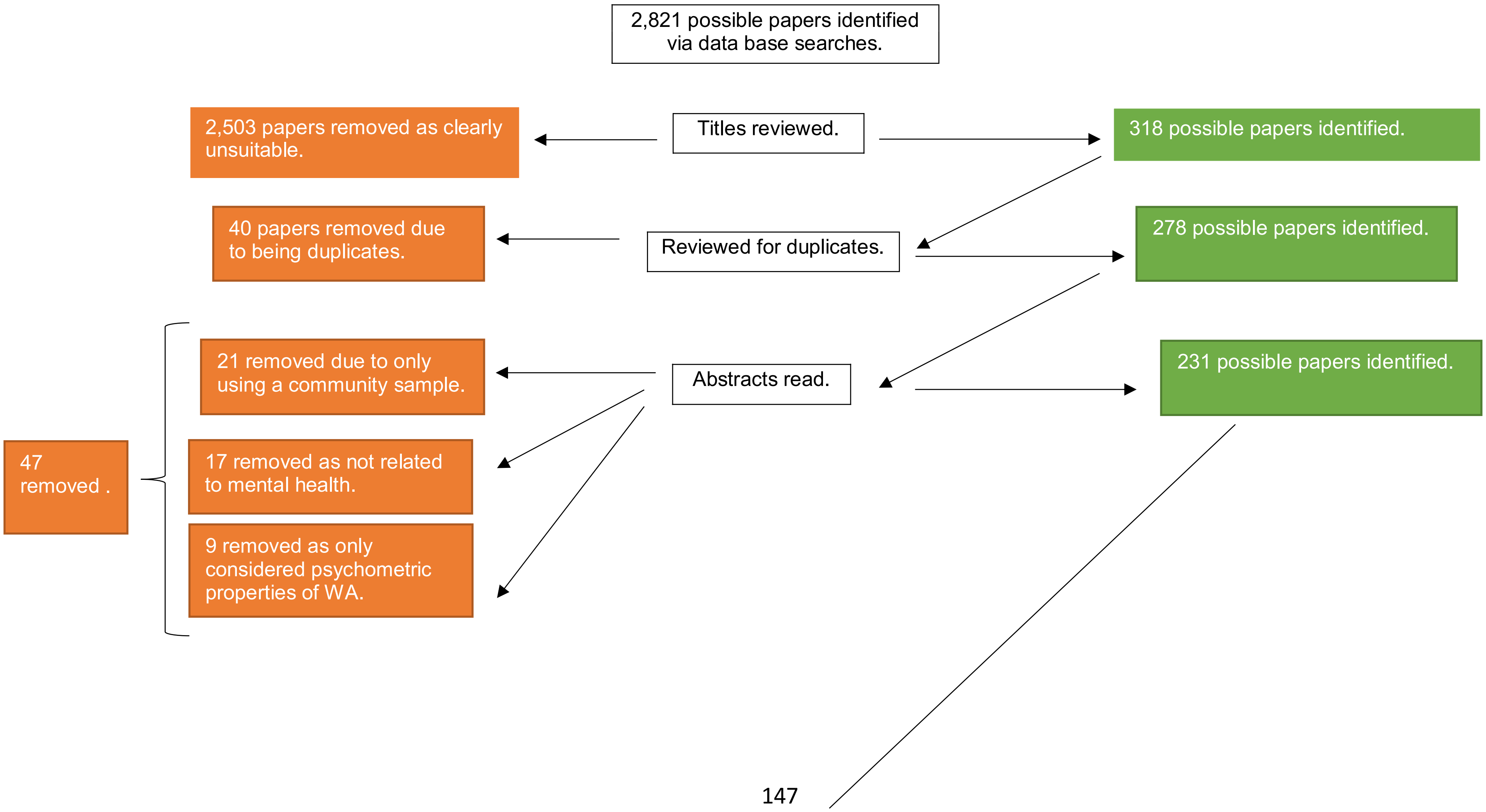
Williams PL & Webb C (1994): The Delphi technique: a methodological discussion. *Journal of Advanced Nursing.* 19: 180-6.

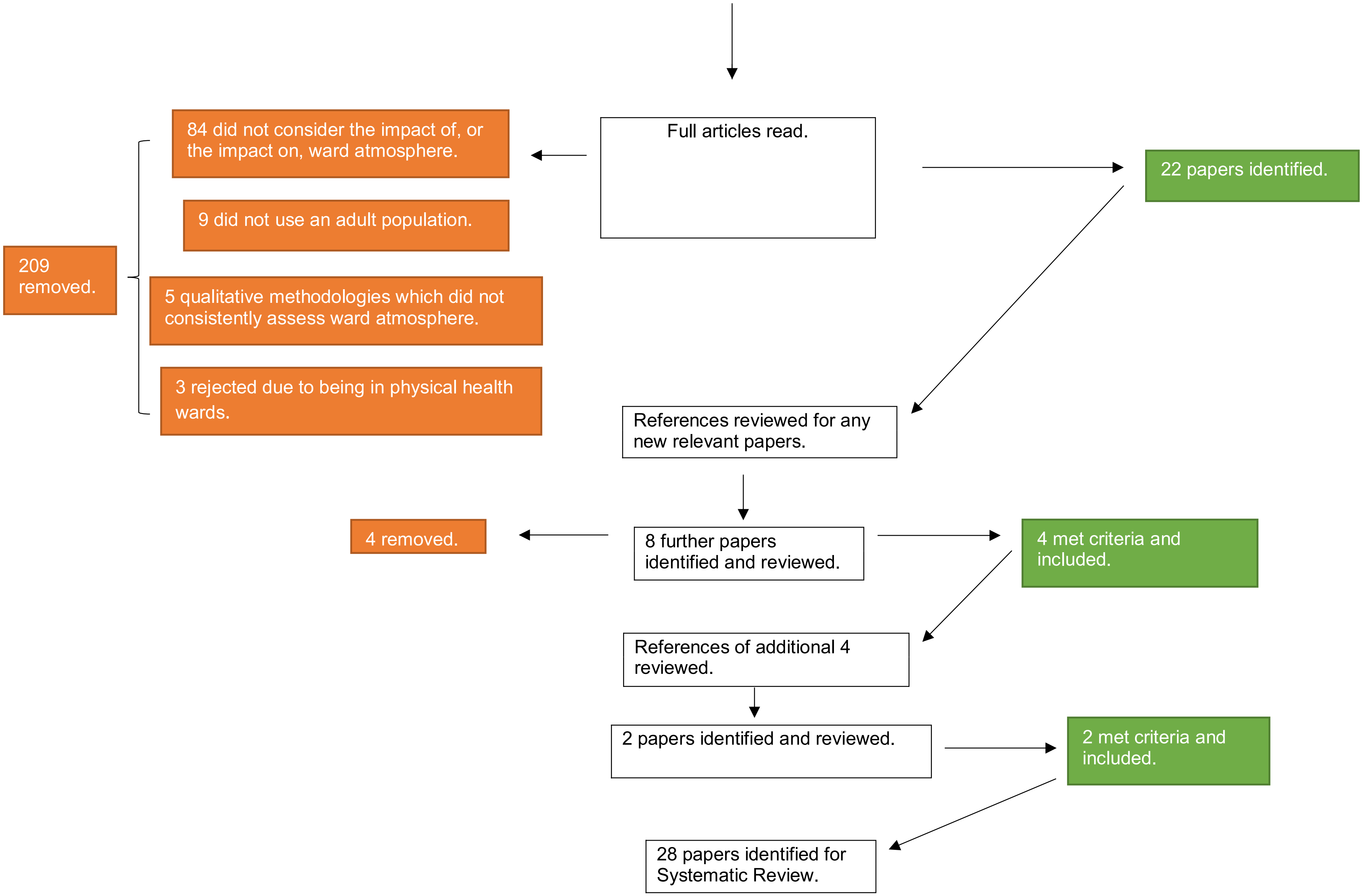
Woolley R.J., Young J.B., Green J.R. & Brooker D.J. (2008). The feasibility of care mapping to improve care for physically ill older people in hospital. *Age Ageing*. 37:390–5.

World Health Organization. (1953*). The community mental hospital: Third report of the expert committee on mental health*, no. 73. Geneva: Switzerland.

Appendices

Appendix 1: Flow chart of review process




Full articles read.
22 papers identified.
84 did not consider the impact of, or the impact on, ward atmosphere.
9 did not use an adult population.
209 removed.
5 qualitative methodologies which did not consistently assess ward atmosphere.
3 rejected due to being in physical health wards.
References reviewed for any new relevant papers.
4 removed.
8 further papers identified and reviewed.
4 met criteria and included.
References of additional 4 reviewed.
2 papers identified and reviewed.
2 met criteria and included.
28 papers identified for Systematic Review.

Appendix 2: Quality Review framework, and scores for papers in the Systematic Review

	Braham, Heasley and Akins (2013)	Brunt (2008)	Kerfoot, Bamford and Jones (2012)	Livingston, Nijdam-Jones and Brink (2012)	Long, Anagnostakis, Fox, Silaule, Somers, West and Webster (2011a)	Long, Lanford, Clay, Craig and Hollin (2011b)	Nesset, Rossberg, Almvik and Friis (2009)	Osborn, Lloyd-Evans, Johnson, Gilburt, Byford, Lee and Slade (2010)	Livingston, Nijdam-Jones, Lapsley, Calderwood and Brink (2013)	Middelboe, Schjodt et al (2001)
Does the title reflect the content?	2	2	1	1	2	2	2	2	2	1
Are the authors credible?	1	2	1	2	2	2	2	2	2	2
Does the abstract summarise the key components?	2	1	1	1	2	2	2	2	2	2
Is the rationale for undertaking the research clearly outlined?	2	2	1	2	2	2	2	2	2	2
Is the literature review comprehensive and up-to-date?	1	1	2	2	2	1	1	1	2	1
Is the aim of the research clearly stated?	2	2	1	1	2	2	2	2	2	2
Are all ethical issues identified and addressed?	1	1	0	1	2	0	1	1	1	0
Is the methodology identified and justified?	1	1	1	2	2	1	2	2	2	2

Is the study design clearly identified, and is the rationale for choice of design evident?	2	2	2	2	2	1	2	1	2	2
Is there an experimental hypothesis clearly stated, are the key variables clearly defined?	1	1	1	1	1	1	2	1	2	1
Is the population identified?	1	2	1	2	2	2	2	1	2	2
Is the sample adequately described and reflective of the population?	1	2	1	2	2	2	2	1	1	1
Is the method of data collection valid and reliable?	1	2	1	2	2	2	2	2	2	2
Is the method of data analysis valid and reliable?	2	2	1	2	2	1	2	2	2	2
Are the results presented in a way that is appropriate and clear?	2	2	2	1	2	2	2	2	2	2
Is the discussion comprehensive?	2	2	2	2	2	2	2	1	2	2
Are the results generalisable?	1	1	1	1	2	1	2	2	2	2
Is the conclusion comprehensive?	1	2	1	1	1	1	1	2	2	2
Total	26	30	21	28	34	27	33	29	34	30

	Kuosmanen, Hatonen et al (2006)	Bressington, Stewart et al (2011)	Gjerden & Moen (2001)	Rossberg, Melle et al (2008)	Rossberg, Melle et al (2005)	Rossberg & Friis (2004)	Brunt & Rask (2007)	Jansson & Eklund (2001)	Schjodt, Middelboe et al (2003)	Jansson & Eklund (2002b)
Does the title reflect the content?	2	2	1	2	2	1	2	2	2	1
Are the authors credible?	2	2	2	2	2	2	2	2	2	2
Does the abstract summarise the key components?	2	2	1	2	2	2	2	1	2	2
Is the rationale for undertaking the research clearly outlined?	2	2	2	2	2	2	2	2	2	2
Is the literature review comprehensive and up-to-date?	2	2	1	1	1	2	1	2	1	2
Is the aim of the research clearly stated?	2	2	1	2	2	2	2	2	2	2
Are all ethical issues identified and addressed?	1	1	0	0	1	0	1	0	0	1
Is the methodology identified and justified?	2	1	2	2	2	2	2	2	2	2
Is the study design clearly identified, and is the rationale for	2	2	2	2	2	2	2	2	1	2

choice of design evident?										
Is there an experimental hypothesis clearly stated, are the key variables clearly defined?	1	1	2	2	2	2	1	1	1	1
Is the population identified?	2	2	2	2	2	2	2	2	2	2
Is the sample adequately described and reflective of the population?	1	2	1	1	1	1	2	1	1	1
Is the method of data collection valid and reliable?	2	2	2	2	2	2	2	2	2	2
Is the method of data analysis valid and reliable?	2	1	2	2	2	2	2	2	2	2
Are the results presented in a way that is appropriate and clear?	2	2	2	2	2	1	2	2	2	2
Is the discussion comprehensive?	2	2	2	2	2	1	2	2	2	2
Are the results generalizable?	1	1	1	2	2	2	1	1	2	1
Is the conclusion comprehensive?	2	1	2	2	2	1	1	2	2	2
Total	32	30	28	32	33	29	31	30	30	31

	Campbell, Allan et al (2014)	Dickens, Suesse et al (2014)	Oshima, Mino et al (2005)	Oshima, Mino et al (2003)	Beazley & Gudjonsson (2011)	Jorgensen, Romma et al (2008)	Johansson & Eklund (2004)	Ros, Van der Helm et al (2013)
Does the title reflect the content?	2	2	2	1	1	2	2	2
Are the authors credible?	2	2	2	2	2	2	2	2
Does the abstract summarise the key components?	2	2	2	1	2	1	2	2
Is the rationale for undertaking the research clearly outlined?	2	2	2	2	2	2	2	2
Is the literature review comprehensive and up-to-date?	2	2	1	1	2	2	2	2
Is the aim of the research clearly stated?	2	2	1	1	2	2	2	2
Are all ethical issues identified and addressed?	1	2	0	0	0	2	1	0
Is the methodology identified and justified?	2	2	1	1	2	2	2	2
Is the study design clearly identified, and is the rationale for choice of design evident?	2	2	2	2	2	2	2	2
Is there an experimental	1	2	1	1	2	2	1	2

hypothesis clearly stated, are the key variables clearly defined?								
Is the population identified?	2	2	2	2	2	2	2	2
Is the sample adequately described and reflective of the population?	1	1	1	1	1	2	1	2
Is the method of data collection valid and reliable?	2	2	2	2	2	2	2	1
Is the method of data analysis valid and reliable?	2	2	2	2	2	2	1	2
Are the results presented in a way that is appropriate and clear?	1	2	1	2	2	2	2	1
Is the discussion comprehensive?	2	2	2	2	2	2	2	2
Are the results generalisable?	1	2	2	2	1	2	2	2
Is the conclusion comprehensive?	1	2	2	2	2	2	2	2
Total	30	35	28	27	31	35	32	32

Appendix 3: Participant Information Sheet: Delphi method

PARTICIPANT INFORMATION SHEET

PROJECT TITLE
Assessing interpersonal interactions in inpatient mental health settings as a measure of the recovery orientation of the service.

INVITATION
My name is Christian Williams and I am training to be a Clinical Psychologist. As part of this I am carrying out a piece of research, which I would be grateful if you would consider taking part in. The research you are being asked to take part in is aiming to develop new ways of assessing the recovery orientation of services in inpatient rehabilitation mental health services. To do this a new tool which uses observation is being designed. The tool works on take observations of interactions between staff and clients in the above mentioned setting and then coding these interactions based upon operationalised items which relate to the delivery of a recovery orientated model of care.

It is the validation of these items which you would be asked to comment on in relation to their relevance and importance of the items which have been developed for this proposed tool.

This projected has had ethical approval from Cardiff and Vale University Health Board Research and Development department as well as from Cardiff University.

WHAT WILL HAPPEN
After you have read this information leaflet, you will have a chance to ask any questions you might have. If you are then happy to take part in this study you will be asked to sign a consent form which says you are happy to take part.

You will then be sent the list of items which have been developed to create the tool; this can be sent to you either via post or e-mail. These items are based on recent qualitative research into the recovery process (Leamy *et al*, 2011; Bonney & Stickley 2008). You will be asked to rate each item on a scale of one to seven with regards its relevance and importance in relation to the setting you work in. You will also be asked to consider if there are any other types of interpersonal interactions which you deem to be important which are not covered by the items.

A consensus will then be taken of all the scores received, the items which fall on or above the mean and any new items which are added to create a new list, which will then be sent out to you, and you will be asked to rate them on the same scales for a second time.

PARTICIPANTS' RIGHTS
You may decide to withdraw without explanation at any point. If this is after the first set of items have been rated, the rating you give may have already been used, if so this data will not be able to be

removed from the study. If you have any questions as a result of reading this information sheet, please contact one of the via the contact details given below.

BENEFITS AND RISKS

There are no known risks or benefits for you in taking part in this research. However, with your help, if the tool proves helpful it may improve the quality of care which can be provided in the future.

COST, REIMBURSEMENT AND COMPENSATION

Your participation in this study is voluntary; however, your time would be greatly appreciated.

CONFIDENTIALITY/ANONYMITY

We will not ask for any identifying information about you. Information you return will be anonymised (such as names or email addresses being removed from scores).

FOR FURTHER INFORMATION

Dr Andrew Vidgen or Christian Williams will be glad to answer your questions about this study at any time. You can contact them on Andrew.vidgen@wales.nhs.uk or WilliamsCR11@cardiff.ac.uk

The final results of this study will be disseminated to anyone who leaves contact details with the research team, or requests the findings at a later date, in which case a copy of the study will be sent out to you.

Date: 01.02.2015
Version 4a

Appendix 4: Consent form, Delphi method

CONSENT FORM

Assessing interpersonal interactions in inpatient mental health settings as a measure of the recovery orientation of the service

Name of Researcher: Christian Williams

Please initial box

1. I confirm that I have read the information sheet dated 01.02.2015 (version 4a) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected. ☐
3. I agree to take part in the above study. ☐

______________________	______________________	______________________
Name of Participant	Date	Signature
______________________	______________________	______________________
Name of Person taking consent	Date	Signature

Appendix 5: Demographic information sheet. Delphi method

Participant Information Form

Please Provide the Following Information

Professional background………...

Current Place of work/Service worked in..........................……….........……………………

Length of Time Working in Rehabilitation and Recovery Services

.....................Years ...…………Months

Gender: Male/Female (please circle)

Age: (please circle)	18-30	31-45	46-60	61+

Thank you very much

Codes and Exemplars

Thank you very much for agreeing to be involved with the development of a new observational tool which is aimed to measure the recovery orientation of inpatient rehabilitation mental health services.

Below is the list of ‘codes’ which have been developed from qualitative research into the Recovery Orientation, and ‘exemplars’ which are operationalised descriptions of the behaviours which may be observed within clinical settings. The codes and exemplars which relate to one another have the same number.

Please rate each code on two dimensions: Firstly, on its Relevance to your service, in relation to how often the type of behaviour the code describes happens (1= no relevance to your service and 7 = very relevant to your service). Secondly, on how much Importance you believe the code has, in relation to the impact it has with regards recovery orientation of the service. Please do NOT consider whether you think that type of behaviour should or should not happen. It is expected that codes may have different ratings on each dimension e.g. it happens a lot in your service but is unimportant, or that it occurs rarely but has a great impact.

Please list any further codes you feel are missing and any additional items you can think of in relation to any of the codes either already developed or codes of your own.

Thank you.

Codes

1. Encouraging engagement in activities.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

2. Forcing engagement in activities.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

3. Encouraging the development of independence.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

4. Non-verbal dismissal

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

5. Discussing client’s personal information (in public settings).

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

6. Offering choice.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

7. Tokenistic involvement

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

8. Staff initiated non-clinical interactions.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

9. Positive responses to client initiated interactions.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

10. Negative responses to client initiated interactions.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

11. Joint involvement in task/activities.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

12. Clients not being engaged with.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

13. Interactions having an ‘illness focus’.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

14. Seeking input from clients.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

15. Not seeking input from clients.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

16. Active listening skills.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

17. Staff not being in the moment with clients.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

18. Recognition of communication needs.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

19. Ignoring communication needs.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

20. Staff recognising strengths and interests.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

21. Punitive interactions.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

22. Set meal times.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

23. Set bed times.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

24. Set getting up times.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

25. Set smoking times.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

26. Encouragement/practical support of pro-social activities, including off the unit.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

Exemplars

1. Discussing available activities and what the client might like to do. Offering support with carrying out tasks e.g. I noticed you have some washing to do, shall I come and give you a hand?
2. Demanding people take part in certain activities e.g. you need to do your washing now. We are all playing bingo, came and join in.
3. Example of doing with, not doing for; ‘shall I help you with X’ as opposed to ‘let me do Y for you, it will be quicker.’
4. Gesturing five minutes with a hand through the window when a client approaches the nurses’ station window.
5. Discussing private aspect of the client’s health/care in communal areas. Making reference to personal events in communal areas e.g. ‘Hurry up the taxi is here for your hospital appointment.’
6. Either open choice: ‘What would you like to do today/this week etc. Or selective choice where necessary e.g. ‘Would you like to do X or Y?’7. Brief ward round where clients are involved minimally e.g. ‘how are you? Everything going ok? Anything you want to ask, no? Good…’
8. Initiating general conversation or engaging in any activity which is not directly related to clinical care e.g. NOT medication/self-care/food.
9. Staff taking the lead from the client
10. Staff brushing off the advance e.g. ‘Not now, I am busy…’
11. Do something with, not to or for someone. ‘Let’s do X together…’
12. Staff members walking past a client sat in the communal area, without making effort to interact.
13. Focus of interactions on symptoms/medication etc. rather than the individual as a person rather than a condition
14. Either the service or individual staff looking for input from clients on issues, rather than simply passing on information. Also relates to providing choice.
15. Telling clients ‘how things are’.
16. Paying attention, giving eye contact, nodding, responding where appropriate.
17. Looking at watches, talking to someone else, staring off into space, checking your phone etc.
18. Language barriers; does someone need an interpreter? Is it more effective to communicate with someone via pictures? Does someone use sign language? Can that client read the care plans you have given them to read? Individual staff knowing and communicating appropriately AND service provision for things such as interpreters.
19. Not recognising the factors listed above.
20. Using information of someone’s strengths to help support them e.g. ‘….well you could try making a list of the things you want to get done if you are worried about not having time to do it all, you are good at planning….’
21. ‘If you don’t do X/continue to do Y, you will…. Miss your next smoking break….not be allowed section 17 leave etc.’
22. Are choices with meals given? Do people eat together? Are the clients fed ‘hotel style’? Are there facilities for clients to prepare their own food/snacks/hot drinks? Do clients have input into their diet?
23. Are clients sent to bed/bedrooms at certain times? Are TV’s shut off? Do communal areas become off limits?
24. Do clients have to get up at a certain time? Are doors locked during the day? Is this individual (e.g. agreed by client/staff/family etc. in a care plans?) or a whole unit system?

25. Can clients control when and how much they smoke?
26. Is going out into the community encouraged/valued and facilitated or is ‘we don’t have the staff’ more common?

Please list any additional codes below. Continue on a separate sheet, if necessary.

Please list additional exemplars below. Please use the numbers to indicate which code the exemplar refers to.

Codes and Exemplars, Round 2

Thank you very much for continuing to be involved with the development of a new observational tool.

Below is the list of 'codes', which have been edited in light of the feedback already received; we would like you to re-rate them.

Please rate each code on two dimensions: Firstly, on its Relevance to your service, in relation to how often the type of behaviour the code describes happens (1= no relevance to your service and 7 = very relevant to your service). Secondly, on how much Importance you believe the code has, in relation to the impact it has with regards recovery orientation of the service. Please do NOT consider whether you think that type of behaviour should or should not happen. It is expected that codes may have different ratings on each dimension e.g. it happens a lot in your service but is unimportant, or that it occurs rarely but has a great impact.

Please list any additional exemplars you can think of in relation to any of the codes.

Thank you.

Codes

Code 2: Forcing engagement in activities.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

Code 4: Non-verbal dismissal

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

Code 5: Discussing **client's** personal information (in public settings).
Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

Code 7: Tokenistic involvement
Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

Code 10: Negative responses to client initiated interactions

Not relevant						Very relevant
1	2	3	4	5	6	7

No importance						Very important
1	2	3	4	5	6	7

Code 12: Clients not being engaged with.

Not relevant						Very relevant
1	2	3	4	5	6	7

No importance						Very important
1	2	3	4	5	6	7

Code 13: Interactions having an ‘illness focus’.

Not relevant						Very relevant
1	2	3	4	5	6	7

No importance						Very important
1	2	3	4	5	6	7

Code 15: Not seeking input from clients.

Not relevant						Very relevant
1	2	3	4	5	6	7

No importance						Very important
1	2	3	4	5	6	7

Code 17: Staff not being in the moment with clients.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

Code 19: Ignoring communication needs.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

Code 21: Punitive interactions.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

Code 22: Set meal times.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

Code 23: Set bed times.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

Code 24: Set getting up times.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

Code 25: Set smoking times.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

Code 27: Offering medication (PRN) instead of talking.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

Code 28: Empathy shown in response to distress.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

Code 29: Doing with, not doing for.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

Code 30: Exploring options with a client.

Not relevant Very relevant

1 2 3 4 5 6 7

No importance Very important

1 2 3 4 5 6 7

Exemplars

2. Demanding people take part in certain activities e.g. 'you need to do your washing now. We are all playing bingo, came and join in.'
3. Example of doing with, not doing for; "shall I help you with X" as opposed to "let me do Y for you, it will be quicker."

5. Discussing private aspect of the client's health/care in communal areas. Making reference to personal events in communal areas e.g. 'Hurry up the taxi is here for your hospital appointment.'
7. Brief ward round where clients are involved minimally e.g. 'how are you? Everything going ok? Anything you want to ask, no? Good…'
10. Staff brushing off the advance e.g. 'Not now I am busy…'
12. Staff members walking past a client sat in the communal area, without making effort to interact.
13. Focus of interactions on symptoms/medication etc. rather than the individual as a person rather than a condition
15. Telling clients 'how things are.'
17. Looking at watches, talking to someone else, staring off into space, checkingyour phone etc.
19. Not recognising the factors listed above.

21. 'If you don't do X/continue to do Y, you will…. Miss your next smoking break….not be allowed section 17 leave etc.'
22. Are choices with meals given? Do people eat together? Are the clients fed 'hotel style'? Are there facilities for clients to prepare their own food/snacks/hot drinks? Do clients have input into their diet?
23. Are clients sent to bed/bedrooms at certain times? Are TVs turned off? Do communal areas become off limits?
24. Do clients have to get up at a certain time? Are doors locked during the day? Is this individual (e.g. agreed by client/staff/family etc. in a care plan?) or a whole unit system?
25. Can clients control when and how much they smoke?
27. 'Would you like some PRN?' being the first response to a client's distress.
28. Recognising that the client is in distress, and validating their feelings.
29. Offering support where necessary, to allow the client to achieve their aim with as little aid as possible; 'what can I do to help you finish X.'
30. 'I understand that Y is problem, perhaps we could discuss some ways in which you might like to deal with it?'

Please list additional exemplars below. Please use the numbers to indicate which code the exemplar refers to.

Appendix 8: Information for Patients

NB: The name of the local collaborator has not been completed in this version as the name will vary depending on which units the information sheets are being distributed within. Everything else will remain the same for each Health Board.

INFORMATION LEAFLET

PROJECT TITLE
Assessing interpersonal interactions in inpatient mental health settings as a measure of the recovery orientation of the service.

INFORMATION
My name is Christian Williams and I am training to be a Clinical Psychologist. As part of this I am carrying out a piece of research, which I have asked some members of staff working on this ward to be part of.

This project is hoping to create a new way of looking at how care is provided in in-patient units. It works by recording the way staff members work with patients.

This research is only interested in whether this new method of assessment works. It is not interested in telling staff members or patients whether they are performing well or not.

This project has been reviewed and approved by an NHS ethics committee and Cardiff and Vale University Health Board.

WHAT WILL HAPPEN
Members of staff will be asked to complete a questionnaire. Observations will take place in the unit with members of staff who have agreed to take part in the study. The observations will only be looking at staff members who have agreed to take part. The observations will not look at patients, visitors or members of staff who do not want to be involved. During the observations, a researcher will sit in public areas of the unit and observe how members of staff interact with people on the ward. The researcher will use time sampling: this means they will observe staff members for five minutes, and then not observe anyone for five minutes. This will last for an hour at a time. The research is only interested in what staff members do, and no one else. Whilst observations are being carried out, the researcher will write down what happens. This information will then be coded using our new tool. The observations will be anonymous; no information will be taken which can identify anyone involved in the interactions. No observations will be carried out in private areas and the observations will only take place on the unit.

WHO SHOULD I CONTACT FOR FURTHER INFORMATION?

If you have any further questions, please feel free to speak to a researcher when you see them, or you can contact Dr Andrew Vidgen or Christian Williams, who will be glad to answer your questions about this study at any time. You can contact them on Andrew.vidgen@wales.nhs.uk or WilliamsCR11@cardiff.ac.uk or on 02920 870582. Otherwise, you can speak to NAME OF LOCAL COLLATOR who is involved with the project.

The final results of this study will be shared across all units that are involved. However, if you would like your own copy, you can email either of the addresses given above and a copy of the study will be sent out to you.

Date: 18/04/2016
Version 3

Appendix 9: Participant Information Sheet: Part 2, ward selection and observations

---

NB: The name of the local collaborator has not been completed in this version as the name will vary depending on which units the information sheets are being distributed within. Everything else will remain the same for each Health Board.

PARTICIPANT INFORMATION LEAFLET

PROJECT TITLE
Assessing interpersonal interactions in inpatient mental health settings as a measure of the recovery orientation of the service.

INVITATION
My name is Christian Williams and I am training to be a Clinical Psychologist. As part of this I am carrying out a piece of research, which I would be grateful if you would consider taking part in. The aim of the research is to test if a new way of assessing the recovery orientation of a service in inpatient Rehabilitation Mental Health services is effective or not. We have developed a tool and now wish to find out if it works.

This new tool works by looking at the interactions between staff and clients to see if these interactions reflect factors which research suggests may support the Recovery process. We aim to assess the tool by finding two wards which have different Recovery-orientations (as assessed by a standardised measure) and then seeing if the new tool can tell them apart.

This research is not assessing the quality of care you or your ward provides, it is looking to assess the new tool.

This project has been reviewed and approved by Cardiff and Vale University Health Board and from Cardiff University.

WHAT WILL HAPPEN

After you have read this information leaflet, you will have chance to ask any questions you might have. You will be given a minimum of 2 weeks to decide if you wish to take part or not.

There are two parts to this study. You will be asked to take part in either both parts, or the first part of the study only.

If you choose to take part in the study. All of the information collected will be anonymous, which means that no one will be able to know what answers you gave, or what information came from you.

In the first part of the study, all members of staff working on the unit who agree to take part in the study will be asked to complete a demographic information sheet (your job title, age range, gender, years in post), this is being collected so that the study can comment on how reflective the sample within the study is, and a questionnaire which is designed to measure the extent to which recovery-supporting practices are evident, in your opinion, in your service; it is called the Recovery Self-Assessment, or RSA. This is a 36-item questionnaire, each item is rated on a 5-point Likert scale. It is estimated to take between 10 -15 minutes to complete. Your answers will be anonymous and you will not be asked for any personal details which could link you to your questionnaire.

The demographic information will be linked to which ward the information came from (using a code) but there will be no way to link the demographic information and the questionnaires.

The average scores for the RSA will be calculated for each unit. The two units which have the scores which are most different from on another will be asked to take part in the second part of the project. If the unit you work on is not one of these wards, you will not be asked to do anything further in the project.

Observations will take place in the unit. This won't require you to do anything. A researcher will sit in public areas of the unit and observe the interactions members of staff and clients which might occur. The researcher will observe some interactions and not others. Whilst observations are being carried out the researcher will write down what happens between the members of staff and the clients. This information will then be coded using our new tool. The observations will be anonymous; no information will be taken which can identify anyone involved i the interactions. No observations will be carried out in private areas and the observations will only take place in the unit.

TIME COMMITMENT

It should take about 15 minutes to complete the questionnaire and demographic sheet.

The observations will last for 1 hour at a time, and will occur between once and ten occasions. You will not be required to do anything for the observations; just act naturally.

PARTICIPANTS’ RIGHTS

You have the right to stop being a part of the research at any point. You may withdraw any information you have provided up until the point it is handed in. Once you have handed in your questionnaires your data is anonymous, which means your name will not be on any of the data, it will therefore not be possible to exclude your data at this point. However, you may still withdraw from taking any further part in the study if you wish.

Whether you take part in the study or not will have no impact on your work or any of your rights. No one else will be told whether you decide to be involved with this research or not.

You have the right to have your questions about the procedures answered. If you have any questions as a result of reading this information sheet, you should ask the researcher before the study begins, you may do this whilst the researcher is on the unit, or via the email addresses provided.

BENEFITS AND RISKS

There are no known risks for you in this study. There are no known personal benefits for you in taking part in this study, however with your help, this tool may be able to help improve the quality of care which can be provided.

COST, REIMBURSEMENT AND COMPENSATION

Taking part in this study will not cost you money. Your participation in this study is voluntary; however, your agreement to take part is greatly appreciated.

CONFIDENTIALITY/ANONYMITY

We will ask for some information about you; this will be your job title and how long you have worked on the unit. We will not ask for any further information such as your name, address or date of birth.

To link the questionnaires and the observations to the unit without identifying it, the researcher will assign each unit a number. Only the researcher will know which unit was assigned which number. This information will not be shared with any other parties.

No identifying features (such as individual's name, date of birthday, address etc) will be recorded. Data will be identified based upon which ward it comes rather than on an individual level. This will be done via a code rather than naming the ward.

The main researcher will have access to all of the data. The other members of the research team will have access to the anonymised data. The data will not be shared with any third party.

During the study, all data will be kept in locked premises on Cardiff and Vale University Health Board property in a locked cabinet. The main research (Christian Williams) and academic supervisor (Dr Andrew Vidgen) will have access to this.

Non-identifiable data will be kept securely in locked premises for 15 years after the end of the study as per Cardiff University policy. The main researcher (Christian Williams) and academic supervisor (Dr Andrew Vidgen) will have access to this.

WHO SHOULD I CONTACT IF I AM UNHAPPY ABOUT SOMETHING IN THE STUDY, OR IF SOMETHING GOES WRONG?

If you are unhappy about any aspect of your involvement in the study and wish to make a complaint, or raise a concern, you can either speak directly with the student researcher, Christian Williams (WilliamsCR11@cardiff.ac.uk), her supervisor, Dr Andrew Vidgen (Andrew.vidgen@wales.nhs.uk 02920870582) or the local collaborator INSERT LOCAL COLLABORATOR'S NAME AND CONTACT DETAILS HERE. If you would like to speak with someone who is not involved in the study, then you can contact Dr Reg Morris (reg.morris@wales.nhs.uk 02920870582).

In the unlikely event that something goes wrong and you are harmed as a result of taking part in the research, then you may be entitled to compensation, but you may have to pay your own legal costs. Insurance for this study is provided by Cardiff University.

FOR FURTHER INFORMATION

Dr Andrew Vidgen or Christian Williams will be glad to answer your questions about this study at any time. You can contact them on Andrew.vidgen@wales.nhs.uk or WilliamsCR11@cardiff.ac.uk

The final results of this study will be disseminated across all units that are involved. However, if you wish for your own copy, you can email either of the addresses given above and a copy of the study will be sent out to you.

Date: 13/04/2016
Version 11

Appendix 10: Consent form: Part 2, ward selection and observations

CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

CONSENT FORM

Assessing interpersonal interactions in inpatient mental health settings as a measure of the recovery orientation of the service.

Name of Researcher: Christian Williams.

Please initial boxes.

1. I confirm that I have read the information sheet dated 13.04.2016 (version 11) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected. ☐

3. I agree to take part in the above study. ☐

______________________	______________________	______________________
Name of Participant	Date	Signature
______________________	______________________	______________________
Name of Person taking consent	Date	Signature

Date: 13.04.2016
Version 11

Appendix 11: Demographic information sheet: Part 2, ward selection

Demographic Information Sheet

1. Which gender would you consider yourself?

Female

Male

Prefer not to say

2. How old are you?

_______years _______months

Prefer not to say

3. How long have you worked at this facility?

_______years _______months

Prefer not to say

4. What is your job title?

Prefer not to say

Version 1
Date 1/11/2015

Appendix 12: Recovery Self-Assessment, Provider Version (RSA-PV)

Code:_____

RSA-R
Provider Version

Please circle the number below which reflects how accurately the following statements describe the activities, values, policies, and practices of this program.

1	2	3	4	5
Strongly Disagree				Strongly Agree

N/A= Not Applicable
D/K= Don't Know

1. Staff make a concerted effort to welcome people in recovery and help them to feel comfortable in this program.	1	2	3	4	5	N/A	D/K
2. This program/agency offers an inviting and dignified physical environment (e.g., the lobby, waiting rooms, etc.).	1	2	3	4	5	N/A	D/K
3. Staff encourage program participants to have hope and high expectations for their recovery.	1	2	3	4	5	N/A	D/K
4. Program participants can change their clinician or case manager if they wish.	1	2	3	4	5	N/A	D/K
5. Program participants can easily access their treatment records if they wish.	1	2	3	4	5	N/A	D/K
6. Staff do not use threats, bribes, or other forms of pressure to influence the behavior of program participants.	1	2	3	4	5	N/A	D/K
7. Staff believe in the ability of program participants to recover.	1	2	3	4	5	N/A	D/K
8. Staff believe that program participants have the ability to manage their own symptoms.	1	2	3	4	5	N/A	D/K
9. Staff believe that program participants can make their own life choices regarding things such as where to live, when to work, whom to be friends with, etc.	1	2	3	4	5	N/A	D/K
10. Staff listen to and respect the decisions that program participants make about their treatment and care.	1	2	3	4	5	N/A	D/K
11. Staff regularly ask program participants about their interests and the things they would like to do in the community.	1	2	3	4	5	N/A	D/K
12. Staff encourage program participants to take risks and try new things.	1	2	3	4	5	N/A	D/K
13. This program offers specific services that fit each participant's unique culture and life experiences.	1	2	3	4	5	N/A	D/K
14. Staff offer participants opportunities to discuss their spiritual needs and interests when they wish.	1	2	3	4	5	N/A	D/K
15. Staff offer participants opportunities to discuss their sexual needs and interests when they wish.	1	2	3	4	5	N/A	D/K
16. Staff help program participants to develop and plan for life goals beyond	1	2	3	4	5	N/A	D/K

O'Connell, Tondora, Kidd, Stayner, Hawkins, and Davidson (2007)

Appendix 13: Observation matrix

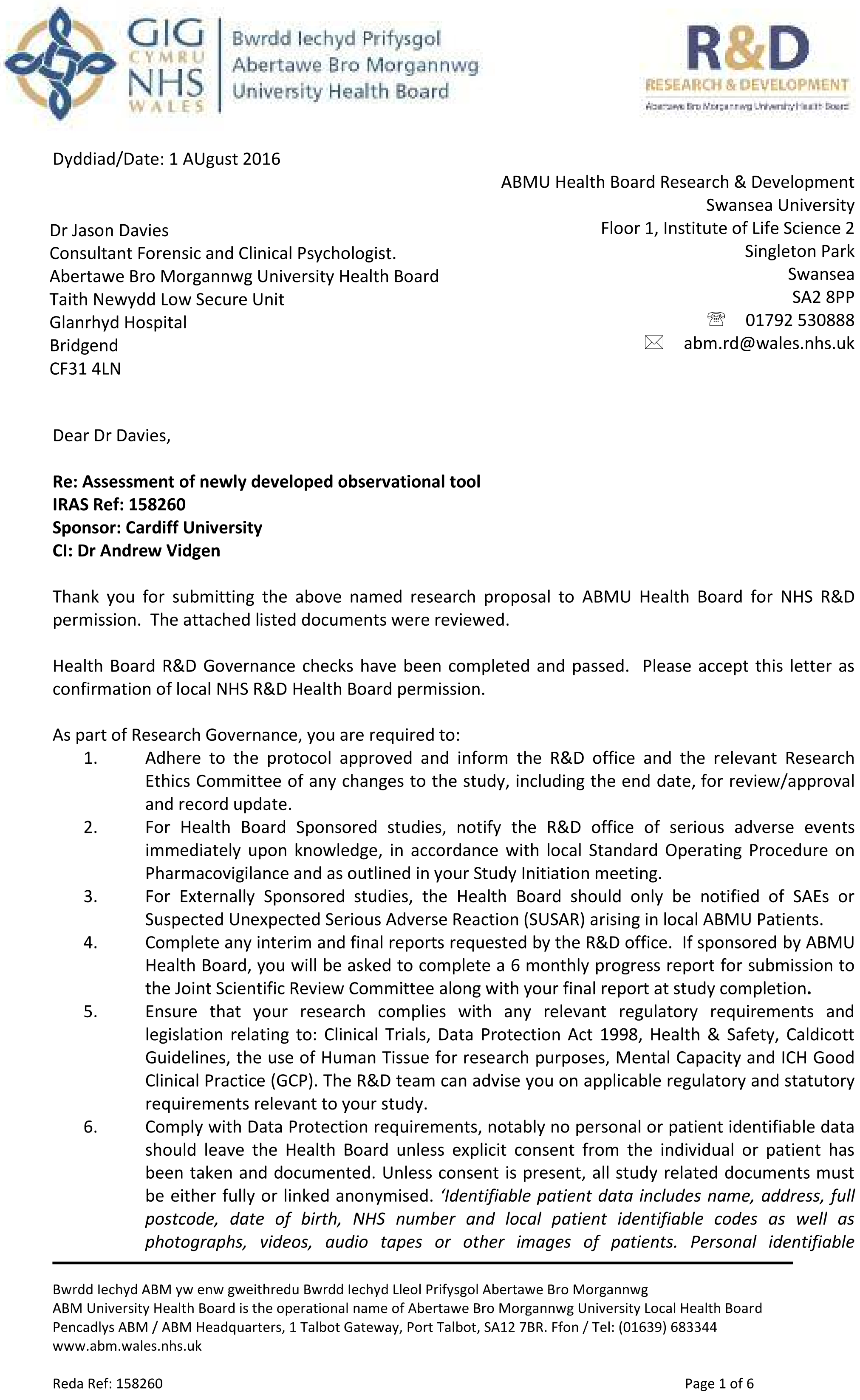
Session	
Description of interaction	Code

No patients present	No staff present	No one present

Appendix 14: Newly developed tool

Code
1. Encouraging engagement in activities.
3. Encouraging the development of independence.
6. Offering choice.
8. Staff initiated non-clinical interactions.
9. Positive responses to client initiated interactions.
11. Joint involvement in task/activities.
14. Seeking input from clients.
16. Active listening skills.
18. Recognition of communication needs.
20. Staff recognising strengths and interests.
26. Encouragement/practical support of pro-social activities, including off the unit.
28: Empathy shown in response to distress.
2. Forcing engagement in activities.
4. Non-verbal dismissal.
5. Discussing client's personal information (in public settings).
10. Negative responses to client initiated interactions.
12. Clients not being engaged with.
13. Interactions having an "illness focus".
15. Not seeking input from clients.
17. Staff not being in the moment with clients.
19. Ignoring communication needs.
24. Set getting up times.
30. Asking how a client would like to deal with a problem; exploring options.

Appendix 15: Letter of Access ABMu



GIG CYMRU NHS WALES
Bwrdd Iechyd Prifysgol
Abertawe Bro Morgannwg
University Health Board

R&D
RESEARCH & DEVELOPMENT
Abertawe Bro Morgannwg University Health Board

Dyddiad/Date: 1 AUgust 2016

ABMU Health Board Research & Development
Swansea University
Floor 1, Institute of Life Science 2
Singleton Park
Swansea
SA2 8PP
☎ 01792 530888
✉ abm.rd@wales.nhs.uk

Dr Jason Davies
Consultant Forensic and Clinical Psychologist.
Abertawe Bro Morgannwg University Health Board
Taith Newydd Low Secure Unit
Glanrhyd Hospital
Bridgend
CF31 4LN

Dear Dr Davies,

Re: Assessment of newly developed observational tool
IRAS Ref: 158260
Sponsor: Cardiff University
CI: Dr Andrew Vidgen

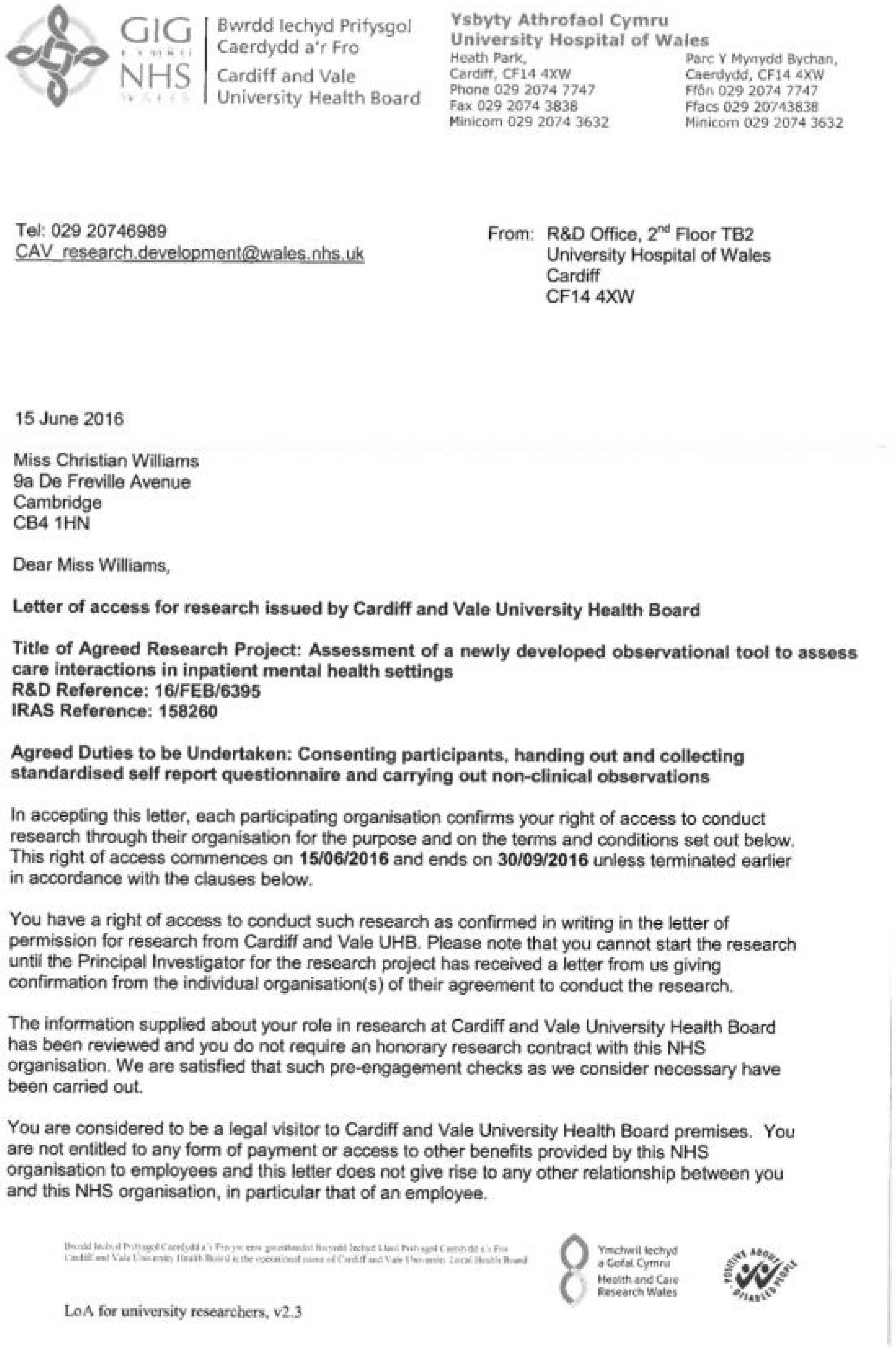
Thank you for submitting the above named research proposal to ABMU Health Board for NHS R&D permission. The attached listed documents were reviewed.

Health Board R&D Governance checks have been completed and passed. Please accept this letter as confirmation of local NHS R&D Health Board permission.

As part of Research Governance, you are required to:

1. Adhere to the protocol approved and inform the R&D office and the relevant Research Ethics Committee of any changes to the study, including the end date, for review/approval and record update.
2. For Health Board Sponsored studies, notify the R&D office of serious adverse events immediately upon knowledge, in accordance with local Standard Operating Procedure on Pharmacovigilance and as outlined in your Study Initiation meeting.
3. For Externally Sponsored studies, the Health Board should only be notified of SAEs or Suspected Unexpected Serious Adverse Reaction (SUSAR) arising in local ABMU Patients.
4. Complete any interim and final reports requested by the R&D office. If sponsored by ABMU Health Board, you will be asked to complete a 6 monthly progress report for submission to the Joint Scientific Review Committee along with your final report at study completion**.**
5. Ensure that your research complies with any relevant regulatory requirements and legislation relating to: Clinical Trials, Data Protection Act 1998, Health & Safety, Caldicott Guidelines, the use of Human Tissue for research purposes, Mental Capacity and ICH Good Clinical Practice (GCP). The R&D team can advise you on applicable regulatory and statutory requirements relevant to your study.
6. Comply with Data Protection requirements, notably no personal or patient identifiable data should leave the Health Board unless explicit consent from the individual or patient has been taken and documented. Unless consent is present, all study related documents must be either fully or linked anonymised. *'Identifiable patient data includes name, address, full postcode, date of birth, NHS number and local patient identifiable codes as well as photographs, videos, audio tapes or other images of patients. Personal identifiable*

Appendix 16: Letter of access Cardiff and Vale



GIG CYMRU
NHS WALES

Bwrdd Iechyd Prifysgol
Caerdydd a'r Fro
Cardiff and Vale
University Health Board

Ysbyty Athrofaol Cymru
University Hospital of Wales

Heath Park,
Cardiff, CF14 4XW
Phone 029 2074 7747
Fax 029 2074 3838
Minicom 029 2074 3632

Parc Y Mynydd Bychan,
Caerdydd, CF14 4XW
Ffôn 029 2074 7747
Ffacs 029 20743838
Minicom 029 2074 3632

Tel: 029 20746989
CAV_research.development@wales.nhs.uk

From: R&D Office, 2nd Floor TB2
University Hospital of Wales
Cardiff
CF14 4XW

15 June 2016

Miss Christian Williams
9a De Freville Avenue
Cambridge
CB4 1HN

Dear Miss Williams,

Letter of access for research issued by Cardiff and Vale University Health Board

Title of Agreed Research Project: Assessment of a newly developed observational tool to assess care interactions in inpatient mental health settings
R&D Reference: 16/FEB/6395
IRAS Reference: 158260

Agreed Duties to be Undertaken: Consenting participants, handing out and collecting standardised self report questionnaire and carrying out non-clinical observations

In accepting this letter, each participating organisation confirms your right of access to conduct research through their organisation for the purpose and on the terms and conditions set out below. This right of access commences on **15/06/2016** and ends on **30/09/2016** unless terminated earlier in accordance with the clauses below.

You have a right of access to conduct such research as confirmed in writing in the letter of permission for research from Cardiff and Vale UHB. Please note that you cannot start the research until the Principal Investigator for the research project has received a letter from us giving confirmation from the individual organisation(s) of their agreement to conduct the research.

The information supplied about your role in research at Cardiff and Vale University Health Board has been reviewed and you do not require an honorary research contract with this NHS organisation. We are satisfied that such pre-engagement checks as we consider necessary have been carried out.

You are considered to be a legal visitor to Cardiff and Vale University Health Board premises. You are not entitled to any form of payment or access to other benefits provided by this NHS organisation to employees and this letter does not give rise to any other relationship between you and this NHS organisation, in particular that of an employee.

Appendix 17: Letter of access Cwm Taf

GIG CYMRU NHS WALES
Bwrdd Iechyd Prifysgol Cwm Taf
University Health Board

Your Ref/ eich cyf
Our ref/ ein cyf: JG/TTW
Date/ dyddiad: 27/05/16
Tel/ ffôn: 01443 443421
Fax/ ffacs: 01443 443420
Email/ ebost: CTUHB_RD@wales.nhs.uk
Dept/ adran: Research & Development Dept

Miss Christian Williams
Trainee Clinical Psychologist
School of Psychology
Cardiff University
Tower Building
70 Park Place
Cardiff
CF10 3AT

Dear Miss Williams

Letter of access for research

This letter confirms your right of access to conduct research through Cwm Taf University Health Board for the purpose and on the terms and conditions set out below. This right of access commences on 27/05/16 and ends on 30/09/16 unless terminated earlier in accordance with the clauses below.

You have a right of access to conduct such research as confirmed in writing in the letter of permission for research from this NHS organisation. Please note that you cannot start the research until the Principal Investigator for the research project has received a letter from us giving permission to conduct the project.

The information supplied about your role in research at Cwm Taf University Health Board has been reviewed and you do not require an honorary research contract with this NHS organisation. We are satisfied that such pre-engagement checks as we consider necessary have been carried out.

You are considered to be a legal visitor to Cwm Taf University Health Board premises. You are not entitled to any form of payment or access to other benefits provided by this NHS organisation to employees and this letter does not give rise to any other relationship between you and this NHS organisation, in particular that of an employee.

While undertaking research through Cwm Taf University Health Board, you will remain accountable to your employer Cardiff University but you are required to follow the reasonable instructions of Mrs Lynne Garwood, Associate Director for Mental Health in this NHS organisation or those given on her/his behalf in relation to the terms of this right of access.

Where any third party claim is made, whether or not legal proceedings are issued, arising out of or in connection with your right of access, you are required to co-operate fully with any investigation by this NHS organisation in connection with any such claim and to give all such assistance as may reasonably be required regarding the conduct of any legal proceedings.

Return Address: Research & Development Department, Royal Glamorgan Hospital, Llantrisant, Rhondda Cynon Taff, CF72 8XR

Chair/Cadeirydd: Dr C D V Jones, CBE
Chief Executive/Prif Weithredydd: Mrs Allison Williams

Cwm Taf University Health Board is the operational name of Cwm Taf University Local Health Board/Bwrdd Iechyd Prifysgol Cwm Taf yw enw gweithredol Bwrdd Iechyd Lleol Cwm Taf

Appendix 18: Examples of Observations

Description of interaction	Code	P or N
Ward 31, Session 1		
1) SA to P1: SA ”you alright X?” P1 “yeah not bad.” SA “What are you thinking of doing today?”		
2) SH walks passed P2, P3, P4 & P5 who are sat watching TV. Looks over, no other interaction.		
3) SA to SB, (P1,2 & 4 present): SA ‘X just put his washing in, it needed to go in on high so will be done after I’m out, can you sort it?”		
4) P5 to SA: P5 “X, let me out for a quick ciggie”. SA checks watch “Ok butt, quick now.”		
5) P2 to SC as SC walks into the unit: P2 “Alright but?” SC “yeah man, nice hat, we won’t miss you with that will we?”		
6) SD administers P4 physical observations and meds in public living space.		
7) SE and P3: P3 “I’ve done it.” Hands SE filled in form. SE “Ah cool, want to put a picture on it?” I could take a really good one.” P3 smiles. SE “is that all I have to do to get a smile?” SE mimes taking a picture, P3 strikes a pose.		
8) SF walks passed P6. P6 “Alight?” SF nods, “Hiya.”		

Description of interaction	Code	P or N
Ward 31, Session 1		
9) SA to P1: SA ”you alright X?” P1 “yeah not bad.” SA “What are you thinking of doing today?”		
10) P3 to SA: P3” That’s a big water bottle, how much?” SA “2.5liters, I am trying to drink two of ‘em a day.” P3 “You must piss like a race horse.” SA “Yeah that is a down side.”		
11) P2 to SC as SC walks passed: P2 “what’s in there?” points at what is being carried. SC “Nothing but, it’s empty.” Lifts the lunch box and shows him.		
12) SB to P3: SC “Sit down now, and I’ll be along in a minute.” Walks into the office and closes the door.		
13) SD sits down next to P3. SD “I’ve got your money here, so that’s £100 in cash and your bank card, are you happy with that?” P3 “yeah.”		

Ward 31, Session 4		
11) SG to P5: SG “me, you and X, are we going out?” P5 “yeah. Hub? Game of pool?” SG “Yeah, when Y is back. Have you signed out?” P5 “No, got a pen?”		
14) SA enters unit, walked into side lounge where P6 and P7 are sat in silence. SA “Alright guys, you up to anything?’ P7 “Nah.” SA “Want to find something to do?”		
15) Sat in the TV lounge. SB to P3: SB “You should be wearing socks” said whilst examining P3’s feet.		
16) P2 stood by kitchen door. SG approaches him. SG “right, what we got then butt? What’s the plan?”		
17) SA wandering around the communal area where P3, P7 and P8 were sat in silence.		
Ward 31/session6/interaction20. P10 to SC: P10 “are you making tea?” SC “There’s milk in the kitchen, why don’t you make yourself a cup?”		
Ward12/session2/interaction18. SG approaches P7 in the corridor: SG “good trip? Where did you go?” P7 makes hand gesture, SG “church?....No” P7 makes another gesture SG “Shopping?” P7 nods.		
Ward 31/session5/interaction7: SD approaches P3: SD “I hear you’ve got cash on you you’re worried about?” P3 “yeah” SD “let me close this…” closes door, room becomes private.		
Ward 31/session4/interaction13: SG in kitchen doing the washing up. P6 enters: P6 “just getting some water.” SG does not look up SG “Ummm.”		
Ward 12/session2/interaction7: SF to P4: SF “you alright love?” P4 “yeah, you?” SF “Good thanks.”		
Ward 12/session6/interaction13: SE and P3 leaving the unit: SG “hope it stays dry for you!” SE to P3 SE ”we don’t mind a bit of rain do we?” P3 “no, we’ve got wet before.”		
Ward 31/session3/interaction11: P9 to SG: P9 “what’s for tea then?” SG looking down at paper in his hand “I dunno.: P9 “Hope its tasty I’m starving.” SG “Hmm” walks away.		
Ward 12/session2/interaction7: P1 to SE: P1 “Alright X?” SE “Hiya, coming to join me? Have you seen this?” Passes newspaper		
Ward 12/session6/interaction13: SB walks passed P8; SB “hello butt, ok?” P* hi, yeah.”		
Ward 31/session 5/interaction 4: SE to P4. SE “Do you want the telly on but?” P4 “No.”		
Ward 12/session 4/interaction 8: P1” Can I go out for a can of pop?” SA “What do you need to do before you can go for one?” SA Typing on computer throughout interaction, no eye contact.		
Ward 31/session 3/interaction 6: P1 to SE. P1 “oi, X” SE (in light hearted tone) “I can’t walk passed without you wanting something eh?” P1 “It’s cos I’m important see.” SA “Ai, you are I suppose!”		

Appendix 19: Number of codes rated by each member of the panel on the Relevance scale. Delphi method, Round 1

Number of participants (out of 34)	Number of codes rated (out of 26)
29	26
5	25
3	24
1	23
1	22

Appendix 20: Number of codes rated by each member of the panel on the Importance scale. Delphi method Round 1

Number of participants (out of 34)	Number of codes rated (out of 26)
23	26
5	25
1	24
2	23
1	22
1	21
1	20
1	10

Appendix 21: Scores for each code, which was scored 1 and 2, and 6 and 7 on the Relevance scale. Delphi method, Round 1.

Code Number	% of ratings which were 1 and 2	% of ratings which were 6 and 7	Conclusion
Code 1	0	100	Included
Code 2	33.3	15.2	Re-rated
Code 3	2.9	94.1	Included
Code 4	37.6	12.5	Re-rated
Code 5	60.6	24.2	Re-rated
Code 6	2.9	85.3	Included
Code 7	25.9	25.8	Re-rated
Code 8	0	81.9	Included
Code 9	6.1	84.8	Included
Code 10	42.4	18.2	Re-rated
Code 11	3	78.8	Included
Code 12	44.1	26.4	Re-rated
Code 13	21.9	31.3	Re-rated
Code 14	2.9	76.4	Included
Code 15	57.6	15.2	Re-rated
Code 16	0	91.2	Included
Code 17	26.5	47.1	Re-rated
Code 18	3	87.9	Included
Code 19	54.6	30.3	Re-rated
Code 20	2.9	91.2	Included
Code 21	53.1	25	Re-rated
Code 22	11.7	41.1	Re-rated
Code 23	14.7	41.1	Re-rated
Code 24	11.7	50	Re-rated
Code 25	35.3	38.3	Re-rated
Code 26	0	100	Included

Appendix 22: Scores for each code, scored 1 and 2, and 6 and 7 on Importance. Delphi method, Round 1

Code Number	% of ratings which were 1 and 2	% of ratings which were 6 and 7	Conclusion
Code 1	0	100	Included
Code 2	32.3	25.9	Re-rate
Code 3	0	94.1	Included
Code 4	31	31	Re-rate
Code 5	42	48.4	Re-rate
Code 6	6	93.9	Included
Code 7	26.7	33.3	Re-rate
Code 8	0	78.8	Included
Code 9	2.9	88.2	Included
Code 10	33.3	33.3	Re-rate
Code 11	33.3	33.3	Re-rate
Code 12	31.3	43.8	Re-rate
Code 13	18.2	36.4	Re-rate
Code 14	0	82.3	Included
Code 15	43.8	31.2	Re-rate
Code 16	0	97.1	Included
Code 17	27.3	39.5	Re-rate
Code 18	0	93.5	Included
Code 19	45.1	45.2	Re-rate
Code 20	0	100	Included
Code 21	46.9	28.1	Re-rate
Code 22	8.8	38.3	Re-rate
Code 23	6	54.5	Re-rate
Code 24	8.8	38.3	Re-rate
Code 25	35.3	35.3	Re-rate
Code 26	0	100	Included

Appendix 23: SPSS Output for Relevance and Importance ratings, Delphi method. Round 1.

Relevance ratings.

Code 1: Encouraging engagement in activities

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	6.0	2	5.9	6.1	6.1
	7.0	31	91.2	93.9	100.0
	Total	33	97.1	100.0	
Missing	System	1	2.9		
Total		34	100.0		

Code 2: Forcing engagement in activities

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.0	8	23.5	24.2	24.2
	2.0	3	8.8	9.1	33.3
	3.0	2	5.9	6.1	39.4
	4.0	9	26.5	27.3	66.7
	5.0	6	17.6	18.2	84.8
	6.0	2	5.9	6.1	90.9
	7.0	3	8.8	9.1	100.0
	Total	33	97.1	100.0	
Missing	System	1	2.9		
Total		34	100.0		

Code 3: Encouraging the development of independence

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.0	1	2.9	2.9	2.9
	5.0	1	2.9	2.9	5.9
	6.0	1	2.9	2.9	8.8
	7.0	31	91.2	91.2	100.0
	Total	34	100.0	100.0	

Code 4: Non-verbal dismissal

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.0	10	29.4	31.3	31.3
	2.0	2	5.9	6.3	37.5

	3.0	1	2.9	3.1	40.6
	4.0	12	35.3	37.5	78.1
	5.0	3	8.8	9.4	87.5
	6.0	3	8.8	9.4	96.9
	7.0	1	2.9	3.1	100.0
	Total	32	94.1	100.0	
Missing	System	2	5.9		
Total		34	100.0		

Code 5: Discussing client's personal information (in public settings)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.0	18	52.9	54.5	54.5
	2.0	2	5.9	6.1	60.6
	3.0	1	2.9	3.0	63.6
	4.0	1	2.9	3.0	66.7
	5.0	3	8.8	9.1	75.8
	6.0	1	2.9	3.0	78.8
	7.0	7	20.6	21.2	100.0
	Total	33	97.1	100.0	
Missing	System	1	2.9		
Total		34	100.0		

Code 6: Offering choice.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.0	1	2.9	2.9	2.9
	3.0	1	2.9	2.9	5.9
	4.0	2	5.9	5.9	11.8
	5.0	1	2.9	2.9	14.7
	6.0	4	11.8	11.8	26.5
	7.0	25	73.5	73.5	100.0
	Total	34	100.0	100.0	

Code 7: Tokenistic involvement

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.0	6	17.6	19.4	19.4
	2.0	2	5.9	6.5	25.8
	3.0	8	23.5	25.8	51.6

	4.0	2	5.9	6.5	58.1
	5.0	5	14.7	16.1	74.2
	6.0	4	11.8	12.9	87.1
	7.0	4	11.8	12.9	100.0
	Total	31	91.2	100.0	
Missing	System	3	8.8		
Total		34	100.0		

Code 8: Staff initiated non-clinical interactions

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	4.0	2	5.9	6.1	6.1
	5.0	4	11.8	12.1	18.2
	6.0	5	14.7	15.2	33.3
	7.0	22	64.7	66.7	100.0
	Total	33	97.1	100.0	
Missing	System	1	2.9		
Total		34	100.0		

Code 9: Positive responses to client initiated interactions

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.0	2	5.9	6.1	6.1
	5.0	3	8.8	9.1	15.2
	6.0	10	29.4	30.3	45.5
	7.0	18	52.9	54.5	100.0
	Total	33	97.1	100.0	
Missing	System	1	2.9		
Total		34	100.0		

Code 10: Negative responses to client initiated interactions

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.0	11	32.4	33.3	33.3
	2.0	3	8.8	9.1	42.4
	3.0	5	14.7	15.2	57.6
	4.0	6	17.6	18.2	75.8
	5.0	2	5.9	6.1	81.8

	6.0	4	11.8	12.1	93.9
	7.0	2	5.9	6.1	100.0
	Total	33	97.1	100.0	
Missing	System	1	2.9		
Total		34	100.0		

Code 11: Joint involvement in task/activities

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.0	1	2.9	3.0	3.0
	4.0	3	8.8	9.1	12.1
	5.0	3	8.8	9.1	21.2
	6.0	6	17.6	18.2	39.4
	7.0	20	58.8	60.6	100.0
	Total	33	97.1	100.0	
Missing	System	1	2.9		
Total		34	100.0		

Code 12: Clients not being engaged with

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.0	10	29.4	29.4	29.4
	2.0	5	14.7	14.7	44.1
	3.0	1	2.9	2.9	47.1
	4.0	4	11.8	11.8	58.8
	5.0	5	14.7	14.7	73.5
	6.0	3	8.8	8.8	82.4
	7.0	6	17.6	17.6	100.0
	Total	34	100.0	100.0	

Code 13: Interactions having an ‘illness focus’

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.0	5	14.7	15.6	15.6
	2.0	2	5.9	6.3	21.9
	3.0	1	2.9	3.1	25.0
	4.0	7	20.6	21.9	46.9
	5.0	7	20.6	21.9	68.8
	6.0	3	8.8	9.4	78.1

	7.0	7	20.6	21.9	100.0
	Total	32	94.1	100.0	
Missing	System	2	5.9		
Total		34	100.0		

Code 14: Seeking input from clients

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.0	1	2.9	2.9	2.9
	4.0	1	2.9	2.9	5.9
	5.0	6	17.6	17.6	23.5
	6.0	3	8.8	8.8	32.4
	7.0	23	67.6	67.6	100.0
	Total	34	100.0	100.0	

Code 15: Not seeking input from clients

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.0	14	41.2	42.4	42.4
	2.0	5	14.7	15.2	57.6
	3.0	1	2.9	3.0	60.6
	4.0	5	14.7	15.2	75.8
	5.0	3	8.8	9.1	84.8
	6.0	2	5.9	6.1	90.9
	7.0	3	8.8	9.1	100.0
	Total	33	97.1	100.0	
Missing	System	1	2.9		
Total		34	100.0		

Code 16: Active listening skills

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	3.0	1	2.9	2.9	2.9
	4.0	1	2.9	2.9	5.9
	5.0	1	2.9	2.9	8.8
	6.0	2	5.9	5.9	14.7
	7.0	29	85.3	85.3	100.0
	Total	34	100.0	100.0	

Code 17: Staff not being in the moment with clients

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.0	7	20.6	20.6	20.6
	2.0	2	5.9	5.9	26.5
	3.0	5	14.7	14.7	41.2
	4.0	4	11.8	11.8	52.9
	5.0	7	20.6	20.6	73.5
	7.0	9	26.5	26.5	100.0
	Total	34	100.0	100.0	

Code 18: Recognition of communication needs

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.0	1	2.9	3.0	3.0
	4.0	2	5.9	6.1	9.1
	5.0	1	2.9	3.0	12.1
	6.0	3	8.8	9.1	21.2
	7.0	26	76.5	78.8	100.0
	Total	33	97.1	100.0	
Missing	System	1	2.9		
Total		34	100.0		

Code 19: Ignoring communication needs

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.0	15	44.1	45.5	45.5
	2.0	3	8.8	9.1	54.5
	4.0	2	5.9	6.1	60.6
	5.0	3	8.8	9.1	69.7
	6.0	1	2.9	3.0	72.7
	7.0	9	26.5	27.3	100.0
	Total	33	97.1	100.0	
Missing	System	1	2.9		
Total		34	100.0		

Code 20: Staff recognising strengths and interests

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2.0	1	2.9	2.9	2.9
	5.0	2	5.9	5.9	8.8
	6.0	2	5.9	5.9	14.7
	7.0	29	85.3	85.3	100.0
	Total	34	100.0	100.0	

Code 21: Punitive interactions

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.0	13	38.2	40.6	40.6
	2.0	4	11.8	12.5	53.1
	3.0	4	11.8	12.5	65.6
	5.0	3	8.8	9.4	75.0
	6.0	5	14.7	15.6	90.6
	7.0	3	8.8	9.4	100.0
	Total	32	94.1	100.0	
Missing	System	2	5.9		
Total		34	100.0		

Code 22: Set meal times

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.0	3	8.8	8.8	8.8
	2.0	1	2.9	2.9	11.8
	3.0	2	5.9	5.9	17.6
	4.0	9	26.5	26.5	44.1
	5.0	5	14.7	14.7	58.8
	6.0	6	17.6	17.6	76.5
	7.0	8	23.5	23.5	100.0
	Total	34	100.0	100.0	

Code 23: Set bed times

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.0	3	8.8	8.8	8.8

	2.0	2	5.9	5.9	14.7
	3.0	2	5.9	5.9	20.6
	4.0	3	8.8	8.8	29.4
	5.0	10	29.4	29.4	58.8
	6.0	8	23.5	23.5	82.4
	7.0	6	17.6	17.6	100.0
	Total	34	100.0	100.0	

Code 24: Set getting up times

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.0	3	8.8	8.8	8.8
	2.0	1	2.9	2.9	11.8
	3.0	1	2.9	2.9	14.7
	4.0	5	14.7	14.7	29.4
	5.0	7	20.6	20.6	50.0
	6.0	10	29.4	29.4	79.4
	7.0	7	20.6	20.6	100.0
	Total	34	100.0	100.0	

Code 25: Set smoking times

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.0	11	32.4	32.4	32.4
	2.0	1	2.9	2.9	35.3
	3.0	3	8.8	8.8	44.1
	4.0	4	11.8	11.8	55.9
	5.0	2	5.9	5.9	61.8
	6.0	4	11.8	11.8	73.5
	7.0	9	26.5	26.5	100.0
	Total	34	100.0	100.0	

Code 26: Encouragement/practical support of pro-social activities, including off the unit

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	6.0	2	5.9	5.9	5.9
	7.0	32	94.1	94.1	100.0
	Total	34	100.0	100.0	

Importance ratings.

Code 1: Encouraging engagement in activities

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	6.00	3	8.8	9.1	9.1
	7.00	30	88.2	90.9	100.0
	Total	33	97.1	100.0	
Missing	System	1	2.9		
Total		34	100.0		

Code 2: Forcing engagement in activities

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	8	23.5	25.8	25.8
	2.00	2	5.9	6.5	32.3
	4.00	8	23.5	25.8	58.1
	5.00	5	14.7	16.1	74.2
	6.00	6	17.6	19.4	93.5
	7.00	2	5.9	6.5	100.0
	Total	31	91.2	100.0	
Missing	System	3	8.8		
Total		34	100.0		

Code 3: Encouraging the development of independence

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	4.00	1	2.9	2.9	2.9
	5.00	1	2.9	2.9	5.9
	6.00	1	2.9	2.9	8.8
	7.00	31	91.2	91.2	100.0
	Total	34	100.0	100.0	

Code 4: Non-verbal dismissal

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	7	20.6	24.1	24.1
	2.00	2	5.9	6.9	31.0
	4.00	6	17.6	20.7	51.7

	5.00	5	14.7	17.2	69.0
	6.00	3	8.8	10.3	79.3
	7.00	6	17.6	20.7	100.0
	Total	29	85.3	100.0	
Missing	System	5	14.7		
Total		34	100.0		

Code 5: Discussing client's personal information (in public settings)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	11	32.4	35.5	35.5
	2.00	2	5.9	6.5	41.9
	5.00	3	8.8	9.7	51.6
	6.00	2	5.9	6.5	58.1
	7.00	13	38.2	41.9	100.0
	Total	31	91.2	100.0	
Missing	System	3	8.8		
Total		34	100.0		

Code 6: Offering choice

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	3.00	1	2.9	3.0	3.0
	4.00	1	2.9	3.0	6.1
	6.00	3	8.8	9.1	15.2
	7.00	28	82.4	84.8	100.0
	Total	33	97.1	100.0	
Missing	System	1	2.9		
Total		34	100.0		

Code 7: Tokenistic involvement

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	5	14.7	16.7	16.7
	2.00	3	8.8	10.0	26.7
	3.00	4	11.8	13.3	40.0
	4.00	2	5.9	6.7	46.7
	5.00	6	17.6	20.0	66.7
	6.00	6	17.6	20.0	86.7
	7.00	4	11.8	13.3	100.0
	Total	30	88.2	100.0	

Missing	System	4	11.8		
Total		34	100.0		

Code 8: Staff initiated non-clinical interactions

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	4.00	2	5.9	6.1	6.1
	5.00	5	14.7	15.2	21.2
	6.00	7	20.6	21.2	42.4
	7.00	19	55.9	57.6	100.0
	Total	33	97.1	100.0	
Missing	System	1	2.9		
Total		34	100.0		

Code 9: Positive responses to client initiated interactions

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	1	2.9	2.9	2.9
	5.00	3	8.8	8.8	11.8
	6.00	6	17.6	17.6	29.4
	7.00	24	70.6	70.6	100.0
	Total	34	100.0	100.0	

Code 10: Negative responses to client initiated interactions

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	7	20.6	25.9	25.9
	2.00	2	5.9	7.4	33.3
	3.00	2	5.9	7.4	40.7
	4.00	3	8.8	11.1	51.9
	5.00	4	11.8	14.8	66.7
	6.00	1	2.9	3.7	70.4
	7.00	8	23.5	29.6	100.0
	Total	27	79.4	100.0	
Missing	System	7	20.6		
Total		34	100.0		

Code 11: Joint involvement in task/activities

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	7	20.6	25.9	25.9
	2.00	2	5.9	7.4	33.3
	3.00	2	5.9	7.4	40.7
	4.00	3	8.8	11.1	51.9
	5.00	4	11.8	14.8	66.7
	6.00	1	2.9	3.7	70.4
	7.00	8	23.5	29.6	100.0
	Total	27	79.4	100.0	
Missing	System	7	20.6		
Total		34	100.0		

Code 12: Clients not being engaged with

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	8	23.5	25.0	25.0
	2.00	2	5.9	6.3	31.3
	4.00	4	11.8	12.5	43.8
	5.00	4	11.8	12.5	56.3
	6.00	4	11.8	12.5	68.8
	7.00	10	29.4	31.3	100.0
	Total	32	94.1	100.0	
Missing	System	2	5.9		
Total		34	100.0		

Code 13: Interactions having an ‘illness focus’

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	3	8.8	9.1	9.1
	2.00	3	8.8	9.1	18.2
	3.00	1	2.9	3.0	21.2
	4.00	6	17.6	18.2	39.4
	5.00	8	23.5	24.2	63.6
	6.00	2	5.9	6.1	69.7
	7.00	10	29.4	30.3	100.0
	Total	33	97.1	100.0	
Missing	System	1	2.9		
Total		34	100.0		

Code 14: Seeking input from clients

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	4.00	1	2.9	2.9	2.9
	5.00	5	14.7	14.7	17.6
	6.00	1	2.9	2.9	20.6
	7.00	27	79.4	79.4	100.0
	Total	34	100.0	100.0	

Code 15: Not seeking input from clients

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	11	32.4	34.4	34.4
	2.00	3	8.8	9.4	43.8
	3.00	2	5.9	6.3	50.0
	4.00	2	5.9	6.3	56.3
	5.00	4	11.8	12.5	68.8
	6.00	1	2.9	3.1	71.9
	7.00	9	26.5	28.1	100.0
	Total	32	94.1	100.0	
Missing	System	2	5.9		
Total		34	100.0		

Code 16: Active listening skills

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	4.00	1	2.9	2.9	2.9
	6.00	2	5.9	5.9	8.8
	7.00	31	91.2	91.2	100.0
	Total	34	100.0	100.0	

Code 17: Staff not being in the moment with clients

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	5	14.7	15.2	15.2
	2.00	4	11.8	12.1	27.3

	3.00	3	8.8	9.1	36.4
	4.00	4	11.8	12.1	48.5
	5.00	3	8.8	9.1	57.6
	6.00	2	5.9	6.1	63.6
	7.00	12	35.3	36.4	100.0
	Total	33	97.1	100.0	
Missing	System	1	2.9		
Total		34	100.0		

Code 18: Recognition of communication needs

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	4.00	1	2.9	3.1	3.1
	5.00	1	2.9	3.1	6.3
	6.00	3	8.8	9.4	15.6
	7.00	27	79.4	84.4	100.0
	Total	32	94.1	100.0	
Missing	System	2	5.9		
Total		34	100.0		

Code 19: Ignoring communication needs

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	13	38.2	41.9	41.9
	2.00	1	2.9	3.2	45.2
	4.00	2	5.9	6.5	51.6
	5.00	1	2.9	3.2	54.8
	6.00	2	5.9	6.5	61.3
	7.00	12	35.3	38.7	100.0
	Total	31	91.2	100.0	
Missing	System	3	8.8		
Total		34	100.0		

Code 20: Staff recognising strengths and interests

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	6.00	3	8.8	9.1	9.1
	7.00	30	88.2	90.9	100.0
	Total	33	97.1	100.0	

Missing System	1	2.9		
Total	34	100.0		

Code 21: Punitive interactions

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	13	38.2	40.6	40.6
	2.00	2	5.9	6.3	46.9
	3.00	1	2.9	3.1	50.0
	4.00	2	5.9	6.3	56.3
	5.00	5	14.7	15.6	71.9
	6.00	4	11.8	12.5	84.4
	7.00	5	14.7	15.6	100.0
	Total	32	94.1	100.0	
Missing	System	2	5.9		
Total		34	100.0		

Code 22: Set meal times

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	2	5.9	5.9	5.9
	2.00	1	2.9	2.9	8.8
	3.00	2	5.9	5.9	14.7
	4.00	11	32.4	32.4	47.1
	5.00	5	14.7	14.7	61.8
	6.00	4	11.8	11.8	73.5
	7.00	9	26.5	26.5	100.0
	Total	34	100.0	100.0	

Code 23: Set bed times

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	1	2.9	3.0	3.0
	2.00	1	2.9	3.0	6.1
	3.00	2	5.9	6.1	12.1
	4.00	6	17.6	18.2	30.3
	5.00	5	14.7	15.2	45.5
	6.00	11	32.4	33.3	78.8
	7.00	7	20.6	21.2	100.0
	Total	33	97.1	100.0	
Missing	System	1	2.9		
Total		34	100.0		

Code 24: Set getting up times

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	1	2.9	2.9	2.9
	2.00	2	5.9	5.9	8.8
	3.00	1	2.9	2.9	11.8
	4.00	6	17.6	17.6	29.4
	5.00	8	23.5	23.5	52.9
	6.00	8	23.5	23.5	76.5
	7.00	8	23.5	23.5	100.0
	Total	34	100.0	100.0	

Code 25: Set smoking times

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	10	29.4	29.4	29.4
	2.00	2	5.9	5.9	35.3
	3.00	4	11.8	11.8	47.1
	4.00	4	11.8	11.8	58.8
	5.00	2	5.9	5.9	64.7
	6.00	4	11.8	11.8	76.5
	7.00	8	23.5	23.5	100.0
	Total	34	100.0	100.0	

Code 26: Encouragement/practical support of pro-social activities, including off the unit

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	7.00	33	97.1	100.0	100.0
Missing	System	1	2.9		
Total		34	100.0		

Appendix 24: SPSS Output, Delphi. Round 2.

Relevance Ratings

Code 2 forcing engagement in activates

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	3	27.3	27.3	27.3
	2	1	9.1	9.1	36.4
	4	1	9.1	9.1	45.5
	5	3	27.3	27.3	72.7
	6	3	27.3	27.3	100.0
	Total	11	100.0	100.0	

Code 4: non-verbal dismissal

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	1	9.1	9.1	9.1
	3	3	27.3	27.3	36.4
	4	1	9.1	9.1	45.5
	5	3	27.3	27.3	72.7
	6	2	18.2	18.2	90.9
	7	1	9.1	9.1	100.0
	Total	11	100.0	100.0	

Code5: Discussing client's personal information in public settings

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	4	36.4	36.4	36.4
	2	2	18.2	18.2	54.5
	5	2	18.2	18.2	72.7
	7	3	27.3	27.3	100.0
	Total	11	100.0	100.0	

Code 7: Tokenistic involvement

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	1	9.1	9.1	9.1
	2	1	9.1	9.1	18.2
	3	1	9.1	9.1	27.3
	4	2	18.2	18.2	45.5
	5	2	18.2	18.2	63.6
	6	3	27.3	27.3	90.9
	7	1	9.1	9.1	100.0

	Total	11	100.0	100.0	

Code 10: Negative responses to client initiated interactions

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	1	9.1	9.1	9.1
	2	2	18.2	18.2	27.3
	3	3	27.3	27.3	54.5
	4	2	18.2	18.2	72.7
	5	1	9.1	9.1	81.8
	7	2	18.2	18.2	100.0
	Total	11	100.0	100.0	

Code 12: Clients not being engaged with

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	1	9.1	9.1	9.1
	3	2	18.2	18.2	27.3
	4	4	36.4	36.4	63.6
	5	1	9.1	9.1	72.7
	6	1	9.1	9.1	81.8
	7	2	18.2	18.2	100.0
	Total	11	100.0	100.0	

Code 13: Interactions having an ‘illness focus’

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	1	9.1	9.1	9.1
	2	2	18.2	18.2	27.3
	4	3	27.3	27.3	54.5
	5	2	18.2	18.2	72.7
	6	3	27.3	27.3	100.0
	Total	11	100.0	100.0	

Code 15: Not seeking input from clients

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	2	18.2	18.2	18.2

	2	1	9.1	9.1	27.3
	3	1	9.1	9.1	36.4
	4	2	18.2	18.2	54.5
	5	2	18.2	18.2	72.7
	6	2	18.2	18.2	90.9
	7	1	9.1	9.1	100.0
	Total	11	100.0	100.0	

Code 17: Staff not being in the moment with clients

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	1	9.1	9.1	9.1
	2	1	9.1	9.1	18.2
	3	1	9.1	9.1	27.3
	4	2	18.2	18.2	45.5
	5	2	18.2	18.2	63.6
	6	2	18.2	18.2	81.8
	7	2	18.2	18.2	100.0
	Total	11	100.0	100.0	

Code 19: Ignoring communication needs

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	3	27.3	27.3	27.3
	3	2	18.2	18.2	45.5
	4	2	18.2	18.2	63.6
	6	1	9.1	9.1	72.7
	7	3	27.3	27.3	100.0
	Total	11	100.0	100.0	

Code 21: Punitive interactions

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	2	18.2	18.2	18.2
	2	3	27.3	27.3	45.5
	4	2	18.2	18.2	63.6
	7	4	36.4	36.4	100.0
	Total	11	100.0	100.0	

Code 22: Set meal times

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	4	3	27.3	27.3	27.3
	5	2	18.2	18.2	45.5
	6	2	18.2	18.2	63.6
	7	4	36.4	36.4	100.0
	Total	11	100.0	100.0	

Code 23: Set bed times

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	1	9.1	9.1	9.1
	3	3	27.3	27.3	36.4
	4	2	18.2	18.2	54.5
	5	2	18.2	18.2	72.7
	6	3	27.3	27.3	100.0
	Total	11	100.0	100.0	

Code 24: Set getting up times

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	3	2	18.2	18.2	18.2
	4	1	9.1	9.1	27.3
	5	6	54.5	54.5	81.8
	6	2	18.2	18.2	100.0
	Total	11	100.0	100.0	

Code 25: Set smoking times

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	1	9.1	9.1	9.1
	2	3	27.3	27.3	36.4
	3	3	27.3	27.3	63.6
	5	1	9.1	9.1	72.7
	6	2	18.2	18.2	90.9
	7	1	9.1	9.1	100.0
	Total	11	100.0	100.0	

Code 27: PRN medication offered as a quick fix/alternative to talking

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	4	36.4	36.4	36.4
	3	3	27.3	27.3	63.6
	4	1	9.1	9.1	72.7
	5	2	18.2	18.2	90.9
	6	1	9.1	9.1	100.0
	Total	11	100.0	100.0	

Code 28: Empathy shown in response to distress

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	4	1	9.1	9.1	9.1
	5	3	27.3	27.3	36.4
	6	3	27.3	27.3	63.6
	7	4	36.4	36.4	100.0
	Total	11	100.0	100.0	

Code 29: Doing with, not doing for

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	3	3	27.3	27.3	27.3
	4	3	27.3	27.3	54.5
	5	3	27.3	27.3	81.8
	6	2	18.2	18.2	100.0
	Total	11	100.0	100.0	

Code 30: Asking how client would like to deal with a problem, exploring options

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	3	2	18.2	18.2	18.2
	4	3	27.3	27.3	45.5
	5	2	18.2	18.2	63.6
	6	2	18.2	18.2	81.8
	7	2	18.2	18.2	100.0
	Total	11	100.0	100.0	

Importance Ratings

Code 2 Forcing engagement in activates

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	1	9.1	9.1	9.1

	2	1	9.1	9.1	18.2
	5	1	9.1	9.1	27.3
	6	2	18.2	18.2	45.5
	7	6	54.5	54.5	100.0
	Total	11	100.0	100.0	

Code 4: Non-verbal dismissal

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	1	9.1	9.1	9.1
	4	1	9.1	9.1	18.2
	5	1	9.1	9.1	27.3
	6	1	9.1	9.1	36.4
	7	7	63.6	63.6	100.0
	Total	11	100.0	100.0	

Code5: Discussing client’s personal information in public settings

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	2	18.2	18.2	18.2
	7	9	81.8	81.8	100.0
	Total	11	100.0	100.0	

Code 7: Tokenistic involvement

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	1	9.1	9.1	9.1
	3	1	9.1	9.1	18.2
	4	2	18.2	18.2	36.4
	6	2	18.2	18.2	54.5
	7	5	45.5	45.5	100.0
	Total	11	100.0	100.0	

Code 10: Negative responses to client initiated interactions

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	2	18.2	18.2	18.2
	6	2	18.2	18.2	36.4
	7	7	63.6	63.6	100.0

Total	11	100.0	100.0	

Code 12: Clients not being engaged with

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 4	1	9.1	9.1	9.1
5	1	9.1	9.1	18.2
6	1	9.1	9.1	27.3
7	8	72.7	72.7	100.0
Total	11	100.0	100.0	

Code 13: Interactions having an 'illness focus'

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 4	2	18.2	18.2	18.2
5	1	9.1	9.1	27.3
6	1	9.1	9.1	36.4
7	7	63.6	63.6	100.0
Total	11	100.0	100.0	

Code 15: Not seeking input from clients

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 4	1	9.1	9.1	9.1
5	1	9.1	9.1	18.2
6	2	18.2	18.2	36.4
7	7	63.6	63.6	100.0
Total	11	100.0	100.0	

Code 17: Staff not being in the moment with clients

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 4	2	18.2	18.2	18.2
5	1	9.1	9.1	27.3
6	3	27.3	27.3	54.5
7	5	45.5	45.5	100.0
Total	11	100.0	100.0	

Code 19: Ignoring communication needs

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	3	1	9.1	9.1	9.1
	4	1	9.1	9.1	18.2
	6	1	9.1	9.1	27.3
	7	8	72.7	72.7	100.0
	Total	11	100.0	100.0	

Code 21: Punitive interactions

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	2	18.2	18.2	18.2
	2	1	9.1	9.1	27.3
	6	1	9.1	9.1	36.4
	7	7	63.6	63.6	100.0
	Total	11	100.0	100.0	

Code 22: Set meal times

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	4	5	45.5	45.5	45.5
	5	1	9.1	9.1	54.5
	6	3	27.3	27.3	81.8
	7	2	18.2	18.2	100.0
	Total	11	100.0	100.0	

Code 23: Set bed times

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	3	2	18.2	18.2	18.2
	4	2	18.2	18.2	36.4
	5	2	18.2	18.2	54.5
	6	4	36.4	36.4	90.9
	7	1	9.1	9.1	100.0
	Total	11	100.0	100.0	

Code 24: Set getting up times

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	4	3	27.3	27.3	27.3
	5	1	9.1	9.1	36.4
	6	4	36.4	36.4	72.7
	7	3	27.3	27.3	100.0
	Total	11	100.0	100.0	

Code 25: Set smoking times

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	1	9.1	9.1	9.1
	3	4	36.4	36.4	45.5
	4	2	18.2	18.2	63.6
	5	1	9.1	9.1	72.7
	6	2	18.2	18.2	90.9
	7	1	9.1	9.1	100.0
	Total	11	100.0	100.0	

Code 27: PRN medication offered as a quick fix/alternative to talking

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	3	2	18.2	18.2	18.2
	4	4	36.4	36.4	54.5
	5	2	18.2	18.2	72.7
	6	3	27.3	27.3	100.0
	Total	11	100.0	100.0	

Code 28: Empathy shown in response to distress

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	4	3	27.3	27.3	27.3
	5	3	27.3	27.3	54.5
	6	3	27.3	27.3	81.8
	7	2	18.2	18.2	100.0
	Total	11	100.0	100.0	

Code 29: Doing with not doing fo.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	3	4	36.4	36.4	36.4
	4	4	36.4	36.4	72.7
	5	3	27.3	27.3	100.0
	Total	11	100.0	100.0	

Code 30: Asking how a client would like to deal with a problem; exploring options

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	4	1	9.1	9.1	9.1
	5	3	27.3	27.3	36.4
	6	3	27.3	27.3	63.6
	7	4	36.4	36.4	100.0
	Total	11	100.0	100.0	

Appendix 25: Percentage of participants who rated each code 5,6 or 7, and whether codes were included or excluded based on the Relevance score Delphi method, Round 2

Code	% of ratings which were 5, 6 or 7	Conclusion
Code 2	54.6	Excluded
Code 4	64.6	Excluded
Code 5	45.5	Excluded
Code 7	64.6	Excluded
Code 10	27.3	Excluded
Code 12	36.4	Excluded
Code 13	45.5	Excluded
Code 15	45.5	Excluded
Code 17	54.6	Excluded
Code 19	36.4	Excluded
Code 21	36.4	Excluded
Code 22	72.8	Excluded
Code 23	45.5	Excluded
Code 24	72.7	Excluded
Code 25	36.4	Excluded
Code 27	27.3	Excluded
Code 28	91	Included
Code 29	25.5	Excluded
Code 30	56.4	Excluded

Appendix 26: Percentage of participants who rated each code 5,6 or 7, and whether codes were included or excluded based on the Importance score Delphi method, Round 2

Code	% of ratings which were 5, 6 or 7	Conclusion
Code 2	81.8	Included
Code 4	81.8	Included
Code 5	81.8	Included
Code 7	63.7	Excluded
Code 10	81.8	Included
Code 12	81.8	Included
Code 13	81.8	Included
Code 15	90.9	Included
Code 17	81.9	Included
Code 19	81.8	Included
Code 21	72.7	Excluded
Code 22	54.6	Excluded
Code 23	63.7	Excluded
Code 24	72.8	Included
Code 25	36.4	Excluded
Code 27	45.5	Excluded
Code 28	72.8	Included
Code 29	27.3	Excluded
Code 30	91	Included

Appendix 27: Codes included in the final tool, and which round they were included in

Codes included in the final tool.	
Code	Phase at which included.
1. Encouraging engagement in activities.	1
3. Encouraging the development of independence.	1
6. Offering choice.	1
8. Staff initiated non-clinical interactions.	1
9. Positive responses to client initiated interactions.	1
11. Joint involvement in task/activities.	1
14. Seeking input from clients.	1
16. Active listening skills.	1
18. Recognition of communication needs.	1
20. Staff recognising strengths and interests.	1
26. Encouragement/practical support of pro-social activities, including off the unit.	1
28: Empathy shown in response to distress.	2
2. Forcing engagement in activities.	2
4. Non-verbal dismissal.	2
5. Discussing client's personal information (in public settings).	2
10. Negative responses to client initiated interactions.	2
12. Clients not being engaged with.	2
13. Interactions having an "illness focus".	2
15. Not seeking input from clients.	2
17. Staff not being in the moment with clients.	2
19. Ignoring communication needs.	2
24. Set getting up times.	2
30. Asking how a client would like to deal with a problem; exploring options.	2

Appendix 28: Codes excluded in the final tool, and in which round they were excluded

Codes excluded from the final tool.	
Code	Phase at which excluded
7. Tokenistic involvement.	2
21. Punitive interactions.	2
22. Set meal times.	2
23. Set bed times.	2
25. Set smoking times.	2
27. PRN medication offered as a quick fix/alternative to talking.	2
29. Doing with not doing for.	2

Appendix 29: Demographic information from participants involved in ward selection

	Ward 12 (n=10)	Ward 13 (n=8)	Ward 21 (n=9)	Ward 22 (n=7)	Ward 31 (n=6)	Ward 32 (n=7)	Ward 33 (n=10)
Profession							
Nursing	6	4	4	3	1	2	2
Clinical Psychologist	x	x	1	x	x	1	1
Ward Manager	x	1	1	x	x	x	x
Health Care Assistant	2	1	3	2	4	2	3
Occupational Therapist	1	1	x	1	x	x	1
Activity Co-ordinator/Therapies Co-ordinator	x	1	x	1	x	x	x
Clinical Lead	x	x	x	x	1	x	1
Psychiatrist	1	x	x	x	x	x	1
Prefer not to say/did not complete	x	x	x	x	x	2	1
Gender							
Female	8	5	6	3	3	3	5
Male	2	3	3	4	3	2	2
Prefer not to say/Not completed.	x	x	x	x	x	2	3
Length of time at service							
0 – 3 years 12 months	6	2	x	4	5	5	4
4 years – 7years 12 months	4	1	5	3	x	x	2
8 years – 11 years 12 months	x	4	x	x	1	x	1
12 years – 15 years 12 months	x	1	2	x	x	x	x
16 years -19 years 12 months	x	x	x	x	x	x	x
20 years – 23 years 12 months	x	x	x	x	x	x	x
24 years 27 years 12 month	x	x	x	x	x	x	1
28 years – 31 years 12 month	x	x	x	x	x	x	1
32 years – 35 years 12 months	x	x	1	x	x	x	x
Prefer not to say/Not completed	x	x	1	x	x	2	1
Age							
18-30	5	x	2	1	2	1	x

31-45	4	4	2	4	2	2	6
46-60	1	4	5	1	2	1	2
60+	x	x	x	x	x	1	1
Prefer not to say/Not completed	x	x	x	1	x	2	1

Appendix 30 SPSS output Wilcoxon signed rank test for codes which support recovery

Ranks

		N	Mean Rank	Sum of Ranks
Ward 31 – Ward 12	Negative Ranks	5[a]	5.00	25.00
	Positive Ranks	7[b]	7.57	53.00
	Ties	1[c]		
	Total	13		

a. Ward31 < Ward12

b. Ward31 > Ward12

c. Ward31 = Ward12

Test Statistics[a]

	Ward31 - Ward12
Z	-1.105[b]
Asymp. Sig. (2-tailed)	.269

a. Wilcoxon Signed Ranks Test

b. Based on negative ranks.

Appendix 31 SPSS output for Wilcoxon signed rank test for codes which hinder recovery

Ranks

		N	Mean Rank	Sum of Ranks
Ward31 - Ward12	Negative Ranks	4[a]	4.38	17.50
	Positive Ranks	4[b]	4.63	18.50
	Ties	1[c]		
	Total	9		

a. Ward31 < Ward12

b. Ward31 > Ward12

c. Ward31 = Ward12

Test Statistics[a]

	Ward31 - Ward12
Z	-.071[b]
Asymp. Sig. (2-tailed)	.943

a. Wilcoxon Signed Ranks Test

b. Based on negative ranks.

Appendix 32: SPSS output for Chi-square test

Chi-square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-square	.446[a]	1	.504		
Continuity Correction[b]	.286	1	.593		
Likelihood Ratio	.447	1	.504		
Fisher's Exact Test				.593	.297
Linear-by-Linear Association	.445	1	.505		
N of Valid Cases	240				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 44.50.

b. Computed only for a 2x2 table